Using Generative Models to Produce Realistic Populations of UK Windstorms

Etron Tsoi[1,2], Kieran Hunt[1,3], Len Shaffrey[1,3], Atta Badii[4], Richard Dixon[1,5,6], Ludovico Nicotina[7]

[1]Department of Meteorology, University of Reading, UK
[2]Department of Marine Systems, Tallinn University of Technology, Tallinn, Estonia
[3]National Centre for Atmospheric Science, University of Reading, UK
[4]Department of Computer Science, University of Reading, UK
[5]CatInsight, London, UK; [6]Oak Re, London, UK; [7]Inigo, London, UK

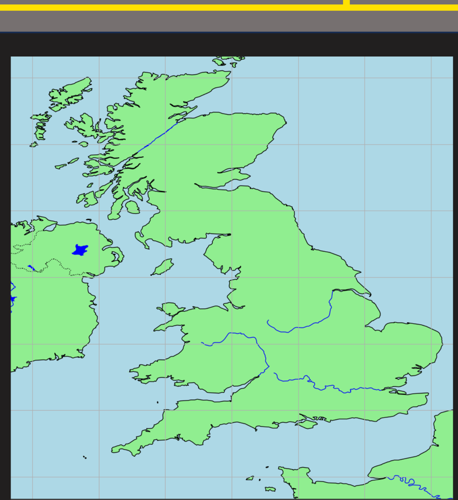
Contact:
yeechun.tsoi@taltech.ee

Preprint Available:

Introduction

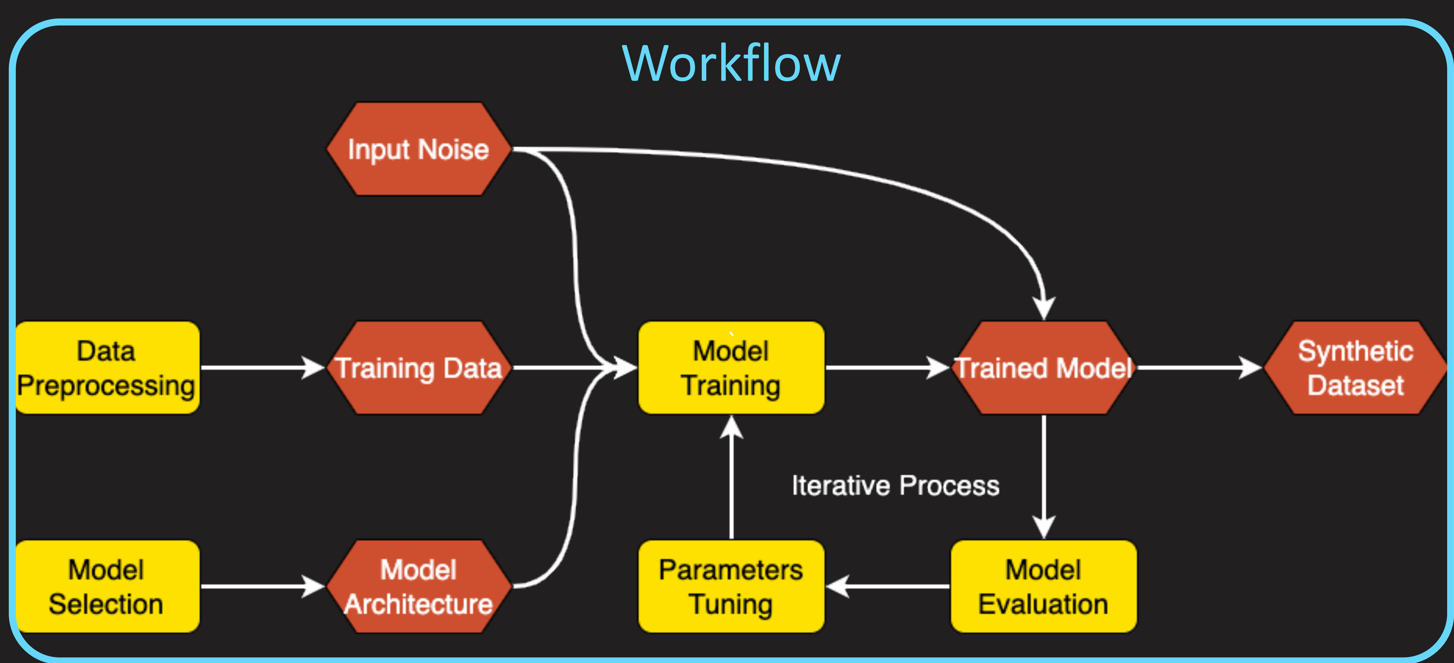
Limited observational data for extreme windstorms pose challenges for developing robust catastrophe models. Current approach of applying regional climate models is too time-consuming for creating extensive hazard datasets. We propose using generative models to produce independent wind fields, offering a scalable solution for risk assessment.

