A Bayesian Framework For Studying Climate Anomalies and Social Conflicts

Ansu Chatterjee

Department of Mathematics and Statistics

University of Maryland Baltimore County

Jan 2, 2025

Co-authors

Ben Bagozzi

Ujjal Mukherjee

Somya Sharma

Outline

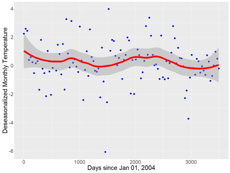
Main issues

- Social science literature is not clear on whether (and how much) is the effect of climate change on various things like migration, political conflicts, supply chain disruption, food-water-energy insecurities.
- A major problem is the sampling bias in the data (people collect data from high conflict areas).
- No uncertainty quantification and scientific hypothesis testing, no accounting for spatio-temporal dependencies.

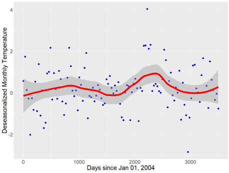
Our data

- Temperature anomalies.
- Conflict event data of various types (collected from news sources, natural language processing and other sources).
- Supply chain data.
- Data aggregated to country-month levels.

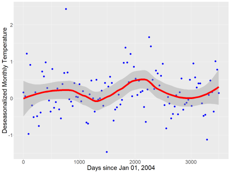
Temperature data (sample)



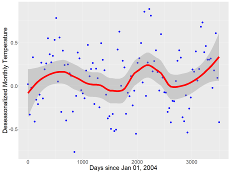
Afghanistan



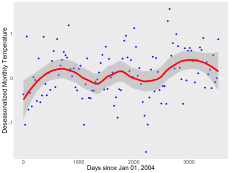
Algeria



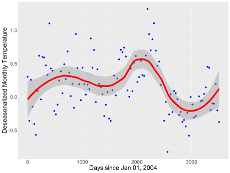
Bangladesh



Brazil



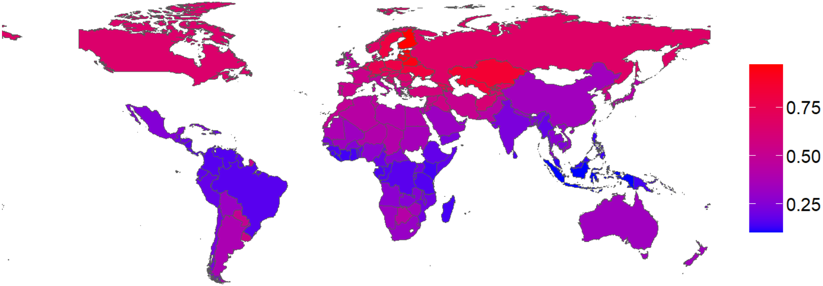
Mexico



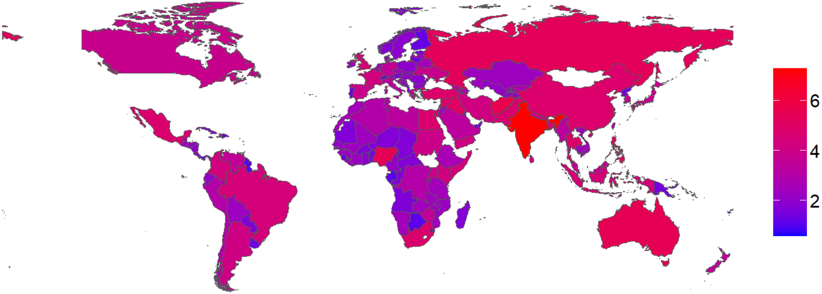
Tanzania

Temperature data $X_{i,t}$ and the seasonality and trend adjusted fits $\hat{X}_{i,t}$ for six representative countries.

Temperature anomalies (absolute)



Log (Average monthly conflict)



Model details

$$
\begin{aligned}
[Y_{i,t} | \mu_{i,t}] &\sim \text{ Negative Binomial}(r_t, p_{i,t}) && (1)\\
&\quad \text{such that } \mathbb{E} Y_{i,t} = \big(1 - p_{i,t}\big)^{-1} p_{i,t} r_t = \mu_{i,t}, \\
\log(\mu_{*,t}) &\sim GP\Big(\eta_{*,t}, K_{\kappa,\phi}\Big) && (2)\\
\eta_{*,t} &= \beta_0 + \Gamma\Big(\log(\mu_{*,t-1}) - \eta_{*,t-1}\Big) + \beta_2 R_{i,t} + \sum_{j=1}^{J_1} \beta_{3,j} Z_{i,t,j} + \sum_{j=1}^{J_2} \beta_{4,j} W_{i,j} && (3)\\
r_t &\overset{i.i.d.}{\sim} Poisson(R), && (4)\\
\theta &= \big(\beta_0, \Gamma, \beta_2, \beta_{3,1}, \ldots, \beta_{3,J_1}, \beta_{4,1}, \ldots, \beta_{4,J_2}, \kappa, \phi, R\big) \sim \pi(\cdot). && (5)
\end{aligned}
$$

