Multi-hazard event set generation using deep learning and statistics of extremes

Alison Peard

OPSIS Weekly Meeting | 5th March 2025

Oxford Programme for Sustainable Infrastructure Systems
UNIVERSITY OF OXFORD

Agenda

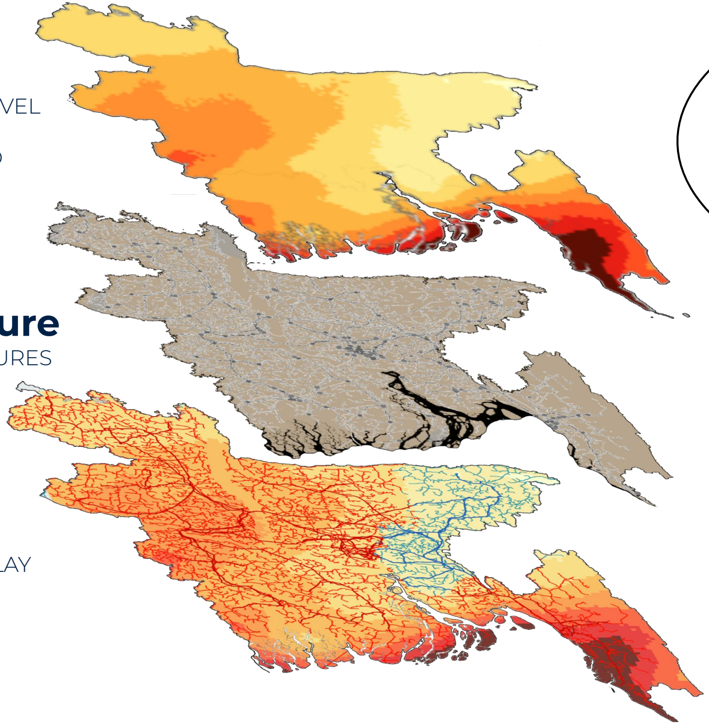
- Refine narrative
- Ground in OPSIS context
- Discuss ideas for applications with OPSIS team (🤞)
- Get feedback on:
  - Figures & results
  - General
    - Anything good
    - Anything unclear
    - Too simple / too complex

Introduction Risk Analysis with Return Level Maps

Hazard
N-YEAR RETURN LEVEL
EPOCH
CLIMATE SCENARIO

Infrastructure
LOCATIONS & FEATURES

Exposure
GEOSPATIAL OVERLAY



+ Vulnerability Curves

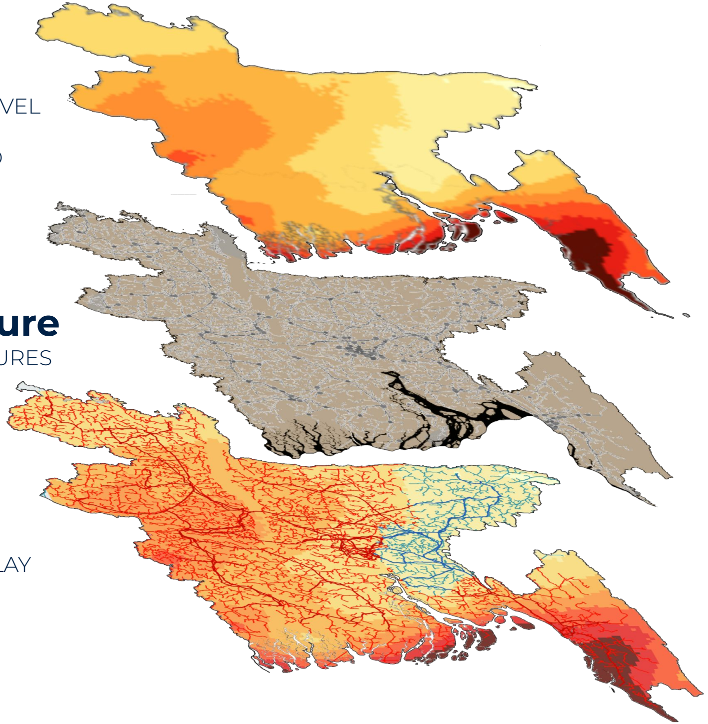
Can you summarise losses over ...? We would like just a single number.

Introduction Risk Analysis with Return Level Maps

Hazard
N-YEAR RETURN LEVEL
EPOCH
CLIMATE SCENARIO

Infrastructure
LOCATIONS & FEATURES

Exposure
GEOSPATIAL OVERLAY



+ Vulnerability Curves

What can you say?

- Risk to a single asset ✅
- Compare single assets ✅

What can you not say?

- Risk across a region/sector/network. ❌
- Risk from multiple hazards co-occuring. ❌

Why?

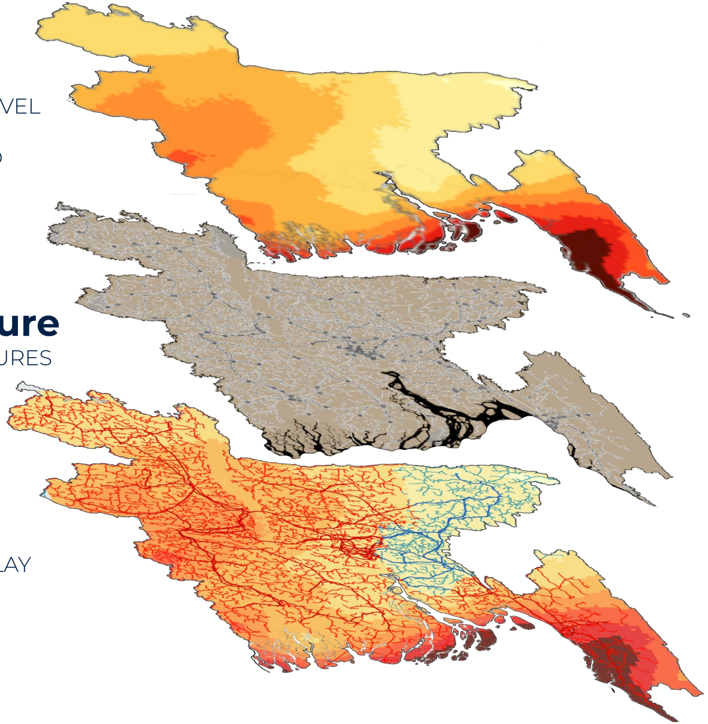
- Hazard maps ≠ realistic hazard events
- Hazard maps are univariate

Introduction Risk Analysis with Return Level Maps

Hazard
N-YEAR RETURN LEVEL
EPOCH
CLIMATE SCENARIO

Infrastructure
LOCATIONS & FEATURES

Exposure
GEOSPATIAL OVERLAY



+ Vulnerability Curves

What can you say?
- Risk to a single asset ✅
- Compare single assets ✅

What can you not say?
- Risk across a region/sector/network. ❌
- Risk from multiple hazards co-occurring. ❌

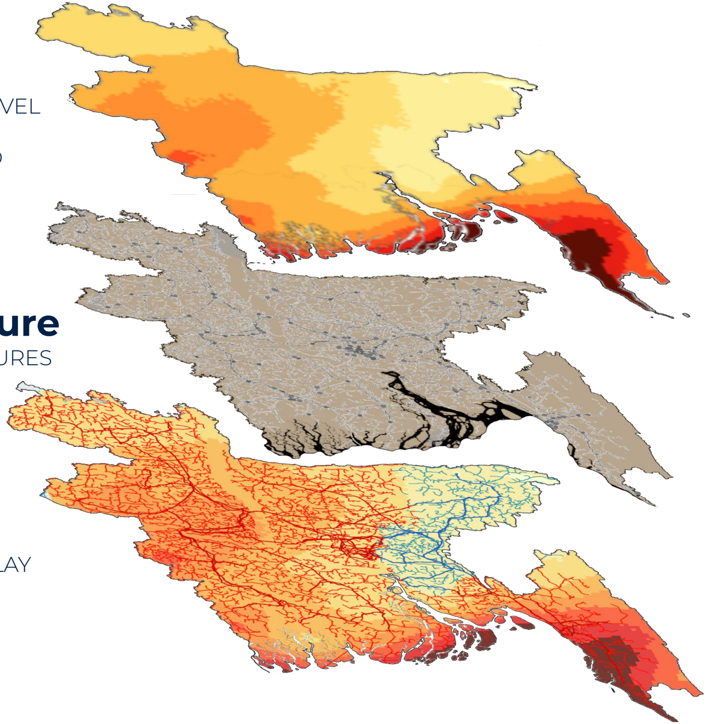
Why?
- Hazard maps ≠ realistic hazard events
- Hazard maps are univariate

Introduction Risk Analysis with Return Level Maps

Hazard
N-YEAR RETURN LEVEL
EPOCH
CLIMATE SCENARIO

Infrastructure
LOCATIONS & FEATURES

Exposure
GEOSPATIAL OVERLAY



+ Vulnerability Curves

What can you say?

- Risk to a single asset ✅
- Compare single assets ✅

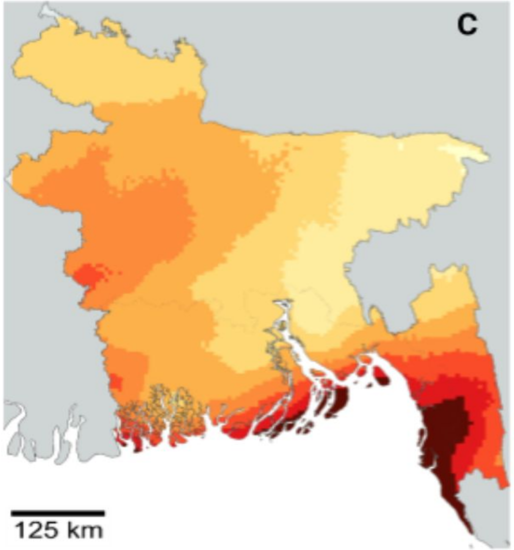
What can you not say?

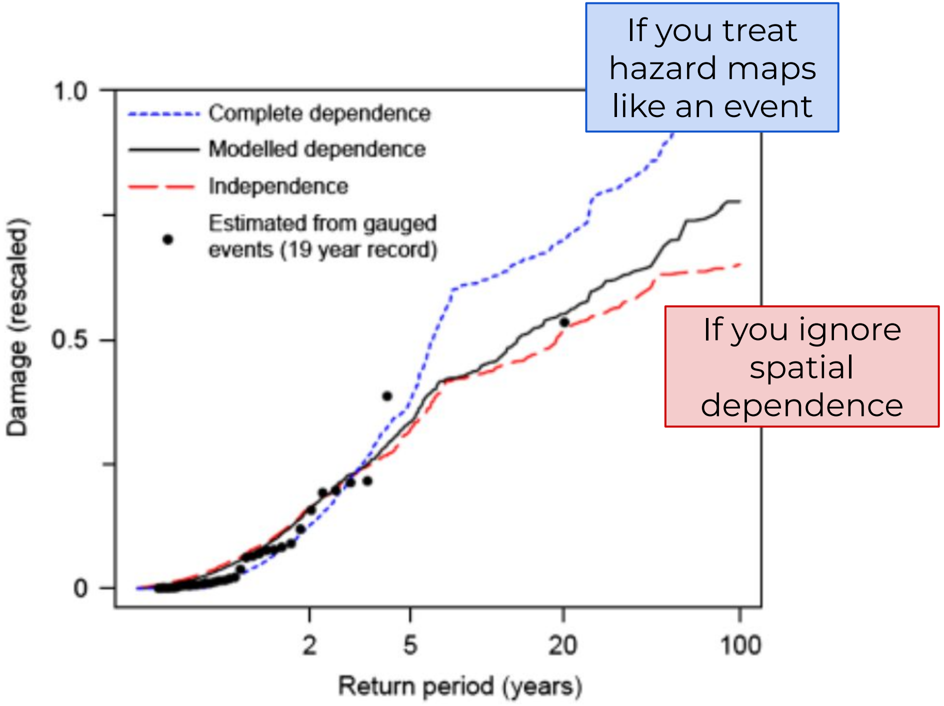
- Risk across a region/sector/network. ❌
- Risk from multiple hazards co-occurring. ❌

Why?

- Hazard maps ≠ realistic hazard events
- Hazard maps are univariate

Introduction Return Level Maps are not Events





1-in-50 year return period map shows

- 50-year wind speed for each point ✅
- 50-year wind storm ❌

1. [ArcGIS | What does expected annual damage mean?]
2. Lamb, Tawn, & Keef (2010) *A new method to assess the risk of local and widespread flooding on rivers and coasts*

Introduction Return Level Maps are Univariate

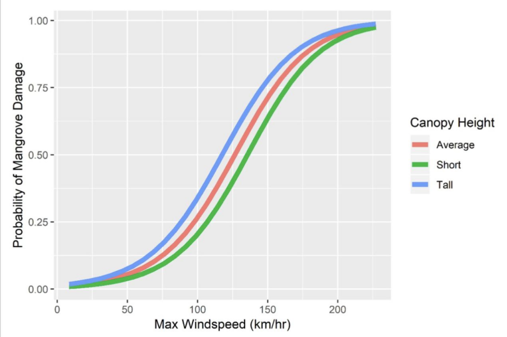
Compound hazards

Two or more hazards may impact the same region and/or time period with impacts different (greater, lesser) than their sum.