Model details (Priors)

$$
\begin{aligned}
\beta_0 &\sim \mathbb{N}(0, \sigma_0^2), \\
\beta_2 &\sim \mathbb{N}(0, \sigma_2^2), \\
\beta_{3,j} &\sim \mathbb{N}(0, \sigma_{j,3}^2)\ \forall j \in \{1, 2, \ldots, J_1\}, \\
\beta_{4,k} &\sim \mathbb{N}(0, \sigma_{k,4}^2)\ \forall k \in \{1, 2, \ldots, J_2\}, \\
{\sigma_0}^{-2}, {\sigma_2}^{-2}, {\sigma_{j,3}}^{-2}, {\sigma_{j,4}}^{-2}, R &\sim \text{Gamma}(4, 4), \\
\gamma_j &\sim \mathbb{N}(0, \tau^2 \lambda_j^2)\ \forall j \in \{1, \ldots, i-1, i+1, \ldots, L\}, \\
\tau &\sim \mathbb{C}^+(0, 1), \\
\lambda_j &\sim \mathbb{C}^+(0, 1).
\end{aligned}
$$

Explanations

- Negative Binomial in (1) owing to the presence of overdispersion in the data.
- The spatial dependencies in the $\{\log(\mu_{i,t})\}$ process is captured through the Gaussian process (GP).
- The mean of this GP has an autoregressive structure and is a function of exogenous variables (including temperature anomaly).
- Autoregression parameter Γ is diagonal, with a sparsity-induced shrinkage using the horseshoe prior (some countries have high level fo conflicts that drag on).
- Matern kernel is used for the GP (other kernels are fine too).

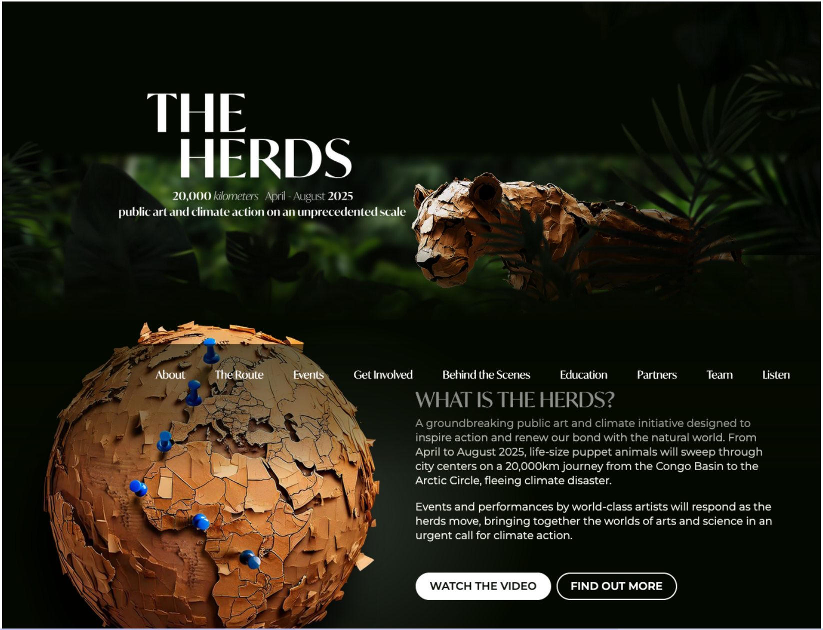
Results on material conflict

Results from MCMC computations when using positive temperature anomalies on the material conflict data.

Variables	Mean	S.D.	Q(2.5%)	Median	Q(97.5%)	$\mathbb{P}$(Par.$\leq$ 0)
(Intercept)	1.18	0.05	1.15	1.15	1.33	0.00
Language	0.38	0.07	0.34	0.35	0.62	0.00
Religion	-0.11	0.02	-0.16	-0.11	-0.11	1.00
Ethnicity	-0.07	0.04	-0.08	-0.08	0.07	0.93
HDI	-0.99	0.06	-1.19	-0.96	-0.96	1.00
Ag. Exports	0.11	0.01	0.09	0.10	0.14	0.00
Manuf.	0.08	0.04	0.01	0.10	0.10	0.02
Commo.	0.12	0.01	0.09	0.13	0.13	0.00
Temp.	0.39	0.03	0.37	0.37	0.47	0.00

What are the limitations?

- Relations between extremes not captured.
- Migration is a missing topic here.
- Separation of temporal and spatial structure.
- GP modeling of the spatial component, which does not scale with data.


THE HERDS
20,000 kilometers April - August 2025
public art and climate action on an unprecedented scale
About
The Route
Events
Get Involved
Behind the Scenes
Education
Partners
Team
Listen
WHAT IS THE HERDS?
A groundbreaking public art and climate initiative designed to inspire action and renew our bond with the natural world. From April to August 2025, life-size puppet animals will sweep through city centers on a 20,000km journey from the Congo Basin to the Arctic Circle, fleeing climate disaster.
Events and performances by world-class artists will respond as the herds move, bringing together the worlds of arts and science in an urgent call for climate action.
WATCH THE VIDEO
FIND OUT MORE

Outline

1. Temperature anomalies, conflicts and supply chains

2. **Hierarchical extremes: Atlantic hurricane modeling**
   - Individual Storm Model and Predictions
   - Seasonal Model and Predictions

3. Capturing complex dependencies: graph neural networks

Motivation

- Tropical cyclones or hurricanes are among the foremost natural phenomena that regularly cause great harm to human communities and infrastructure.
- The relationship between economic loss and a tropical cyclone's size, intensity, storm surge, and other important climatic factors, is complex but important.

Prediction at two levels

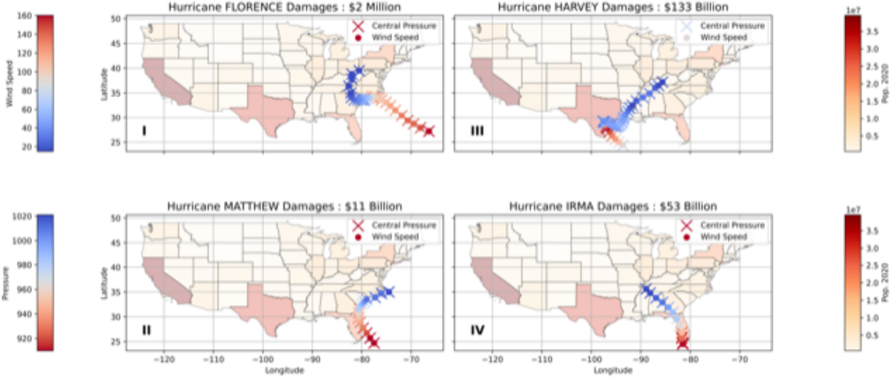
- *Individual predictions.* Predict the probability that a given tropical cyclone may be damage-inflicting, and the amount of damage that it can cause.

- *Seasonal predictions.* Predictive model that forecasts the number of cyclone events that will take place, the probability that a given cyclone will inflict damages, and the monetized value of damages for that season.

Data Sources

1. HURDAT2: Atlantic Tropical Cyclone Attributes Data
   - Available from 1851-2022 but we only use data post 1960.
   - Storm measurements on location (latitude, longitude), maximum winds, central pressure, etc. taken every 6 hours during the storm lifetime
2. ICAT: United States Tropical Cyclone Damage Data
   - Comprehensive damages from 1900-2022 attributed to publicly available sources (Monthly Weather Review)
   - Normalized (2024 $) based on inflation, wealth, and population

Individual Storm Analysis



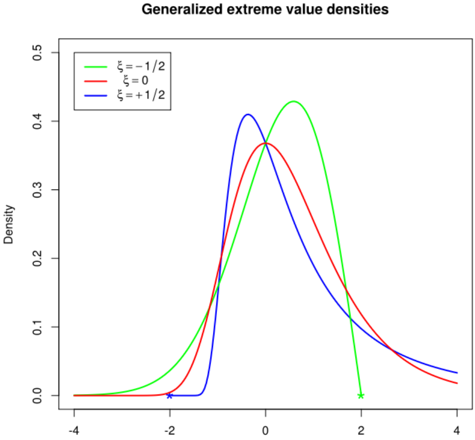
The trajectory of four tropical cyclones in the North Atlantic basin, along with wind-speed (circles), central pressure (crosses), and associated monetary damages (title).

Notations

- Let Z_1 represents log(minCP), Y_1 represents log(maxWS) and Y_2 is for log(damages).
- Let X denote the design matrix with $p = 11$ covariates, namely, average latitude, average longitude, start month, year of occurence, and the May-June averages of following climate indices:
    - Atlantic Multidecadal Oscillation (AMO)
    - North Atlantic Oscillation (NAO)
    - Southern Oscillation Index (SOI)
    - Niño 3.4 anomaly series
    - Atlantic Sea Surface Temperature (SST)
    - July-June Average Sunspot number (SSN)

Generalized Extreme Value distribution

The generalized extreme value (GEV) distribution is a family of continuous probability distributions developed within extreme value theory to combine the Gumbel, Fréchet and Weibull families a.k.a. type I, II and III extreme value distributions



The Hierarchical Model for extremes

The hierarchical model. We model each of the variables using a *generalized extreme value distribution* (GEV).

$$Z_1|X \sim \text{GEV}(\mu_{z_1}(X), \sigma_{z_1}, \xi_{z_1}) \tag{6}$$
$$Y_1|(Z_1, X) \sim \text{GEV}(\mu_{y_1}(Z_1, X), \sigma_{y_1}, \xi_{y_1}) \tag{7}$$
$$Y_2|(Z_1, Y_1, X) \sim \text{GEV}(\mu_{y_2}(Z_1, Y_1, X), \sigma_{y_2}, \xi_{y_2}), \tag{8}$$

where, for $\alpha \in \mathbb{R}^p, \beta \in \mathbb{R}^{p+1}$ and $\gamma \in \mathbb{R}^{p+2}$.

$$\mu_{z_1}(X) = X\alpha$$
$$\mu_{y_1}(Z_1, X) = [Z_1, X]\beta$$
$$\mu_{y_2}(Z_1, Y_1, X) = [Z_1, Y_1, X]\gamma.$$