- Types
- Times
- Locations

Interaction of hazards

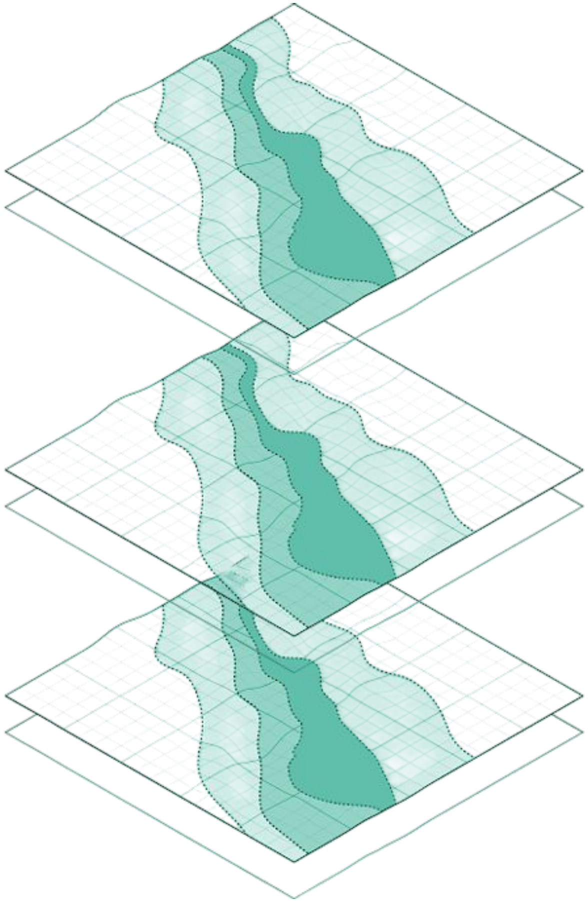
- None → “Multi-layer single hazards”
- Hazard level → Multi-hazard
- Vulnerability level → Multi-risk



	Estimate	95% CI
Intercept	−1.65	−1.1, −2.22
Max wind	1.91	2.24, 1.59
Max wind^2	0.02	0.22, −0.18
Rain	0.39	0.48, 0.29
StormHist1[a]	−0.88	−0.25, −1.49
StormHist2[b]	−0.35	0.29, −0.97
Canopy height	0.35	0.51, 0.19
Wind duration	0.12	0.33, −0.09

1. MYRIAD-EY *D1.2 Handbook of Multi-hazard, Multi-Risk Definitions and Concepts*
2. Taillie (2021) Widespread mangrove damage resulting from the 2017 Atlantic mega hurricane season

Introduction Spatially-Resolved Event Sets (Cat Models)



Features
Portfolio of 1000s of events
Modelling is expensive
Numerical, stochastic, or hybrid simulations
Calculate statistics over losses from entire portfolio
Multivariate or region events

Introduction Existing Event Sets

Tropical cyclones

- Bloemendaal et al. (2020) | STORM

Floods

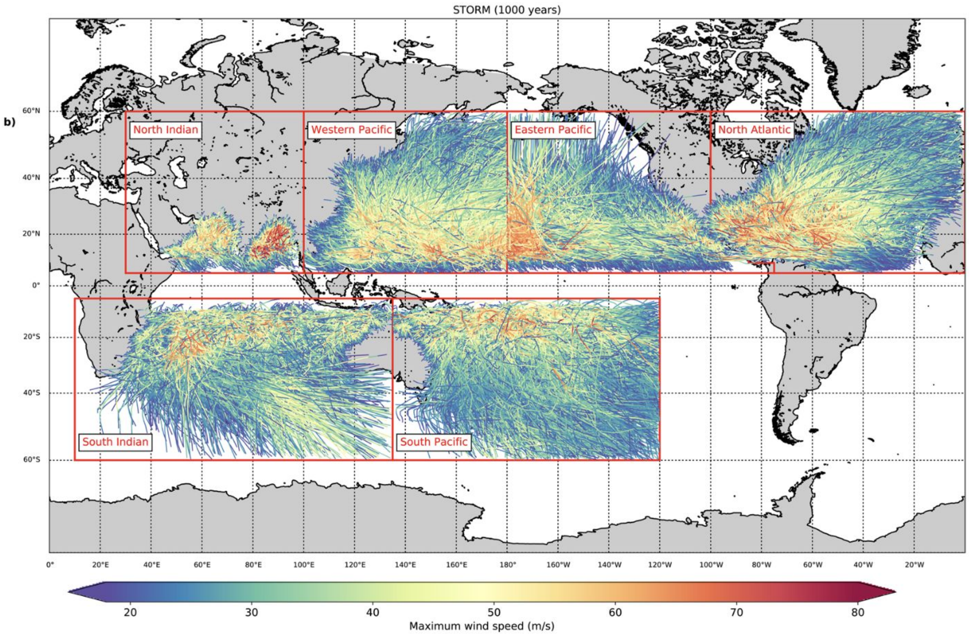
- JBA | Global Flood Event set
- Fathom | Global Flood Cat

1. Bloemendaal et al. (2020) *Generation of a global synthetic tropical cyclone hazard dataset using STORM*

Introduction STORM

STORM

- Based on IBTrACS
- Statistical-empirical method
- Key variables
    - Wind speed **Vt**
    - Pressure **P**
    - Maximum Potential Intensity **MPI**
    - Sea surface temperature **SST**



Simulate track

Genesis points → 18 ms-1 wind → Initial pressure → Simulate pressure → Simulate winds

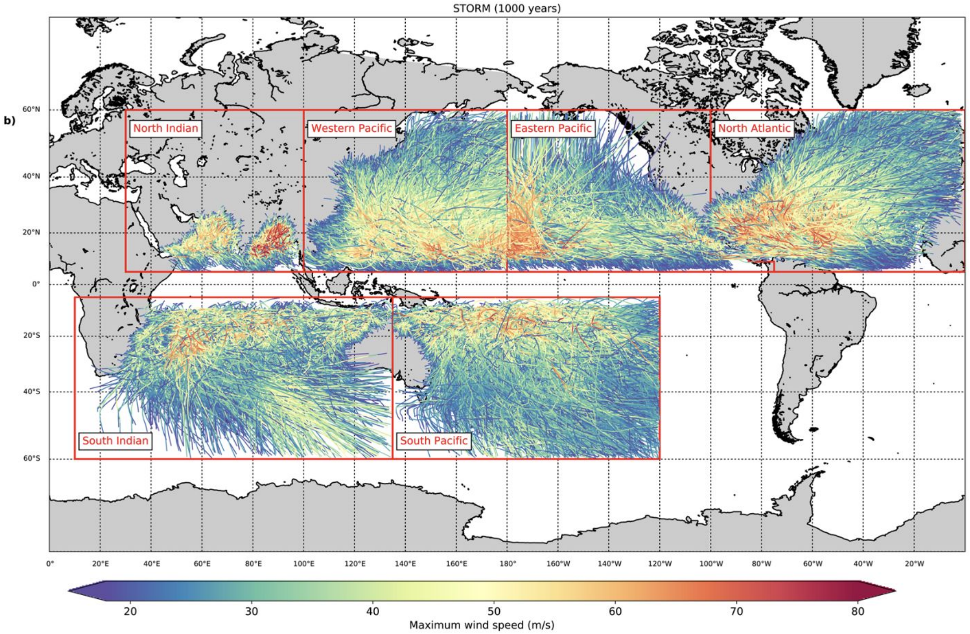
stochastic **empirical** **stochastic** **empirical**

1. Bloemendaal et al. (2020) *Generation of a global synthetic tropical cyclone hazard dataset using STORM*

Introduction STORM

STORM

- Based on IBTrACS
- Statistical-empirical method
- Key variables
  - Wind speed **Vt**
  - Pressure **P**
  - Maximum Potential Intensity **MPI**
  - Sea surface temperature **SST**



$$V_t = a(P_{env} - P_t)^b$$

$$P_{env} - P = A + Be^{C(SST - T_0)}, \quad T_0 = 30.0\ ^\circ\mathrm{C}$$

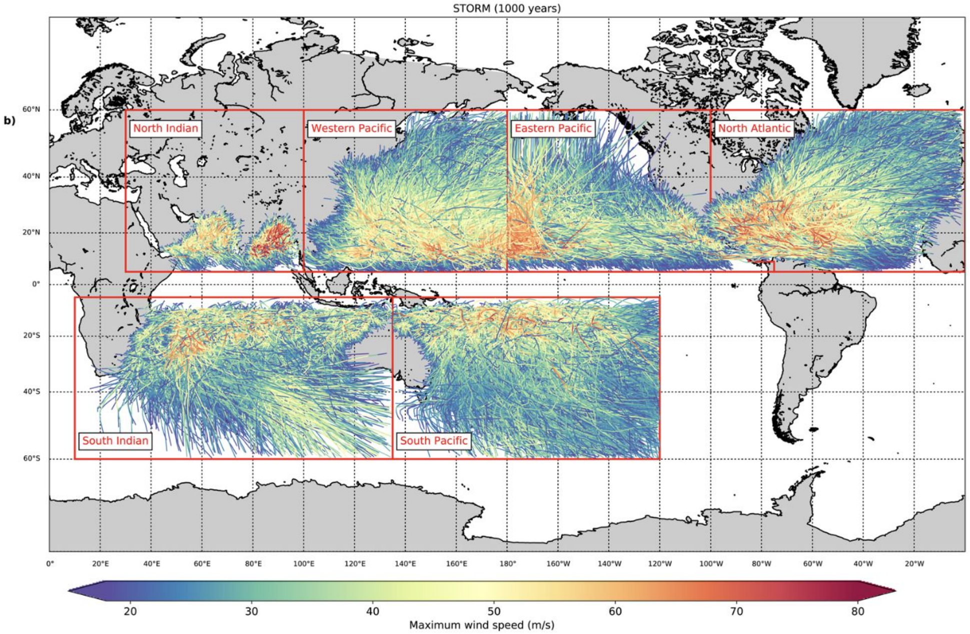
$$\Delta P_t = c_0 + c_1 \Delta P_{t-1} + c_2 e^{-c_3 X}, \qquad c_2 > 0,\ X = \max\{0,\ P_t - MPI\}$$

1. Bloemendaal et al. (2020) *Generation of a global synthetic tropical cyclone hazard dataset using STORM*

Introduction STORM

STORM

- Based on IBTrACS
- Autoregressive
- Key variables
  - Wind speed **Vt**
  - Pressure **P**
  - Maximum Potential Intensity **MPI**
  - Sea surface temperature **SST**

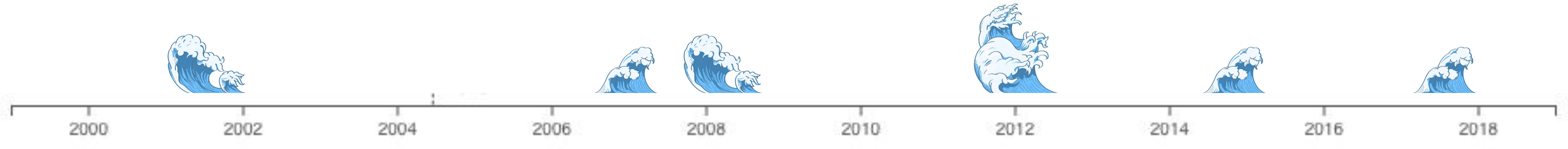


Wind	=	**f(Pressure)**
MPI	=	**f(SST)**
ΔPressure	=	**Autoregression with constraint f(MPI)**

1. Bloemendaal et al. (2020) *Generation of a global synthetic tropical cyclone hazard dataset using STORM*

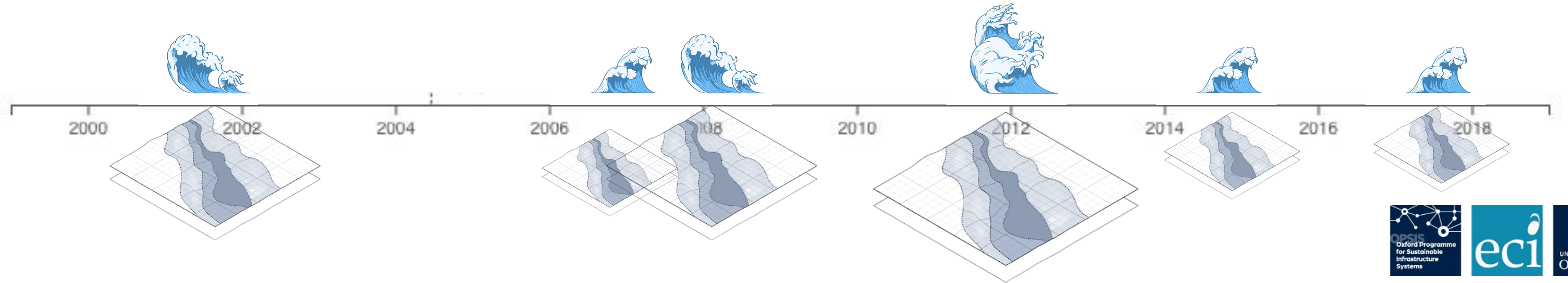
Method Overview

- Identify independent hazard events
- Extract footprints
- Estimate
  - Marginal distributions
  - Dependence structure (fields **and** space)
- Generate new samples



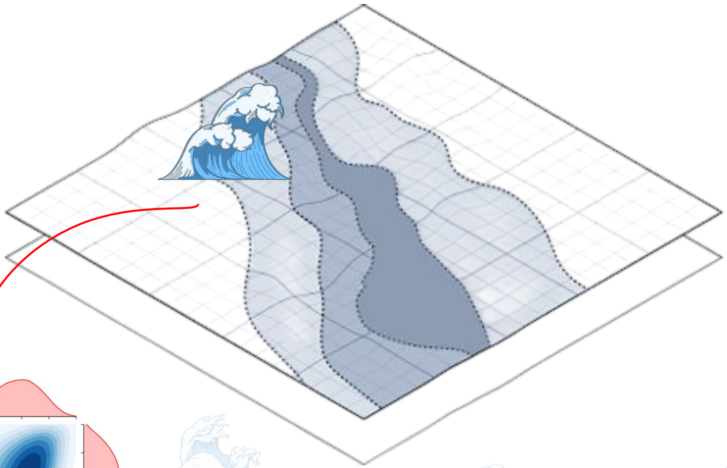
Method Overview

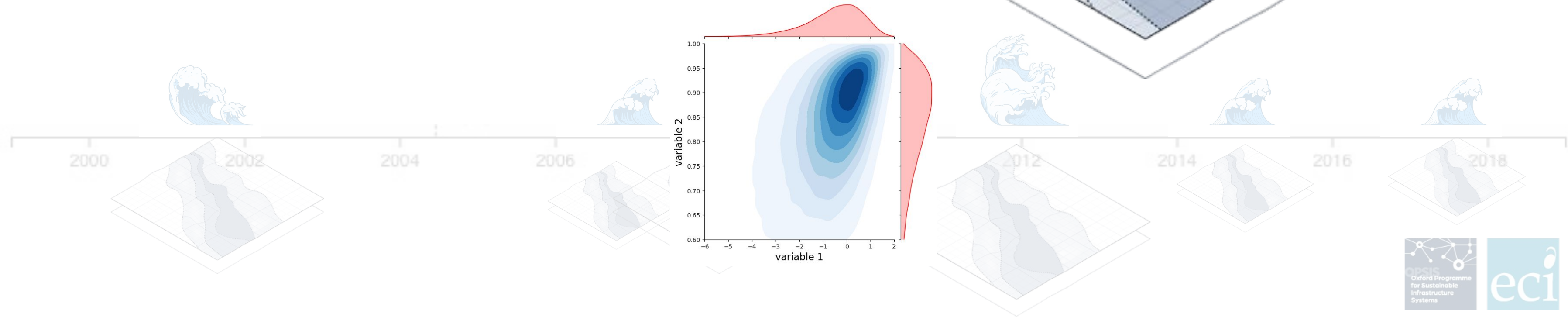
- Identify independent hazard events
- Extract footprints
- Estimate
  - Marginal distributions
  - Dependence structure (fields **and** space)
- Generate new samples