Posterior Sampling for the Hierarchical Bayesian model

The joint density can then be written as,

$$
\begin{aligned}
f(Z_1, Y_1, Y_2|X,\theta) &= f(Y_2|Y_1, Z_1, X)f(Y_1|Z_1, X)f(Z_1|X) \\
&= \frac{1}{\sigma_{Y_2}}(t(Y_2))^{\xi_{Y_2}+1}\exp(-t(Y_2))\frac{1}{\sigma_{Y_1}}(t(Y_1))^{\xi_{Y_1}+1} \\
&\quad \times \exp(-t(Y_1))\frac{1}{\sigma_{Z_1}}(t(z_1))^{\xi_{z_1}+1}\exp(-t(Z_1)),
\end{aligned}
$$

where,

$$
t(x) = \begin{cases} (1+\xi(\frac{x-\mu}{\sigma}))^{-1/\xi} & \xi \neq 0 \\ \exp\left(-\frac{x-\mu}{\sigma}\right) & \xi = 0. \end{cases}
$$

Posterior Sampling for the Hierarchical Bayesian model

The joint density can then be written as,

$$\begin{aligned}
f(Z_1, Y_1, Y_2|X,\theta) &= f(Y_2|Y_1,Z_1,X)f(Y_1|Z_1,X)f(Z_1|X) \\
&= \frac{1}{\sigma_{Y_2}}(t(Y_2))^{\xi_{Y_2}+1}\exp(-t(Y_2))\frac{1}{\sigma_{Y_1}}(t(Y_1))^{\xi_{Y_1}+1} \\
&\quad \times \exp(-t(Y_1))\frac{1}{\sigma_{z_1}}(t(z_1))^{\xi_{z_1}+1}\exp(-t(Z_1)),
\end{aligned}$$

where,

$$t(x) = \begin{cases} (1+\xi(\frac{x-\mu}{\sigma}))^{-1/\xi} & \xi \neq 0 \\ \exp\left(-\frac{x-\mu}{\sigma}\right) & \xi = 0. \end{cases}$$

- No closed form conditional densities, use the Metropolis Hastings sampling scheme (MCMC) for $N = 10^6$ steps with step-sizes chosen to achieve 20% acceptance rate.

Priors

$$
\begin{aligned}
\alpha_i &\overset{iid}{\sim} \mathrm{N}(0, 10^2), i = 1, \ldots, p \\
\beta_i &\overset{iid}{\sim} \mathrm{N}(0, 10^3), i = 1, \ldots, p+1 \\
\gamma_i &\overset{iid}{\sim} \mathrm{N}(0, 10^2), i = 1, \ldots, p+2 \\
\sigma_{Z_1}, \sigma_{Y_1}, \sigma_{Y_2} &\overset{iid}{\sim} \mathrm{IG}(\alpha = 1, \beta = 3) \\
\xi_{Z_1} &\sim \mathrm{Unif}(-1, 1) \\
\xi_{Y_1}, \xi_{Y_2} &\overset{iid}{\sim} \mathrm{Unif}(-0.55, 0.5).
\end{aligned}
$$

Selected variables

Variables	Min CP	Max WS	Damages
Intercept	0.0000	0.0000	0.0000
Min CP (scaled)	NA	0.0000	0.0177
Max WS (scaled)	NA	NA	0.2479
Avg. Latitude	0.0000	0.8318	0.2290
Avg. Longitude	0.0009	0.9698	0.9965
StartMonth	0.3456	0.9949	0.3351
Year	0.1122	0.9620	0.5877
NAO	0.9956	0.9973	0.9991
SOI	0.9912	1.0000	0.4244
AMO	0.4748	0.9967	0.5665
ANOM.3.4	0.9513	0.9974	0.0486
Atl_SST	0.0021	0.9996	0.0091
Sunspots	0.9461	0.9830	0.9998
ξ	0.0000	0.0000	0.0018
σ	0.0000	0.0000	0.0000

Data depth to measure how relevant a variable is in each of the layers of the model.

Posterior means

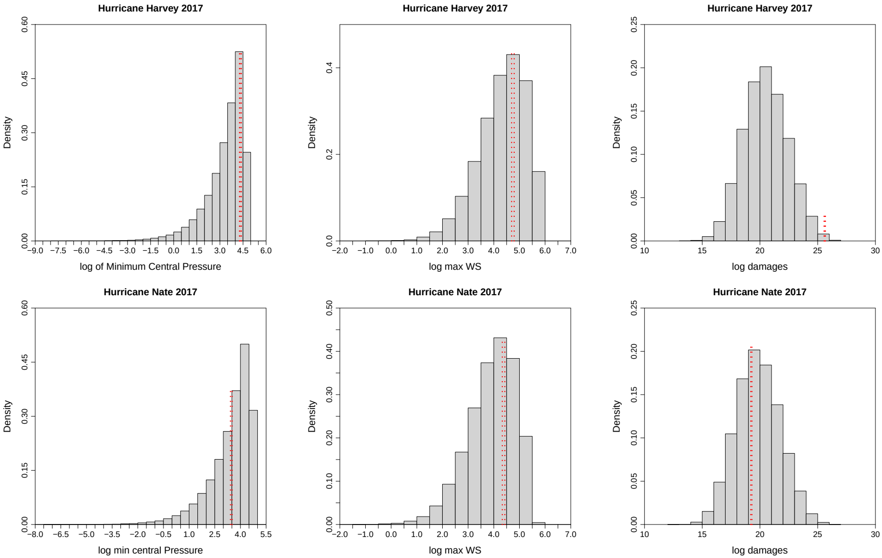
	Posterior means	Posterior standard deviation
Min CP		
Intercept	3.1797	0.0767
Avg. Latitude	-0.2251	0.0483
Avg. Longitude	-0.1257	0.0449
Start month	0.0890	0.0643
Year	-0.0981	0.0465
AMO	-0.1145	0.0662
Atlantic SST	0.2508	0.0484
ξ	-0.9331	0.0470
σ	1.2931	0.0659
Max WS		
Intercept	3.7737	0.0900
Min CP (scaled)	0.3696	0.0792
ξ	-0.5403	0.0109
σ	1.0166	0.0516
Damages		
Intercept	19.4191	0.1867
Max WS (scaled)	0.8941	0.2968
Min CP (scaled)	0.6707	0.2853
Avg. Latitude	-0.3315	0.1650
Start month	-0.2466	0.1687
Year	0.1405	0.1954
SOI	0.1912	0.2044
AMO	0.2318	0.2503
ANOM.3.4	0.4064	0.1906
Atlantic SST	-0.7354	0.2391
ξ	-0.2618	0.0380
σ	1.8824	0.0756

Posterior mean and standard deviation estimates for models (6),(7), (8) for the fully hierarchical Bayesian Generalized Extreme Value (GEV) model with selected covariates.

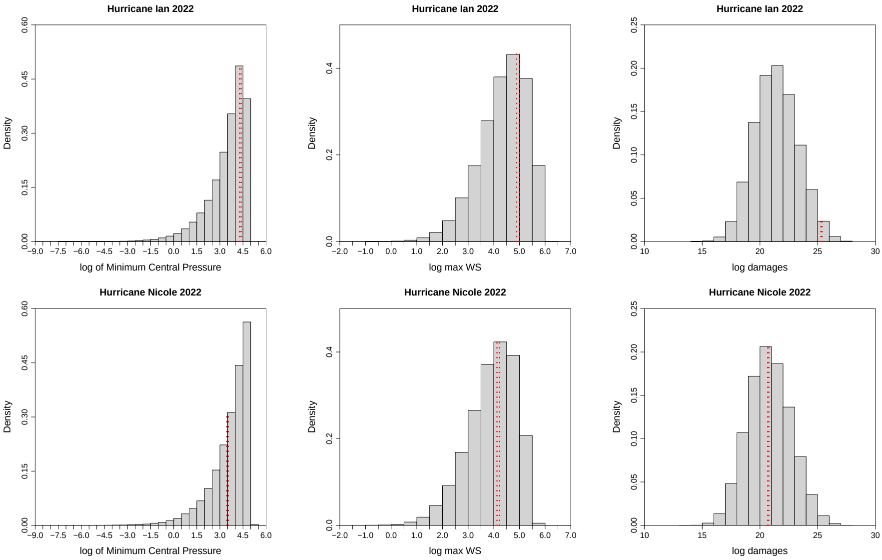
Model interpretations

- Shape (ξ) and scale (σ) parameters highly significant.
- Statistically significant variables
  1. For minCP: Average latitude, average longitude, AMO, and Atlantic SST.
  2. For maxWS: minCP.
  3. for log(damages): maximum wind speed, minimum central pressure, average latitude, ANOM 3.4, and Atlantic SST.

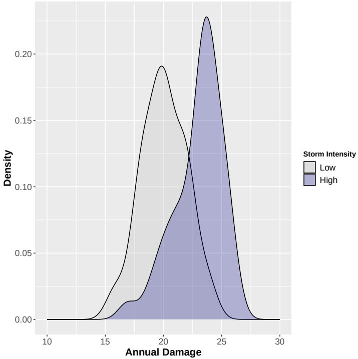
Posterior Predictive Distributions for Storms in 2017:



Posterior Predictive Distributions for Storms in 2022:

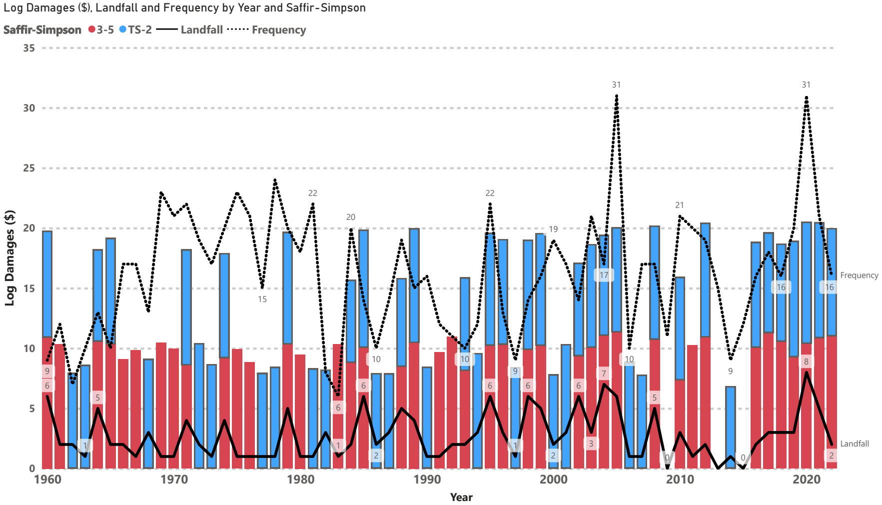


Bimodal nature of damages by the storm intensity



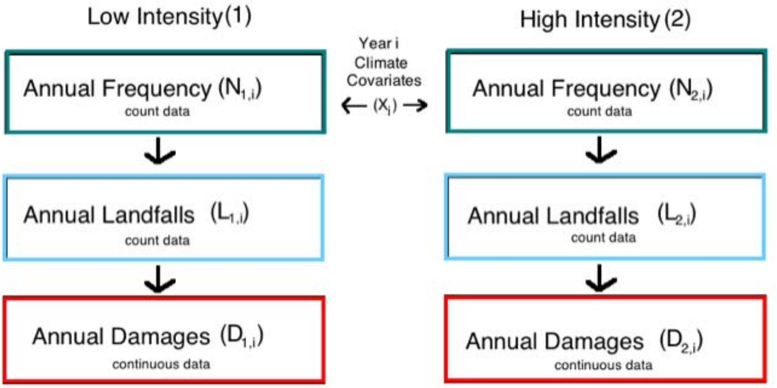
Estimated Densities of Annual Log Damage in US dollars for Cyclones in Low and High-Intensity Categories illustrate a bimodality, with high-intensity storms expected to cause greater damages.

Seasonal damages by landfalls and frequency of occurence



Seasonal predictions of damages as functions of annual landfalls and frequency of occurrence by storm intensity.

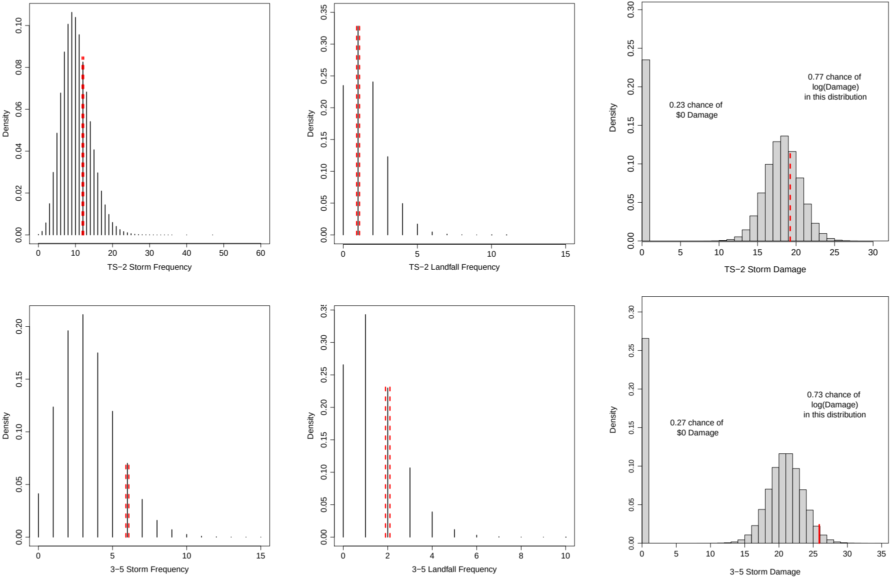
Our approach: Bayesian Hierarchical Model



Separately model *low intensity* (tropical storm to Category 2) and *high intensity* (Category 3-5) tropical cyclones based on their clearly distinguished differences in frequency, landfall probability, and damage capacity.

Model Summary

- NAO, SOI, Sunspots coefficients not statistically significant
- Damages
  - Low intensity: damage causing probability is 13.5%
  - High intensity: damage causing probability is 43.4%
- Damage
  - Low intensity damages: 507.7 million dollars/year (292.8, 877.2).
  - High intensity damages: 9.76 billion dollars/year (5.17, 18.43).
- U.S. should budget for 12.597 billion dollars in annual damage.



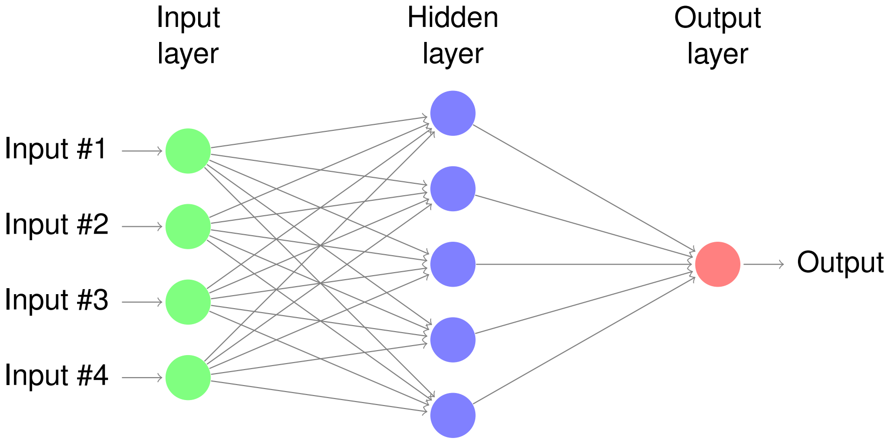
Posterior Predictive Distributions for 2017 tropical cyclones.