Method Overview

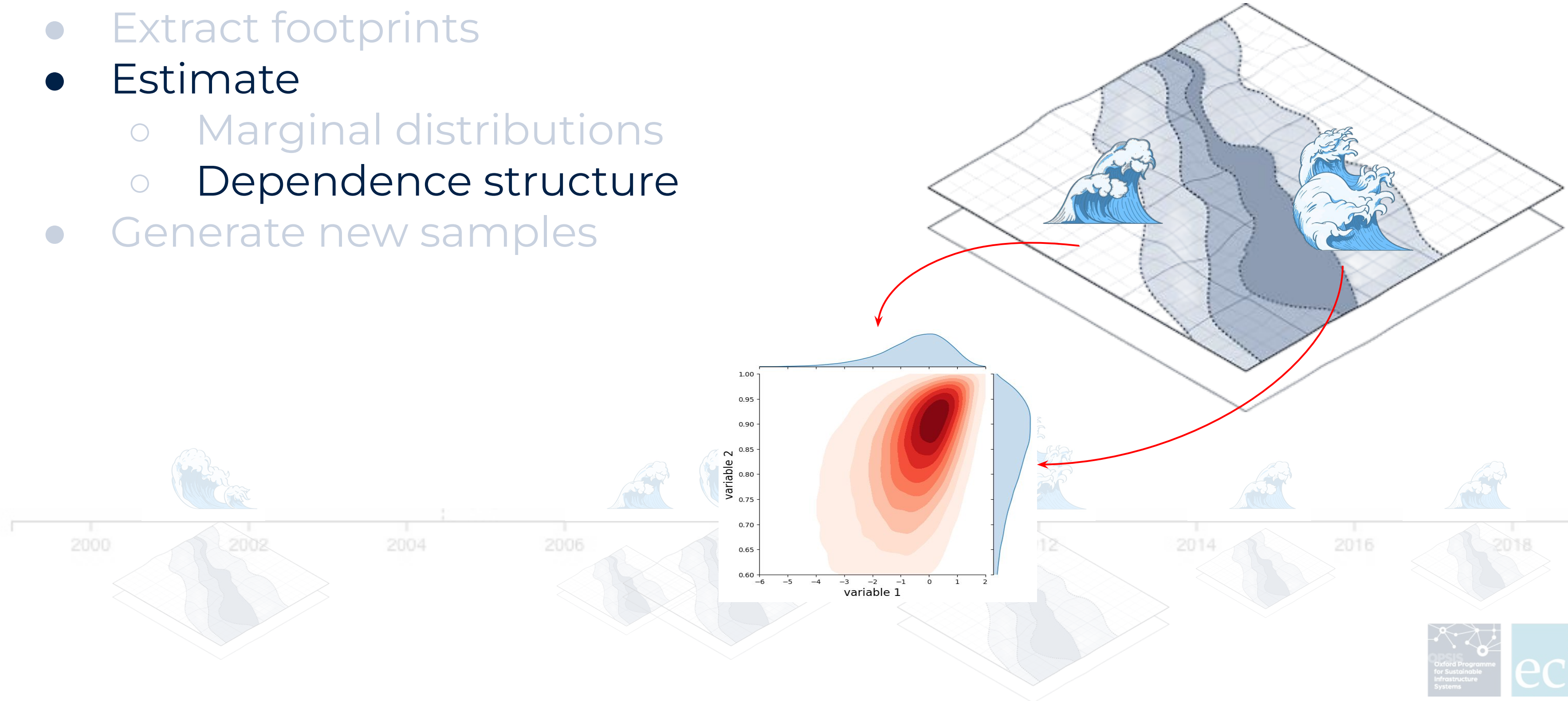
- Identify independent hazard events
- Extract footprints
- Estimate
    - Marginal distributions
    - Dependence structure
- Generate new samples





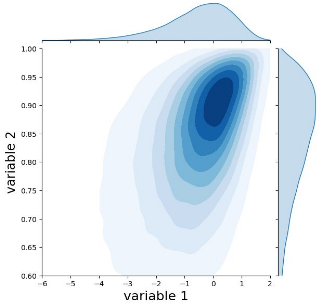
Method Overview

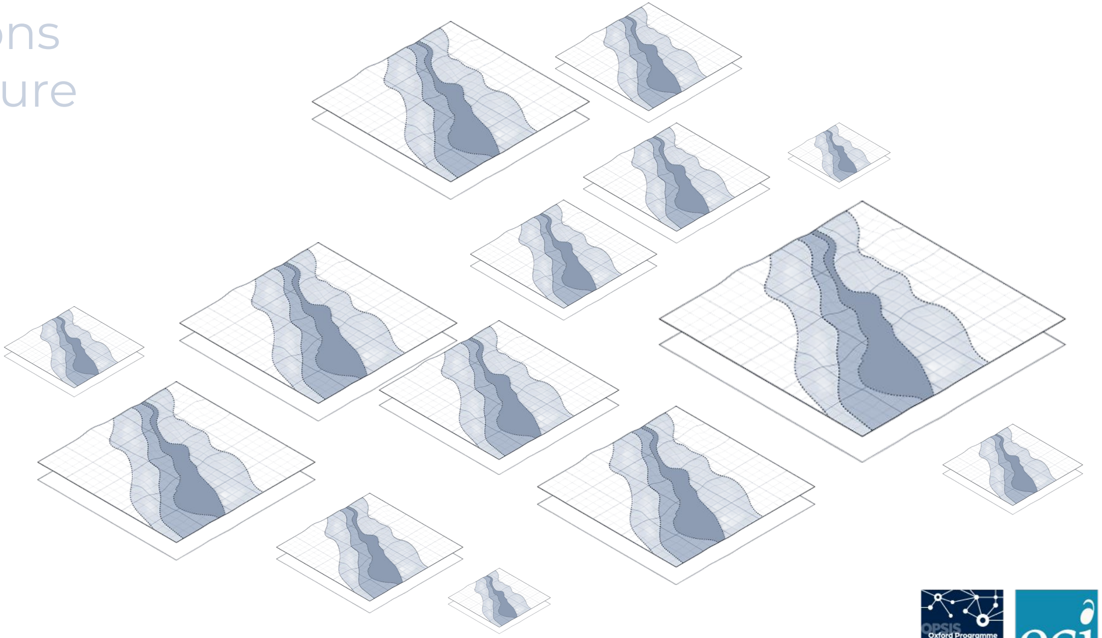
- Identify independent hazard events
- Extract footprints
- Estimate
  - Marginal distributions
  - Dependence structure
- Generate new samples



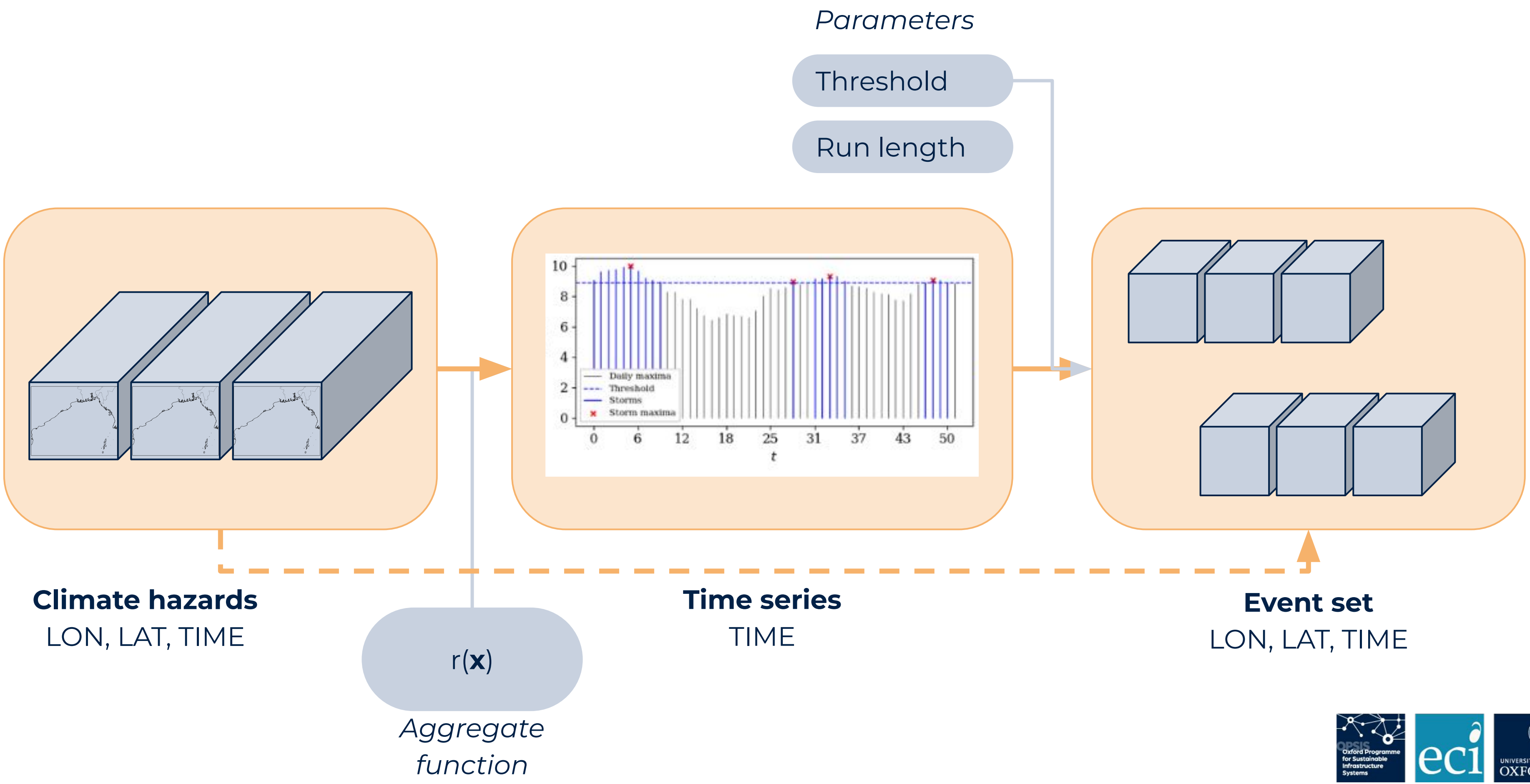
Method Overview

- Identify independent hazard events
- Extract footprints
- Estimate
  - Marginal distributions
  - Dependence structure
- Generate new samples

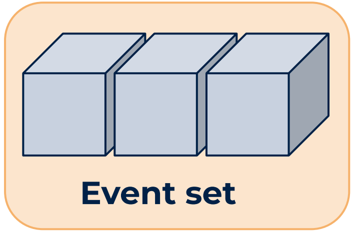


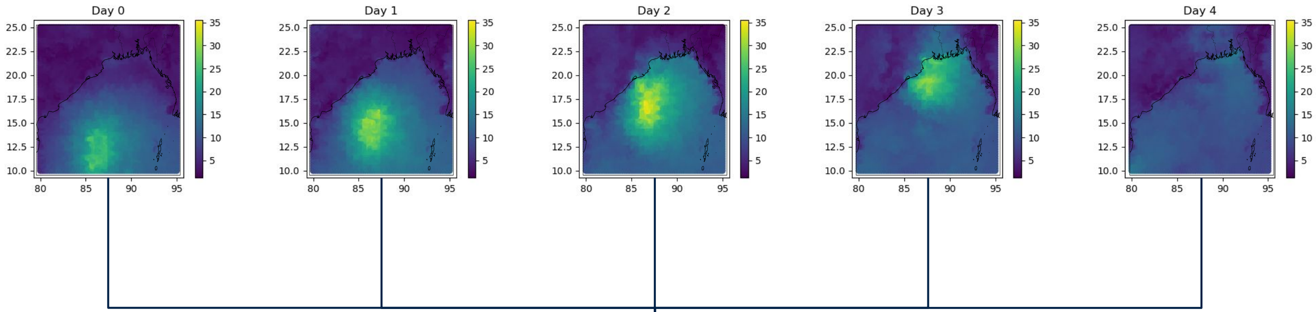


Method Identify independent hazard events

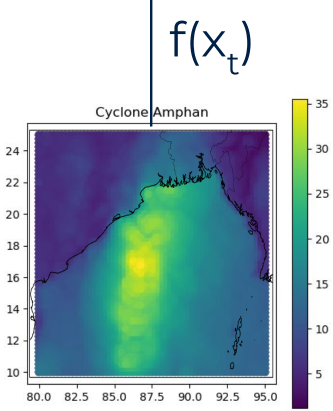


Method Extract Footprints



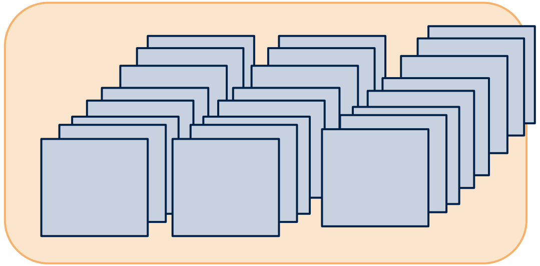


t



$f(x_t)$

- Maximum
- Minimum
- Cumulative sum
- Standard deviation



Hazard footprints

Method Estimate Marginal Distributions

Marginal Distribution of a single random variable

Estimate marginal distribution of every gridcell

Standard methods can underestimate extremes

e.g., Gaussian distribution, autoregression

⇒ Statistics of extremes

Asymptotically–justified

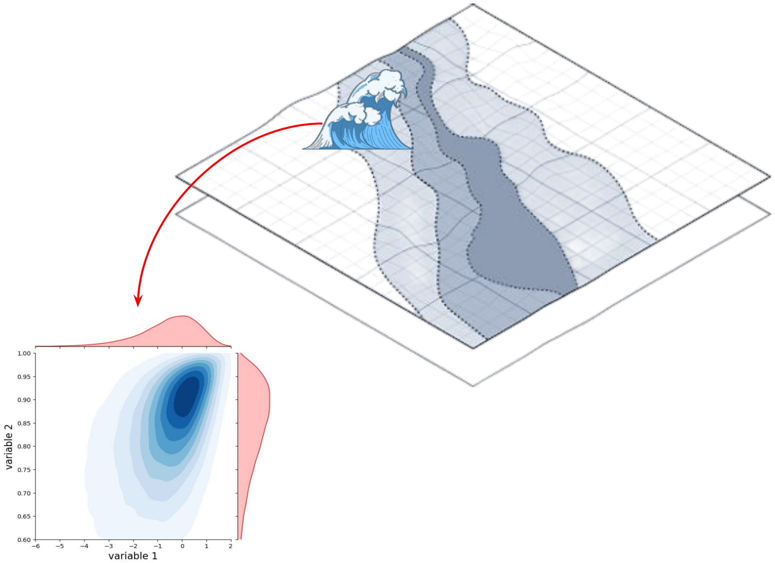
Supported by mathematical theory

- Block maxima
- Peak-over-threshold (POT)

Empirically-justified

Supported by strong empirical evidence

- Wind speeds
    - *"The overwhelming weight of evidence suggests that the parent distribution of wind speeds is of the Weibull type"* [1]



1. Harris (2005) *Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage?*

Method Estimate Marginal Distributions

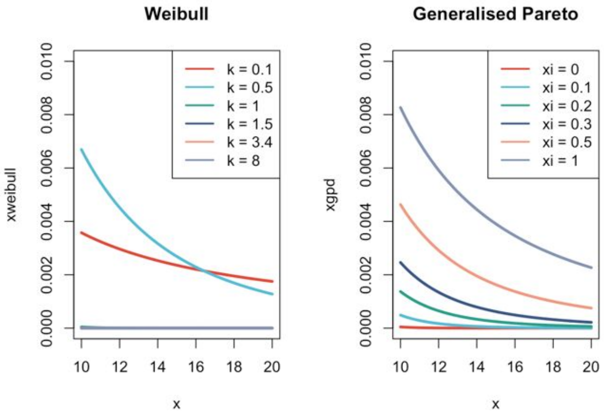
Marginal Distribution of a single random variable

Estimate marginal distribution of every gridcell

Standard methods can underestimate extremes
e.g., Gaussian distribution, autoregression
⇒ Statistics of extremes

Asymptotically–justified
Supported by mathematical theory
- Block maxima → Generalized Extreme Value (GEV)
- Peak-over-threshold (POT) → Generalized Pareto (GPD)

Empirically-justified
Supported by strong empirical evidence
- Wind speeds
  - *“The overwhelming weight of evidence suggests that the parent distribution of wind speeds is of the Weibull type”* [1]

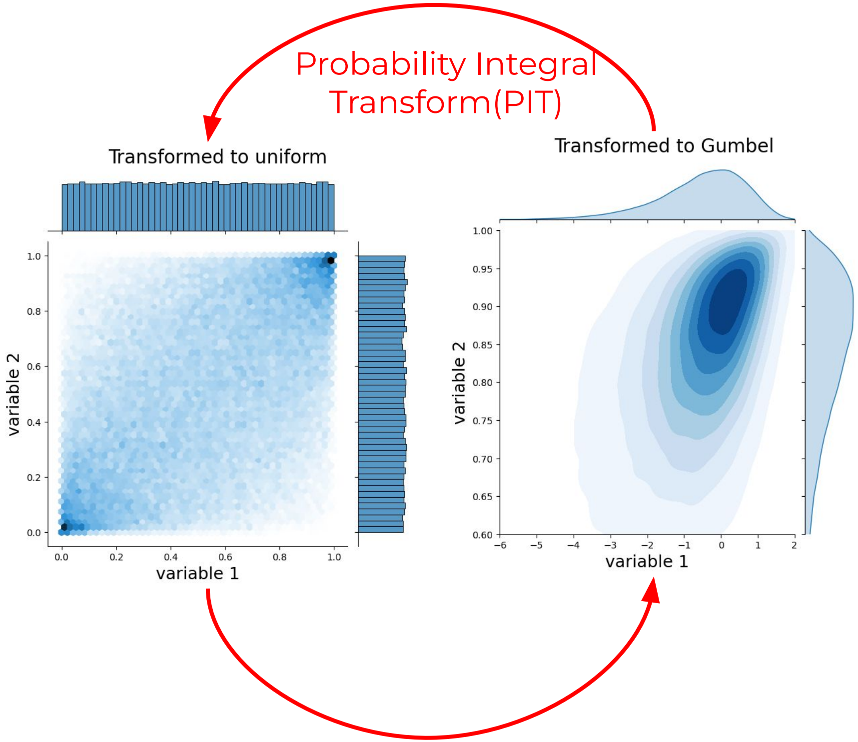


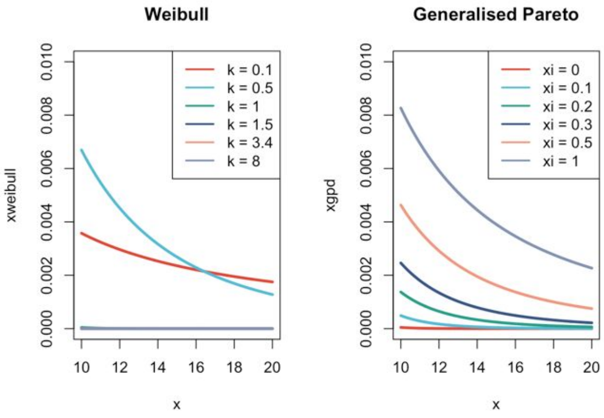
Wind ~ Weibull(σ,μ;k)
Precipitation ~ GPD(μ,σ,ξ)
Sea level pressure ~ GPD(μ,σ,ξ)

→ Estimate parameters

1. Harris (2005) *Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage?*

Method Estimate Marginal Distributions





Wind ~ Weibull(σ,ξ;k)
Precipitation ~ GPD(μ,σ,ξ)
Sea level pressure ~ GPD(μ,σ,ξ)

1. Harris (2005) *Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage?*

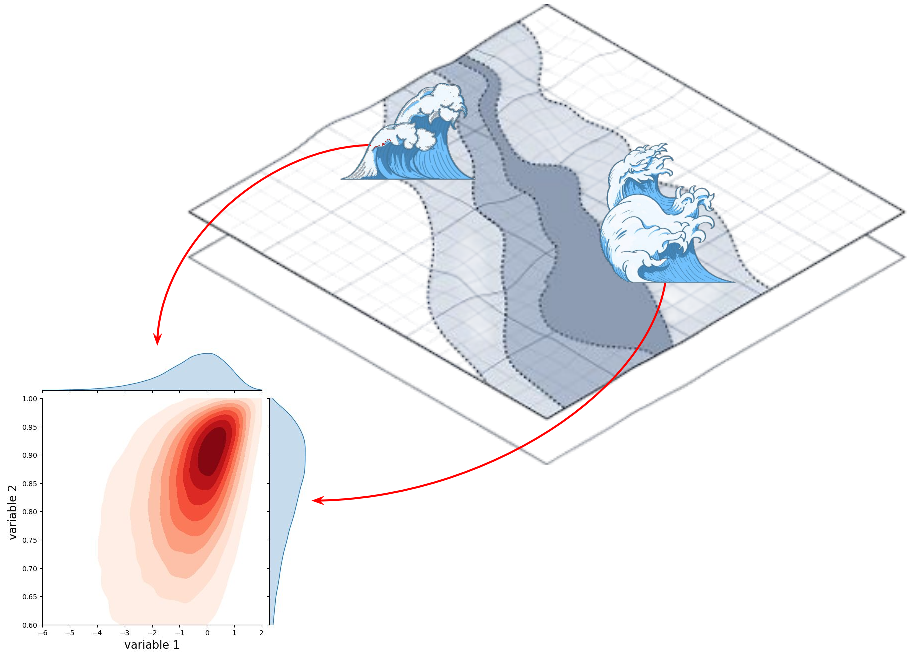
Method Estimate Dependence Structure

Classical

- High dimensionality
  - T x H x W x N variables
- Assumptions/tail flexibility
- Extremes
- E.g., Multivariate Gaussian, copulas

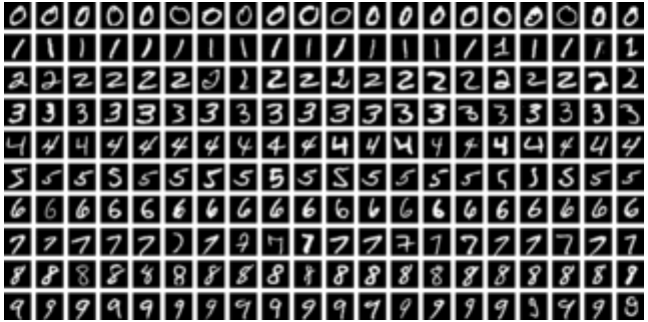
Solution

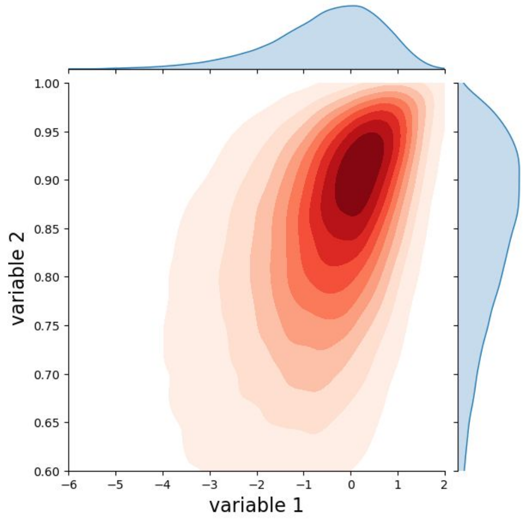
- Deep learning
  - *"Implicitly" learns the distribution*



Generative Adversarial Networks (GANs)

1. Data: ≥10,000 training samples
   ⇒ Need large dataset
2. Focus on mode
   ⇒ Rarest samples ignored
3. Equally sensitive across all ranges
   ⇒ Training data distribution matters

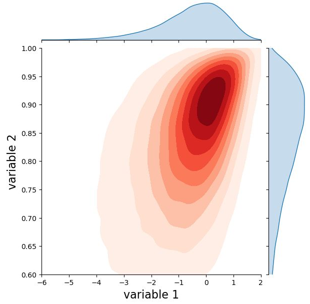




Method Modelling the dependence structure

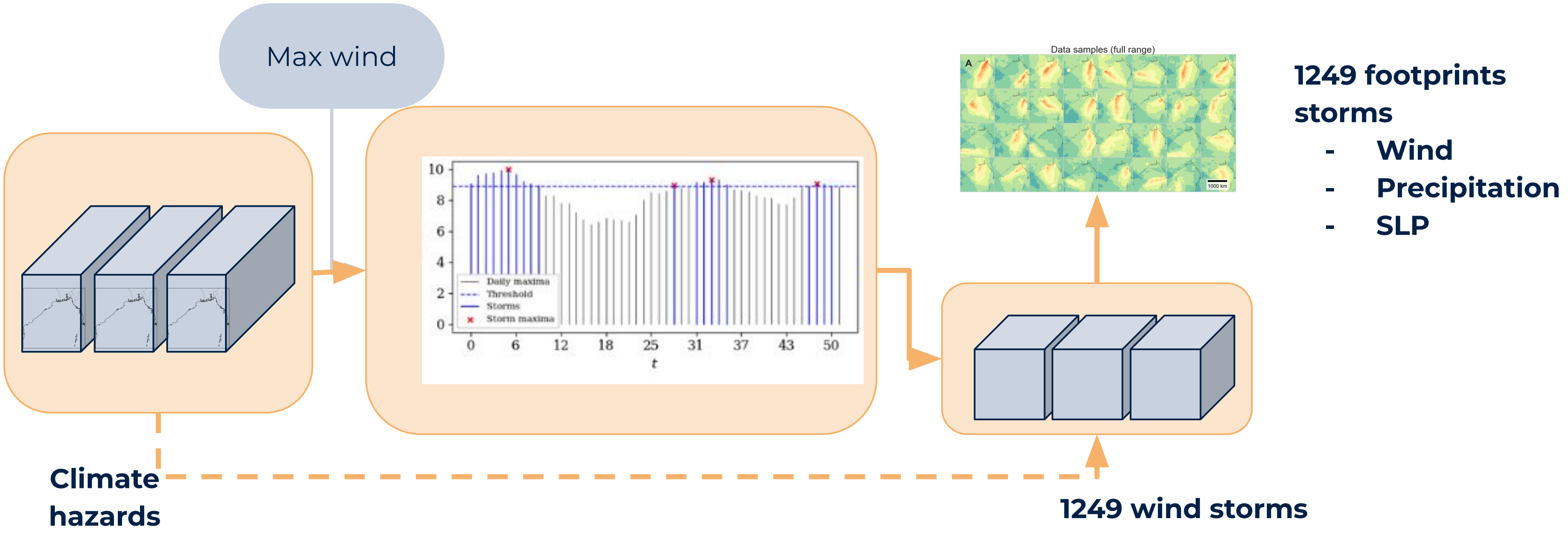
StyleGAN with Differentiable Augmentation