Prediction at two levels

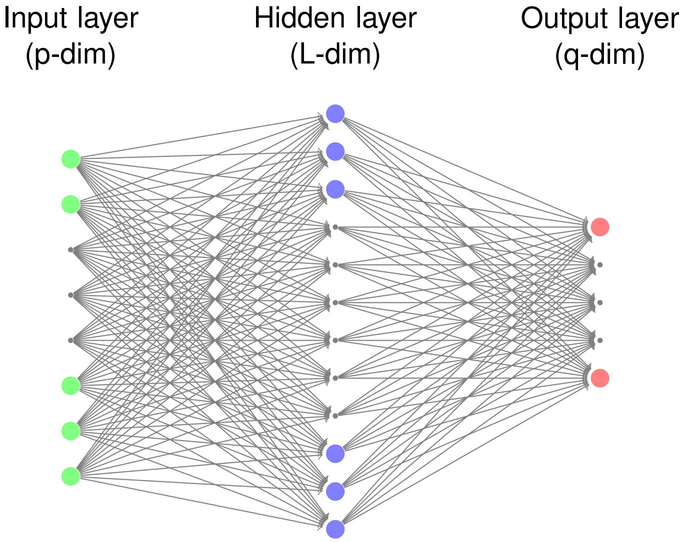
- Currently, no significant migration patterns due to Atlantic tropical storms.
- Currently, no significant social conflict attributable to these storms.
- However, significant effect on insurance and reinsurance industry.
- The models do not control social/political networks and other dependencies.

Outline

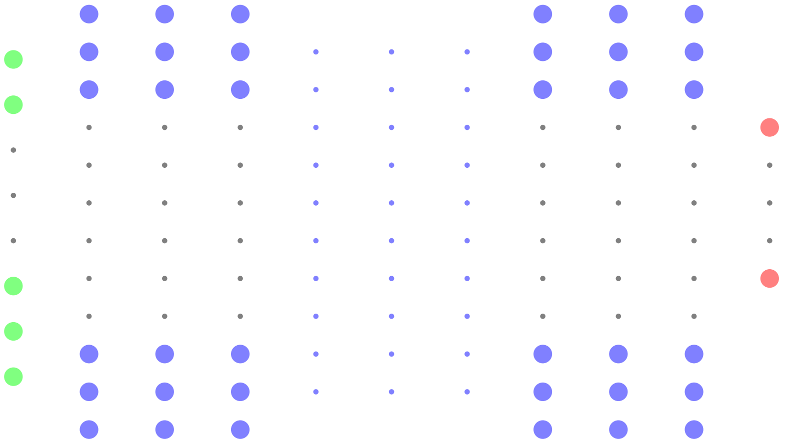
MLP in 1K words



ANN in 1K words



This is the *classical artificial neural network* (ANN)



This is the *deep artificial neural network* (DANN, DNN)

DNN: Architecture

- Generic observation $(X \equiv X_0, Y) \in \mathbb{R}^{p_0} \times \mathbb{R}$.
- First, we obtain a linear transformation of the previous layer:

$$H_\ell = W_\ell X_{\ell-1} + b_\ell \in \mathbb{R}^{p_\ell}, \ell = 1, \ldots, L.$$

- Then, we use an ***activation function*** to transform H_ℓ to the ℓ-th layer output X_ℓ, $\ell = 1, \ldots, L$.

Example (Activation functions)

$$A(x) = \left[1 + \exp\{-x\}\right]^{-1}, \text{(sigmoid)},$$
$$A(x) = \tanh(x), \text{(hyperbolic tangent)},$$
$$A(x) = \max\left(0, x\right), \text{(ReLU)}.$$

Linear transform on ℓ-th layer, i-th observation:

$$H_{\ell,i} = W_\ell X_{\ell-1,i} + b_\ell \in \mathbb{R}^{p_\ell},\ \ell = 1, \ldots, L,\ \ i = 1, \ldots, n.$$

The ℓ-th layer, i-th observation, j-th element:

$$\begin{aligned} X_{\ell,i,j} &= A\big(H_{\ell,i,j}\big), \\ &\equiv A\Big(\sum_k W_{\ell,j,k} X_{\ell-1,i,k} + b_{\ell,j}\Big), \\ &\qquad \ell = 1, \ldots, L,\ \ i = 1, \ldots, n,\ \ j = 1, \ldots, p_\ell. \end{aligned}$$

Graph NN: Structure notations

DNN

The ℓ-th layer, i-th observation, j-th element:

$$X_{\ell,i,j} = A\Big(\sum_k W_{\ell,j,k} X_{\ell-1,i,k} + b_{\ell,j}\Big).$$

Graph Neural Network (one version)

The ℓ-th layer, i-th observation, j-th element:

$$X_{\ell,i,j} = A\Big(\sum_k W_{\ell,j,k} X_{\ell-1,i,k} + b_{\ell,j} + \sum_{u \in N(i)} \sum_k U_{\ell,j,k} X_{\ell-1,u,k}\Big).$$

Here, $N(i)$ is the set of neighbors of the i-th node.