1. Good results with ~100 images
2. Only train on top 150 storms
3. Use with Gumbel PIT



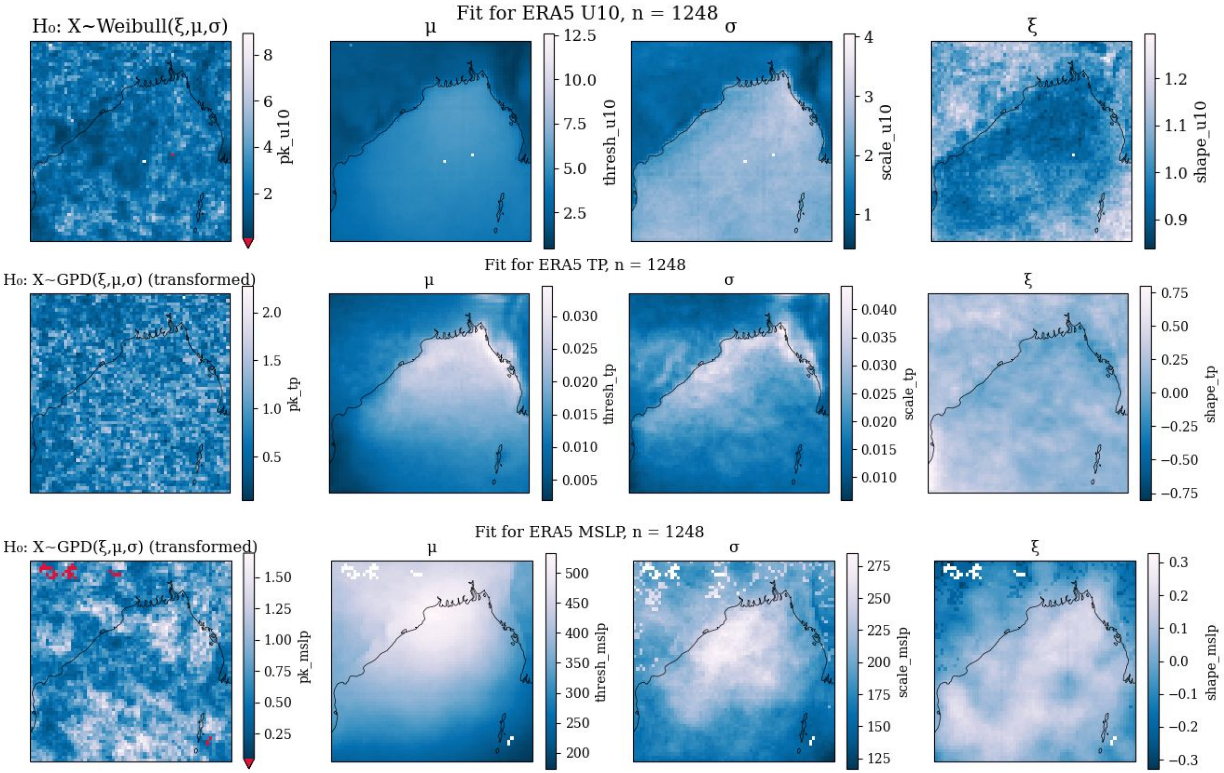
https://github.com/mit-han-lab/data-efficient-gans/tree/master

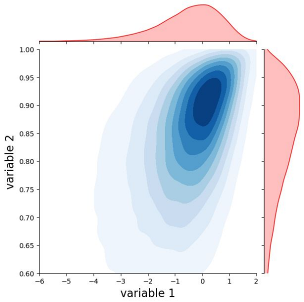
Application of Method ERA5 Reanalysis Data



Variable	Unit	Dataset	Temporal Range	Spatial Range	Aggregate	Extreme	In-Resolution	Out-Resolution
10m wind speed	mps	ERA5	1941-2022	10-25°N, 80-95°E	Maxima	Maxima	0.25° / ~26km	0.23° / ~24 km
Total precipitation	m	ERA5	1941-2022	10-25°N, 80-95°E	Sum	Maxima	0.25° / ~26km	0.23° / ~24 km
Sea level pressure	Pa	ERA5	1941-2022	10-25°N, 80-95°E	Minima	Minima	0.25° / ~26km	0.23° / ~24 km

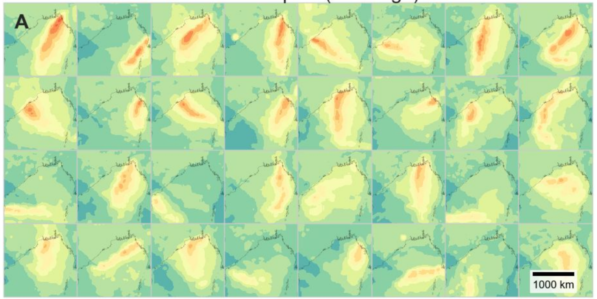
Application of Method Marginal distributions



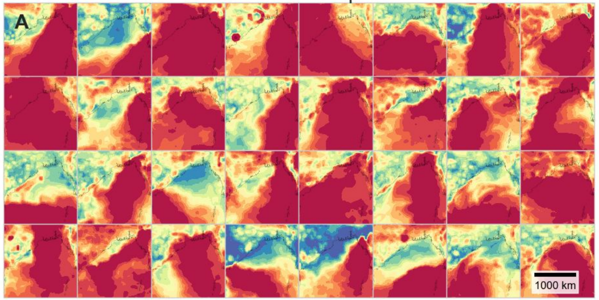


Application of Method Transforming fields

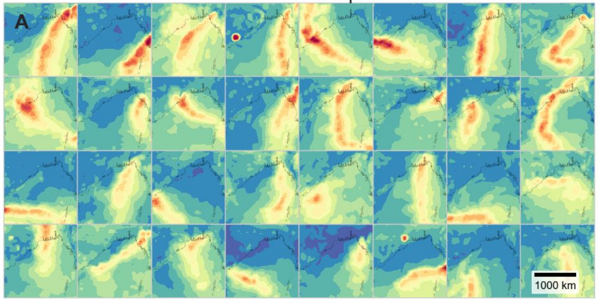
Data samples (full range)

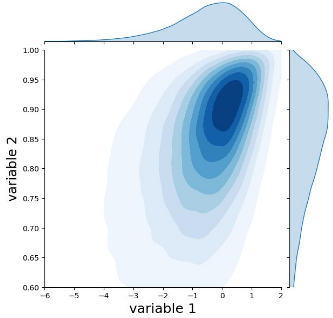


Uniform samples

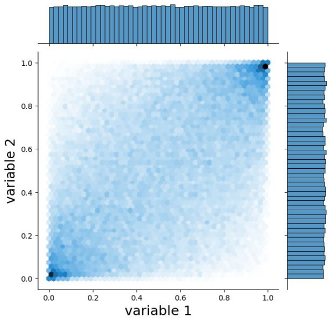


Gumbel samples

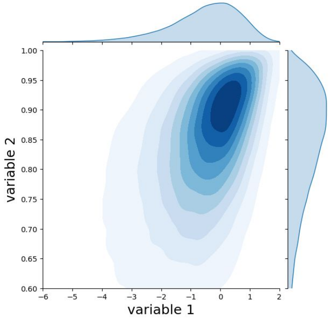




Transformed to uniform



Transformed to Gumbel



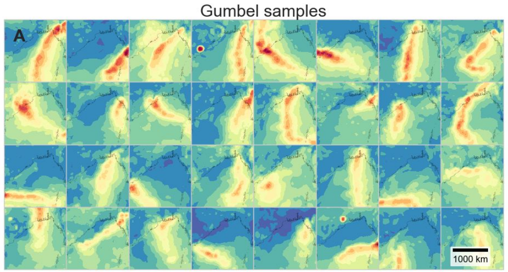
$F(x;\xi,\mu,\sigma)$

$F^{-1}_{Gumbel}(x)$

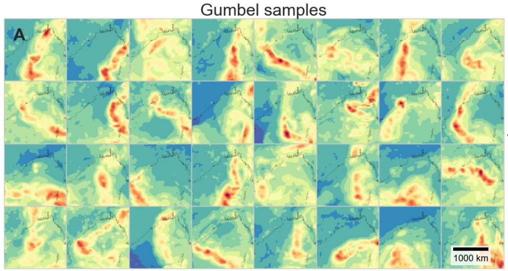
Probability Integral Transform(PIT)

Application of Method **StyleGAN Training**

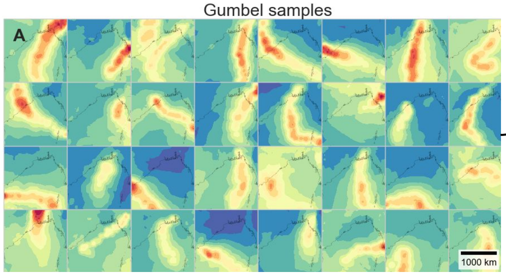
wind

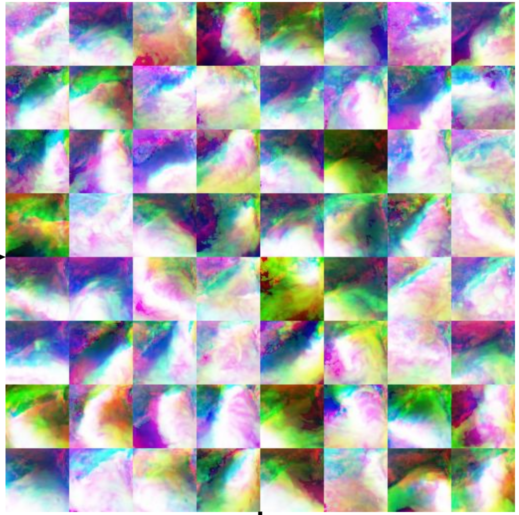


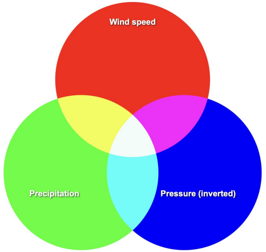
precipitation



pressure

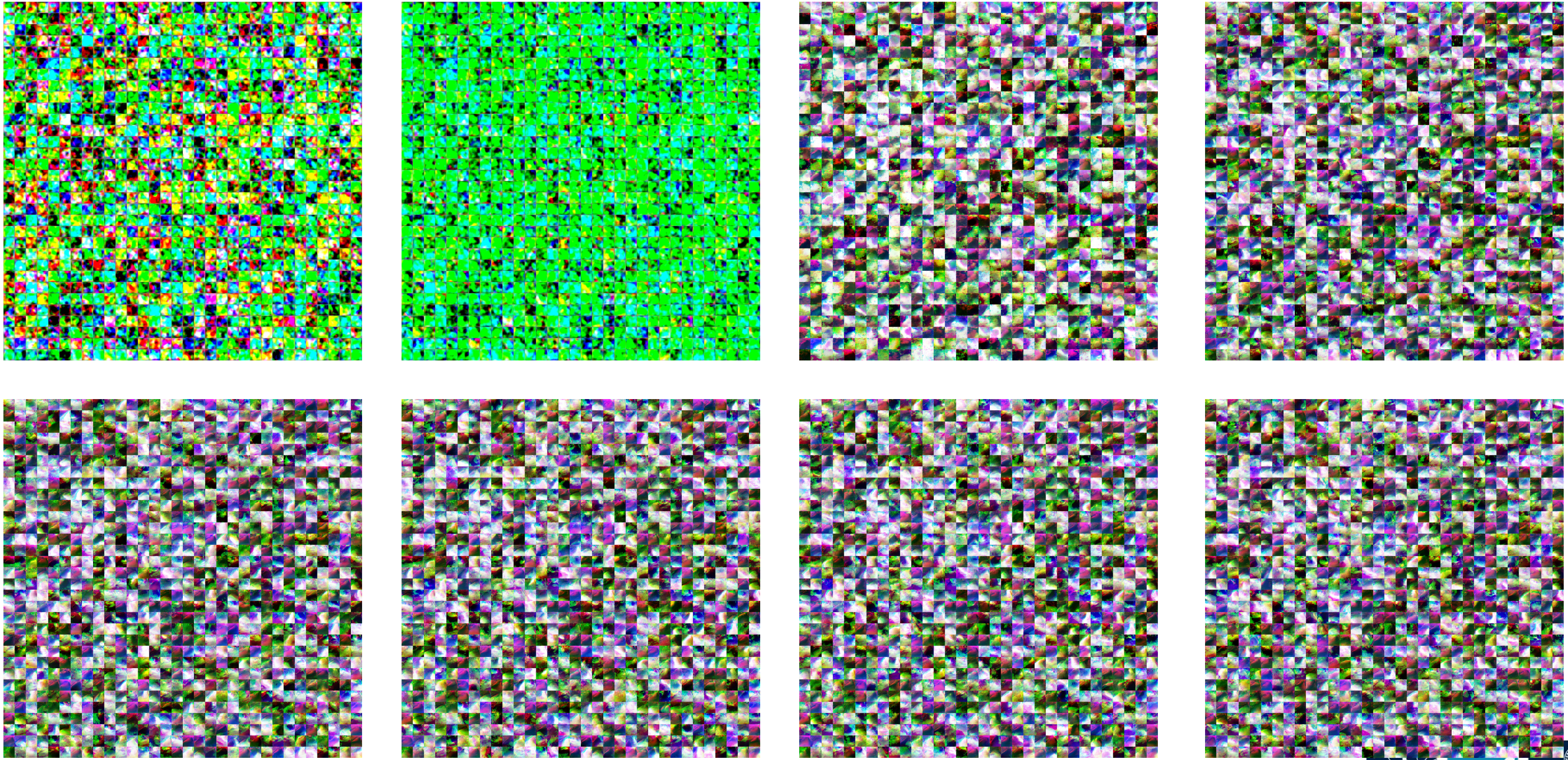




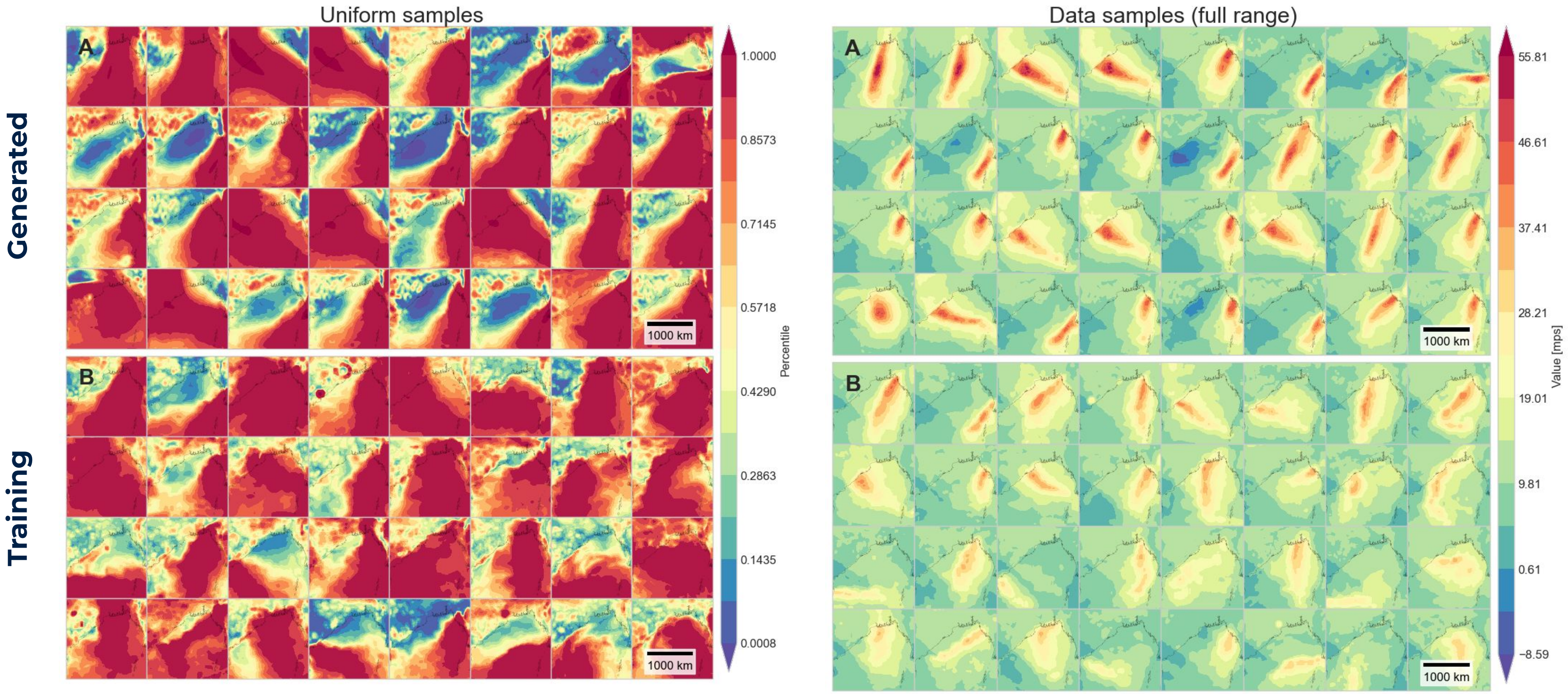


StyleGAN

Application of Method **StyleGAN Training**

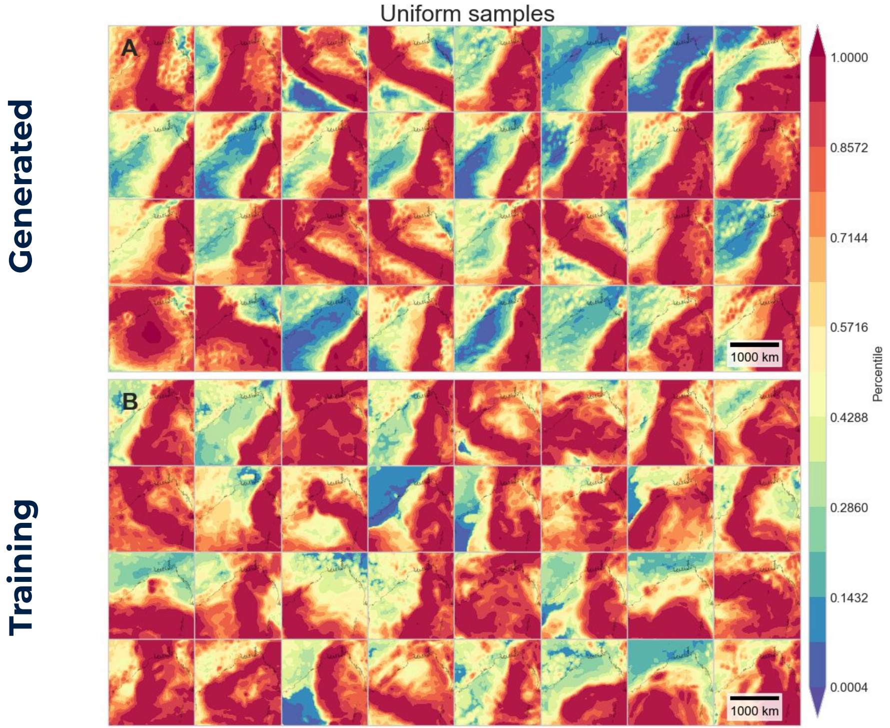


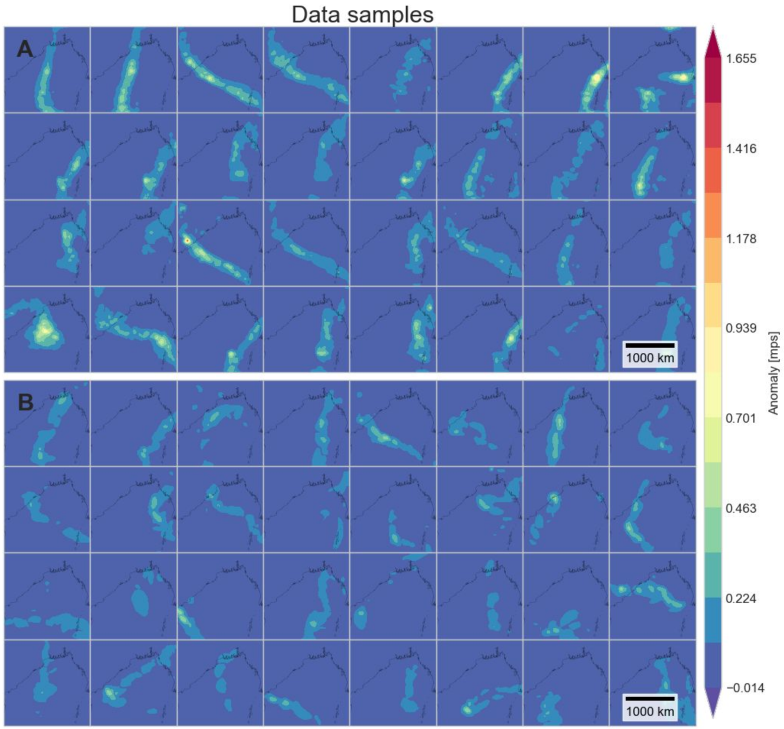
Application of Method Results (wind)



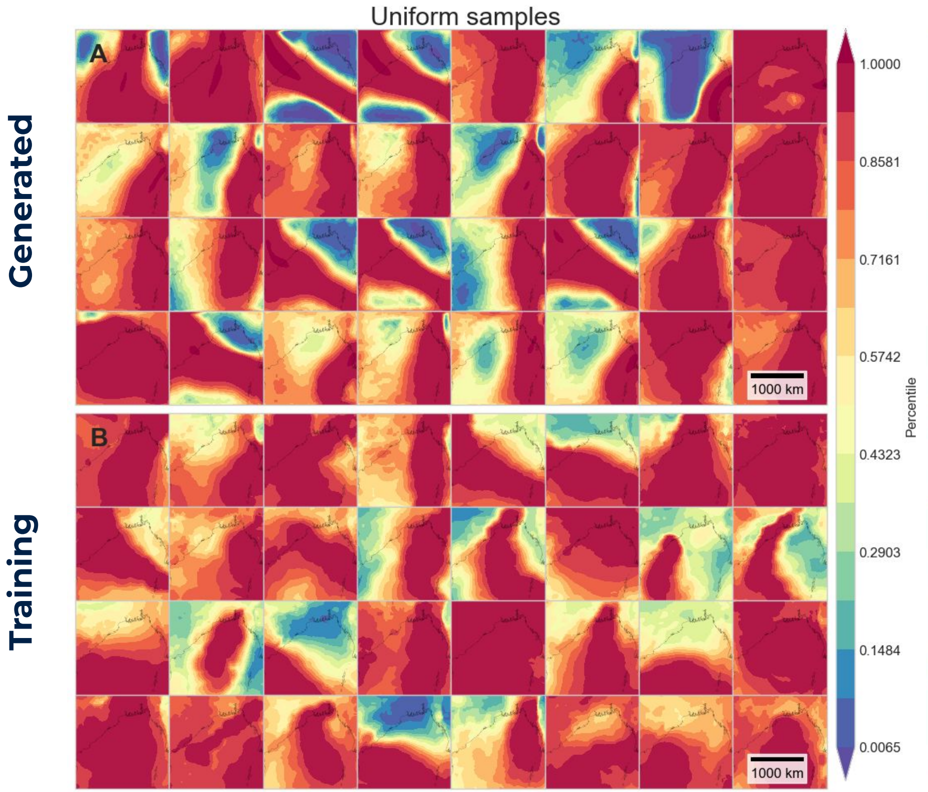
5000 samples

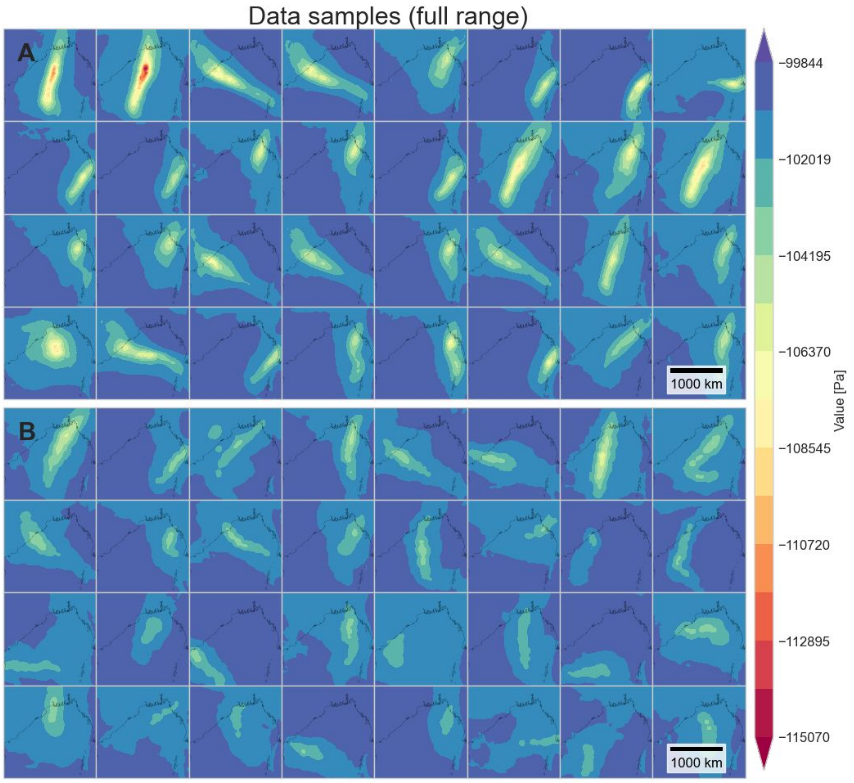
Application of Method Results (precipitation)



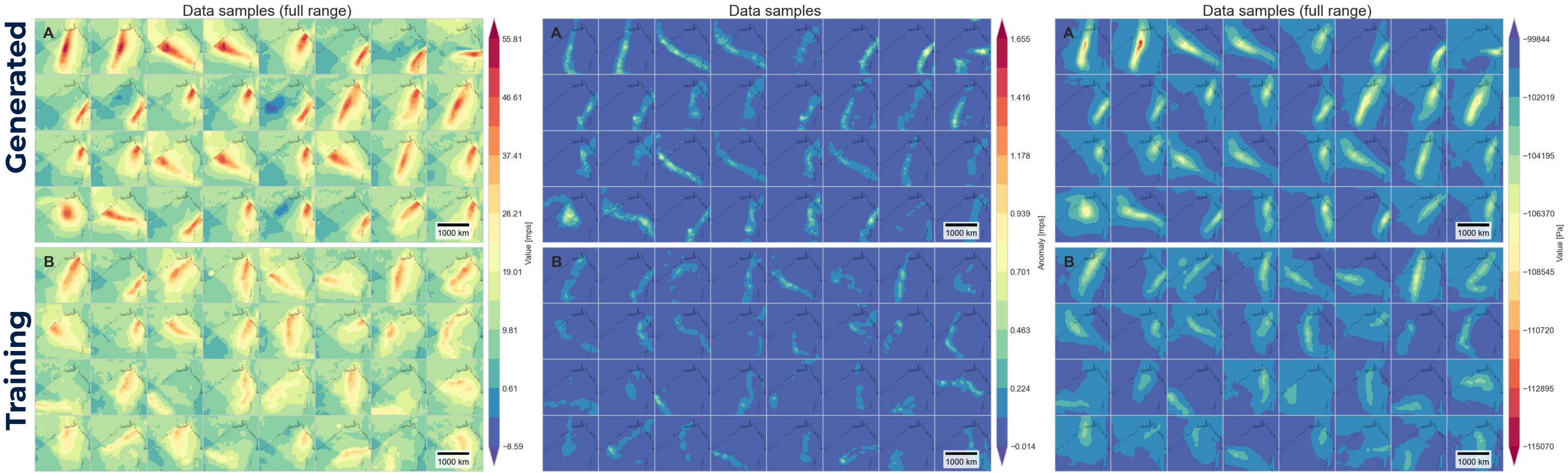


Application of Method Results (sea level pressure)