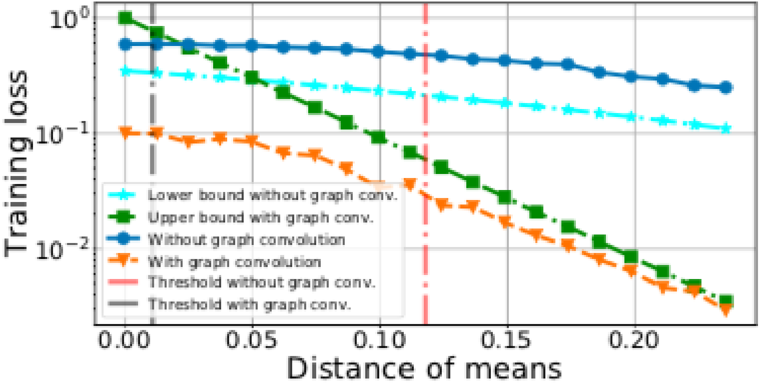
Contextual Stochastic Block Model (CSBM)

- Assume two classes: $Y \in \mathcal{Y} = \{-1, 1\}$, with $\mathbb{P}[Y = 1] = 1/2 = \pi_1$.
- A network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$ is constructed as follows:
  1. Nodes set $\mathcal{V}$: $|\mathcal{V}| = n$, and Y_i's iid for each node according to the above distribution, $i = 1, \ldots, n$.
  2. Given Y_i, the attributes X_i come from f_1 or f_{-1}, two densities.
  3. Nodes (i, j) are connected in $\mathcal{E}$ with probability p if $Y_i = Y_j$ and with probability q if $Y_i \neq Y_j$.
  4. If $p > q$ ($p < q$) we have homophily (heterophily).

Graph Convolution Network

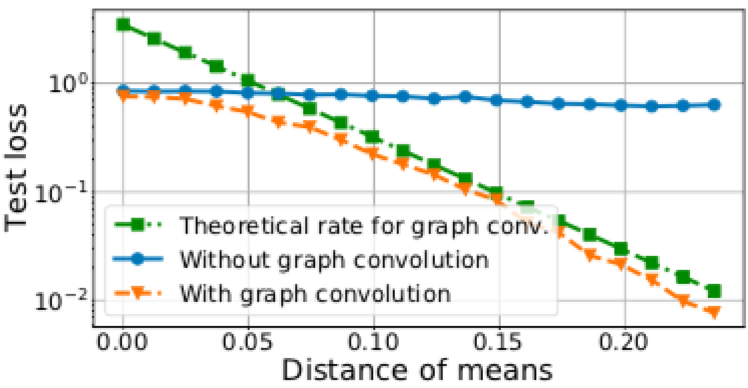
- A Graph Convolution network (GCN) with ReLU activation build on the above CSBM achieves Bayes optimality (result due to Wei *et al.* (2022 NeurIPS)).
- Computations are challenging, but some approximations are known in the literature.

GCN Examples-1: training error



(a) Training loss vs distance of means

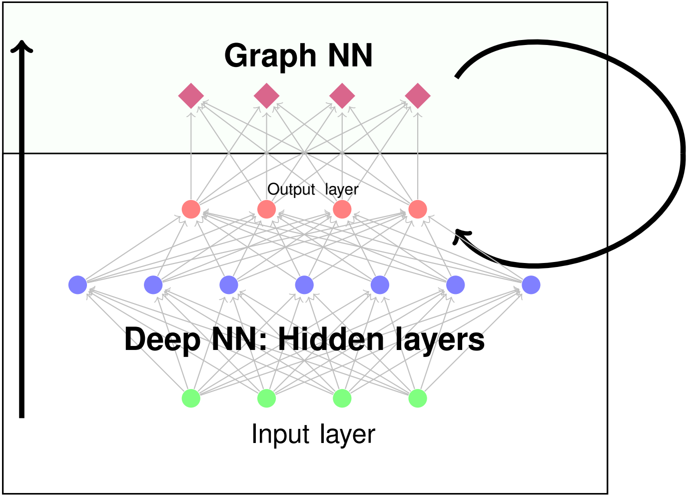
GCN Examples-1: test error



(b) Test loss vs distance of means

Graph Convolution Network: challenges

- Still no uncertainty quantification: the above version only produces Bayes estimates.
- Shoehorns the nonlinear mapping from covariate/feature to response and the (potentially linear) smoothing over graph vertices in the same framework.
- No discovery low-dimensional structures.


Graph NN
Output layer
Deep NN: Hidden layers
Input layer

Proposed architecture: advantages

- Bayesian uncertainty quantification: fairly complex modeling and computation, *work is ongoing (early results are very promising)*.
- Separates the nonlinear mapping from covariate/feature to response and the (potentially linear) smoothing over graph vertices.
- Can exploit (known) spatio-temporal structures.
- Computationally scalable and more versatile than standard GP-based approaches.

Acknowledgment:

This research is partially supported by the National Science Foundation (NSF) under grants # DMS-2436549.

Thank you