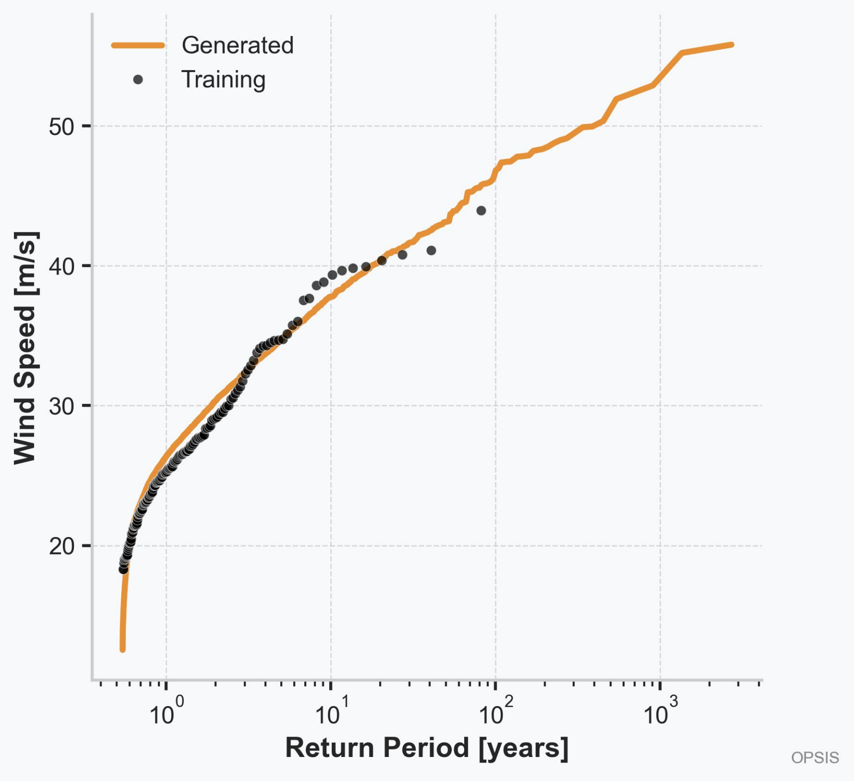


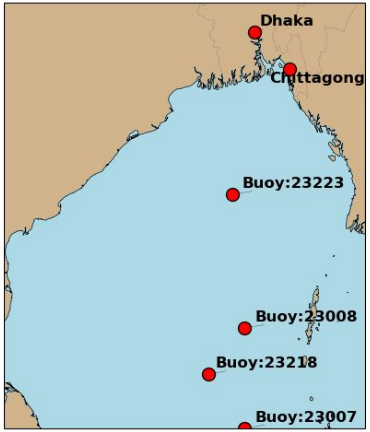


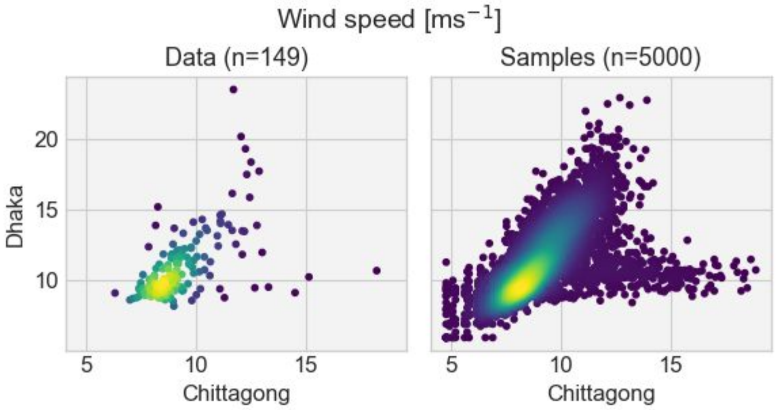
Application of Method Results



Application of Method Evaluation

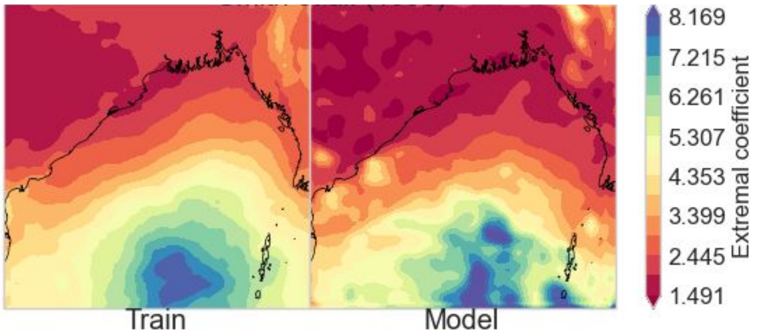


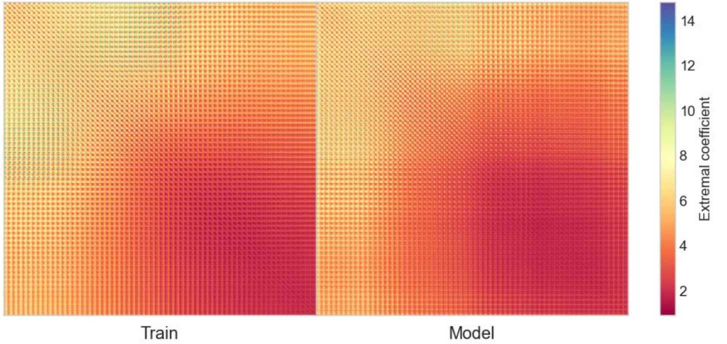




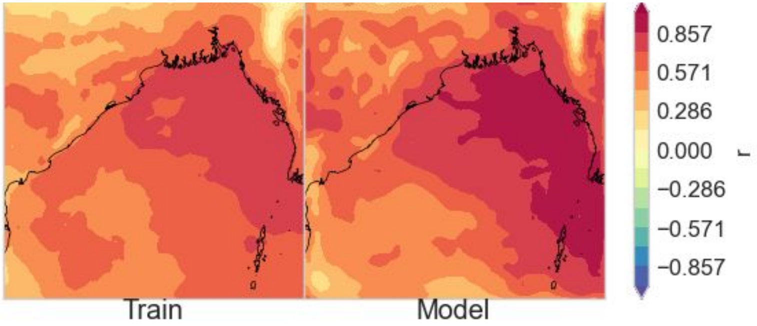
Application of Method Evaluation

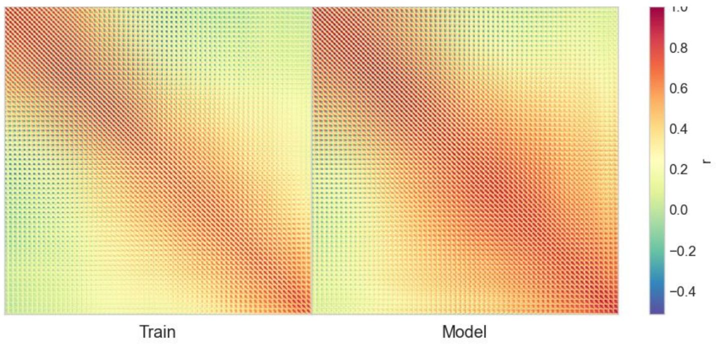
Extremal correlation





Pearson correlation

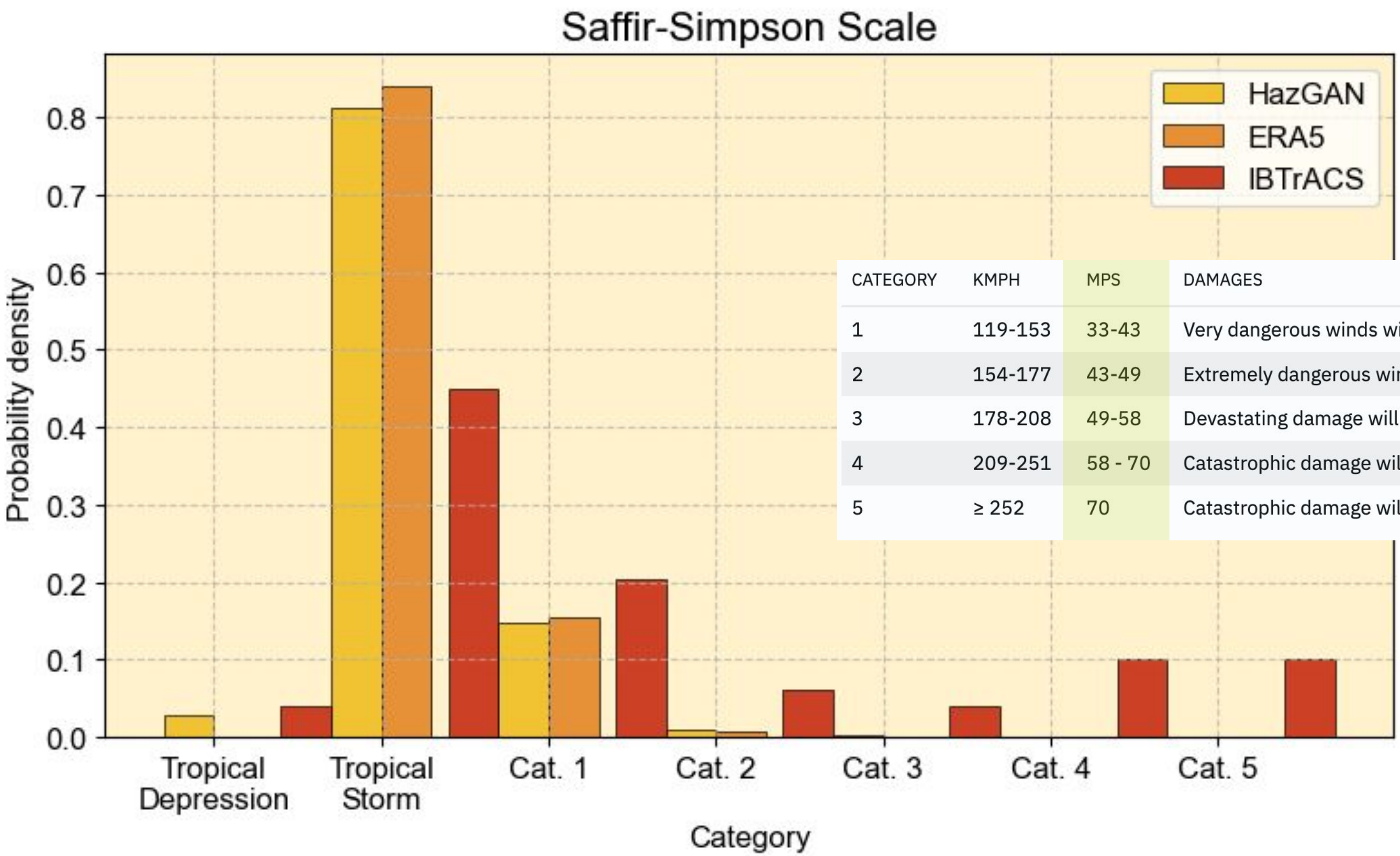




Wind - Precipitation

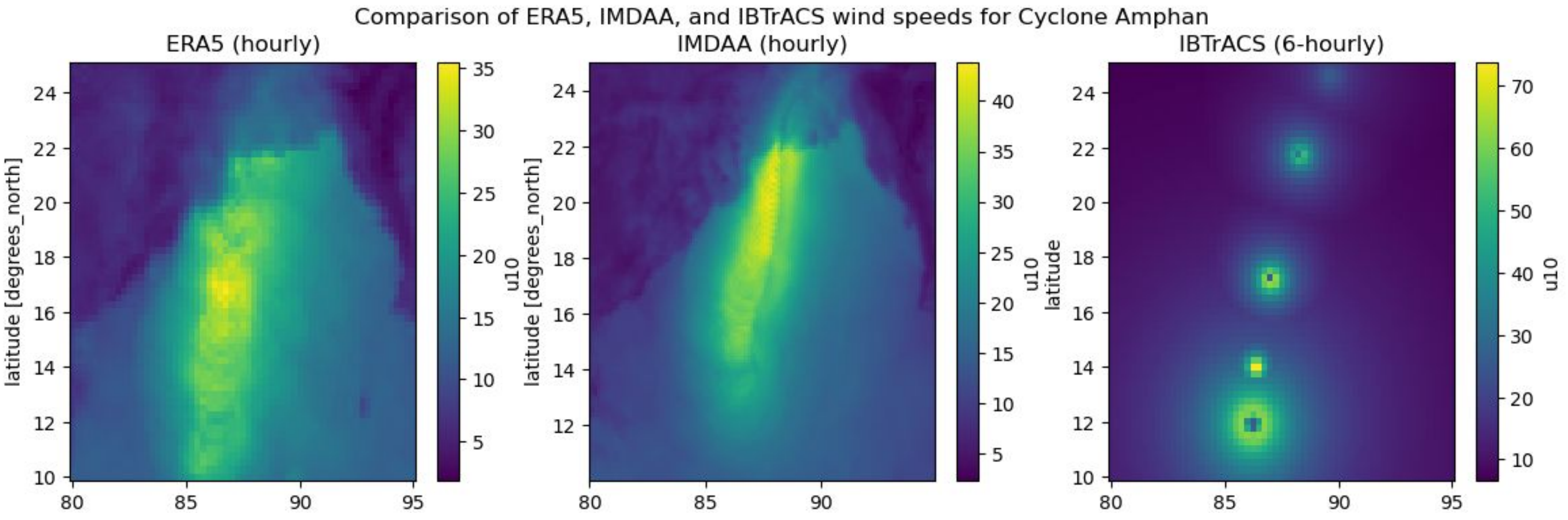
Across Space

Application of Method Evaluation



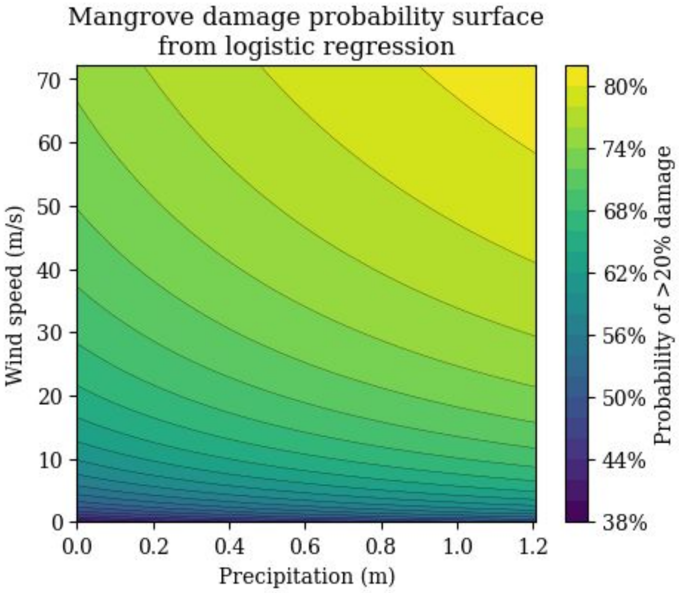
CATEGORY	KMPH	MPS	DAMAGES
1	119-153	33-43	Very dangerous winds will produce some damage
2	154-177	43-49	Extremely dangerous winds will cause extensive dam
3	178-208	49-58	Devastating damage will occur
4	209-251	58 - 70	Catastrophic damage will occur
5	≥ 252	70	Catastrophic damage will occur

Discussion Comparison to STORM



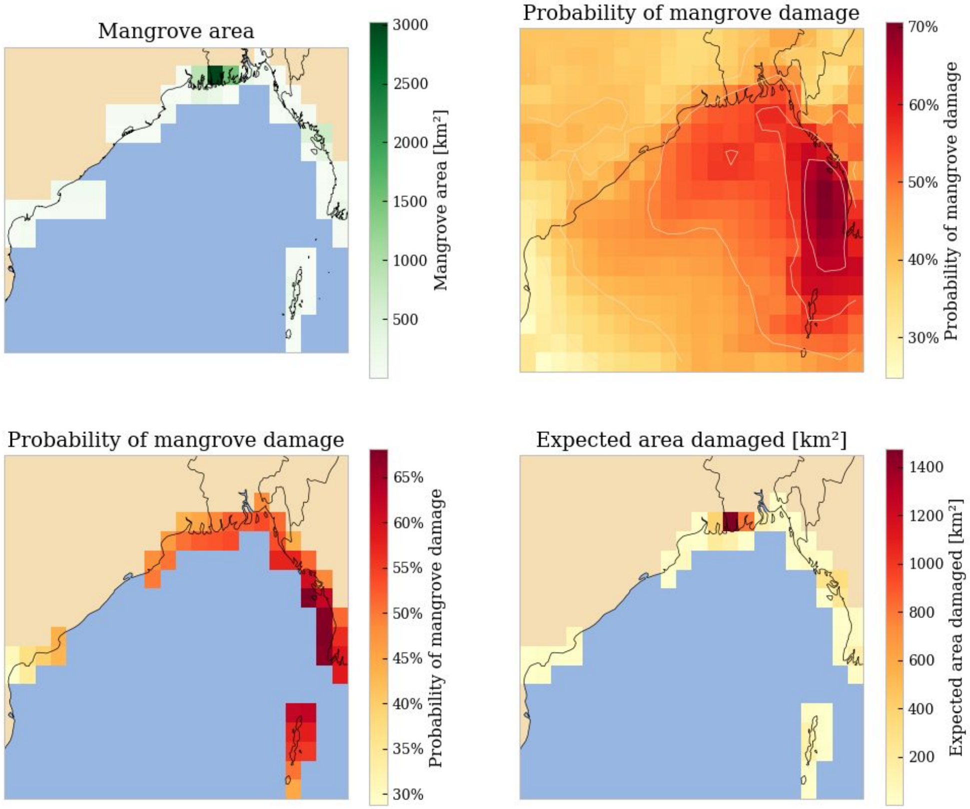
CATEGORY	KMPH	MPS	DAMAGES
1	119-153	33-43	Very dangerous winds will produce some damage
2	154-177	43-49	Extremely dangerous winds will cause extensive damage
3	178-208	49-58	Devastating damage will occur
4	209-251	58 - 70	Catastrophic damage will occur
5	≥ 252	70	Catastrophic damage will occur

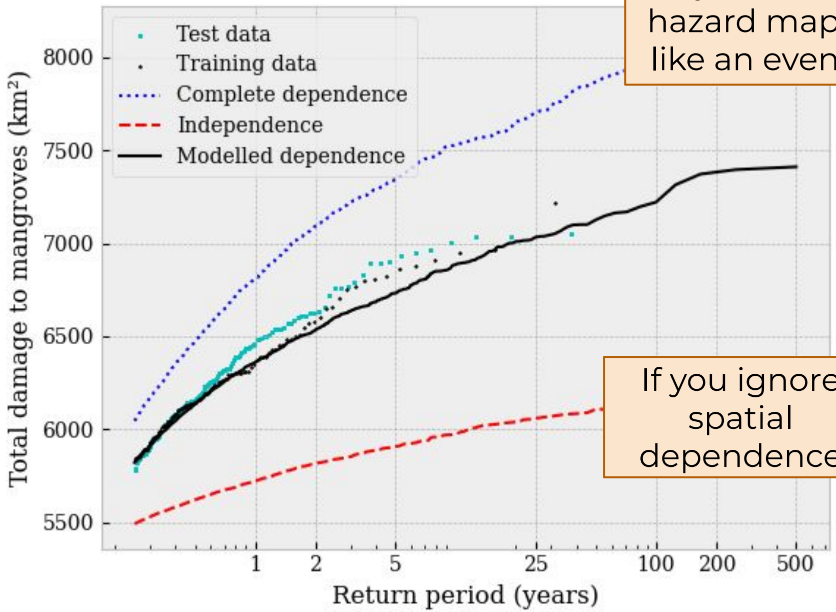
Mangrove risk Mangrove damages (old results)



Damage = ≥20% loss of NDVI

Mangrove risk Mangrove damages (old results)

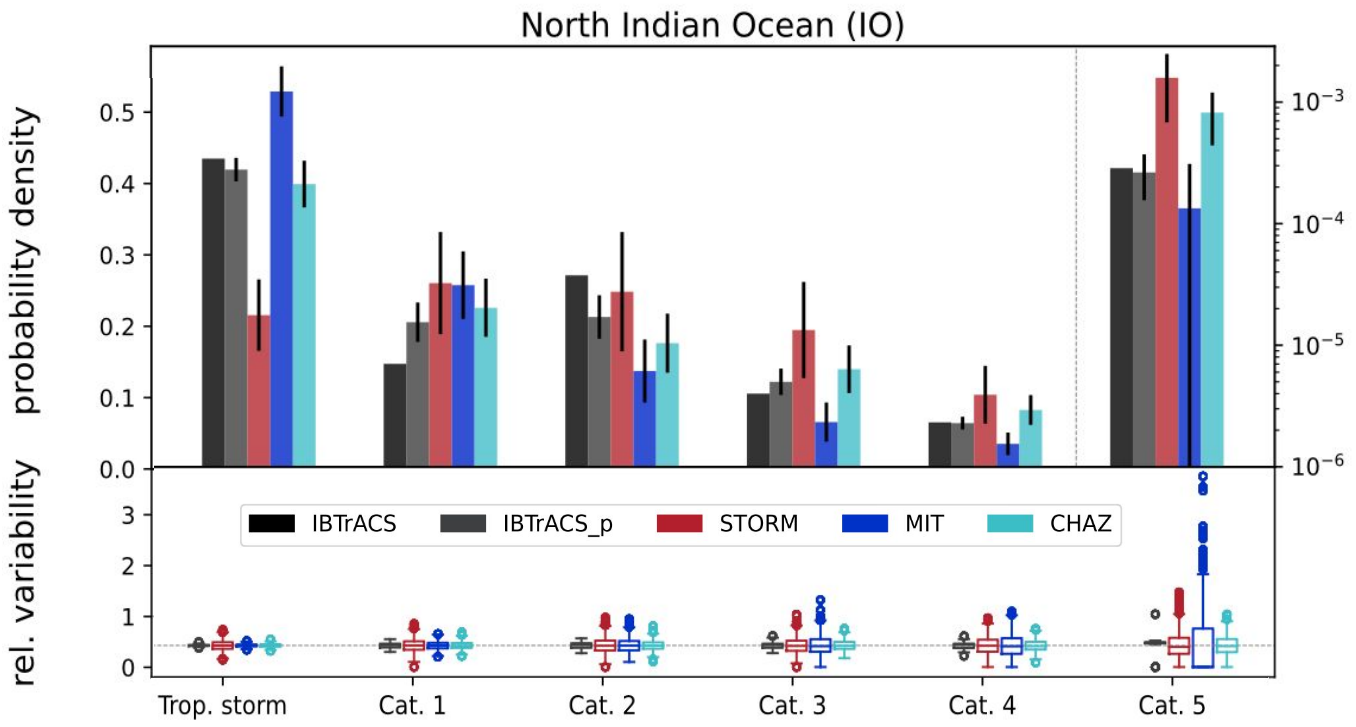




If you treat hazard maps like an event

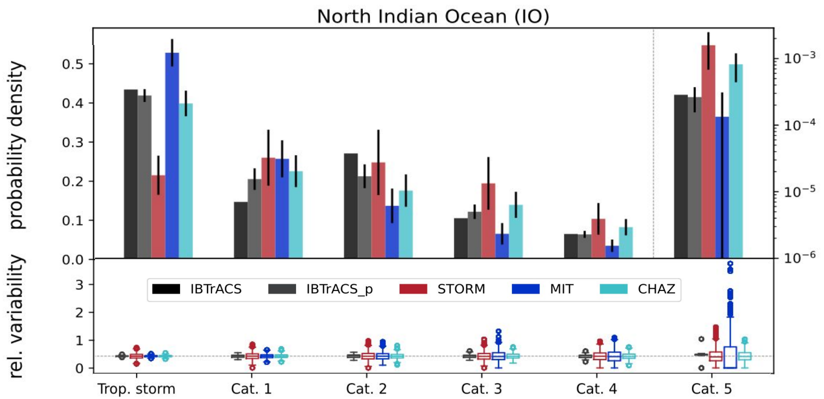
If you ignore spatial dependence

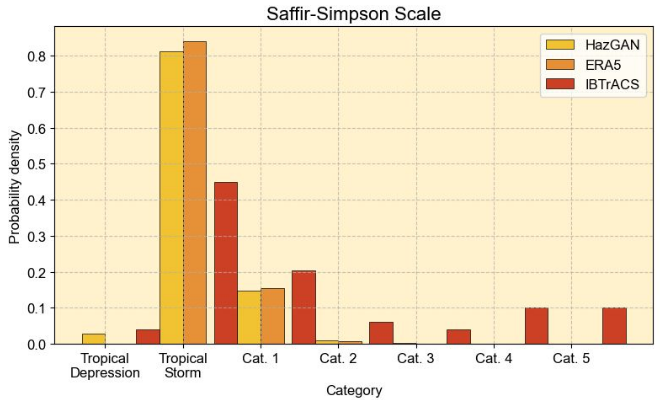
Discussion Comparison to STORM

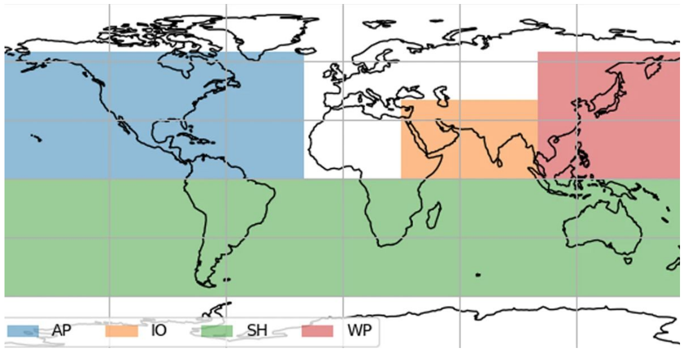


1. Meiler (2022) Supplementary Information for article"Intercomparison of regional loss estimates from global synthetic tropical cyclone models"

Discussion Comparison to STORM







CATEGORY	KMPH	MPS	DAMAGES
1	119-153	33-43	Very dangerous winds will produce some damage
2	154-177	43-49	Extremely dangerous winds will cause extensive damage
3	178-208	49-58	Devastating damage will occur
4	209-251	58 - 70	Catastrophic damage will occur
5	≥ 252	70	Catastrophic damage will occur

1. Meiler (2022) Supplementary Information for article"Intercomparison of regional loss estimates from global synthetic tropical cyclone models"

Discussion Comparison to STORM

	HazGAN	STORM
Accuracy	Better matches training data ✅	Overestimates extremes
Training data	ERA5 underestimates extremes	IBTrACS ✅
Extremes	Characterised ✅	Not characterised
Coverage	Training area	Global ✅
Variables	Any ✅	Wind and pressure
Future projections	No	No
Intepretability	Limited	Straightforward ✅

Case study 1: Damage to electricity networks in the UK

Create enough ensemble members to inform cost-benefit analysis of 30 years of losses with/without adapting UK electricity network.

Focus on low-impact storms which are not covered by insurance.

Variables:

1. Wind speed | ERA5
2. Wind direction | ERA5
3. Antecedent 30 day precipitation | HadUK

Key metric: Number of faults

Case study 2: Wind and solar drought in the UK

Model the probability of nationwide solar and wind droughts and its effect on the UK grid

Expertise needed

Variables:

1. Wind speed | ERA5
2. Cloud cover | ERA5
3. Solar radiation | ERA5

https://docs.google.com/spreadsheets/d/1TkQs0vy4CqpcNhXkIMQ1jxJBW6M531t5_OGymaUn8pw/edit?gid=0#gid=0

Thanks for listening!