Data Description

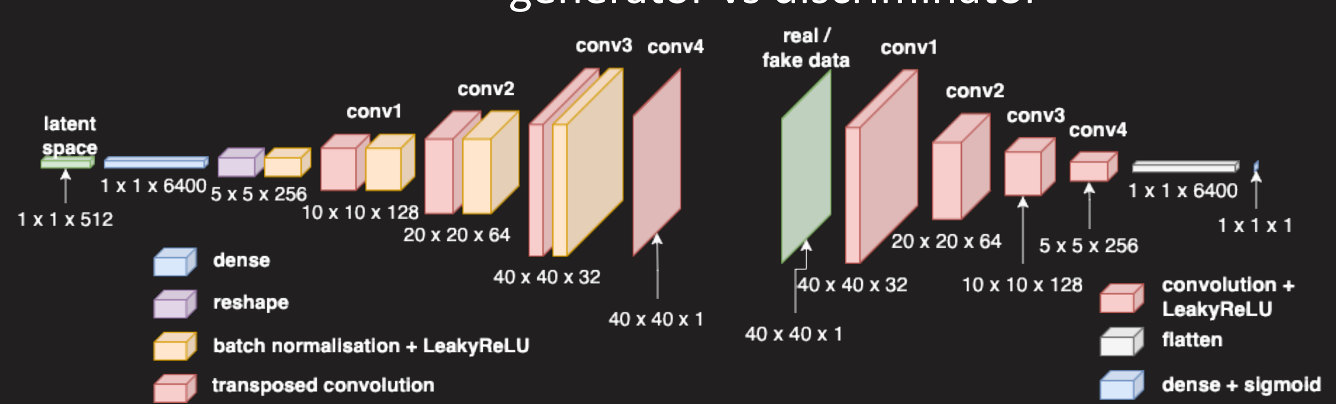


ERA5 reanalysis from ECMWF
- Spatial domain: 49°N to 59°N, 8°W to 2°E (0.25° × 0.25°)
- Period: 1940 – 2022
- Variables: 10-m wind speed
- Pre-processed to a normalized range of [0,1]

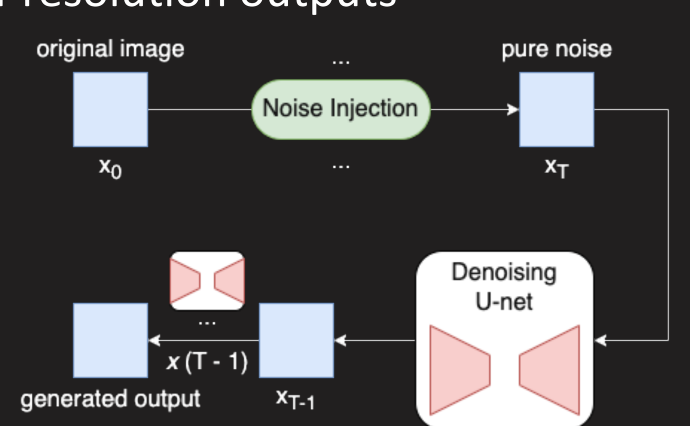
Methodology



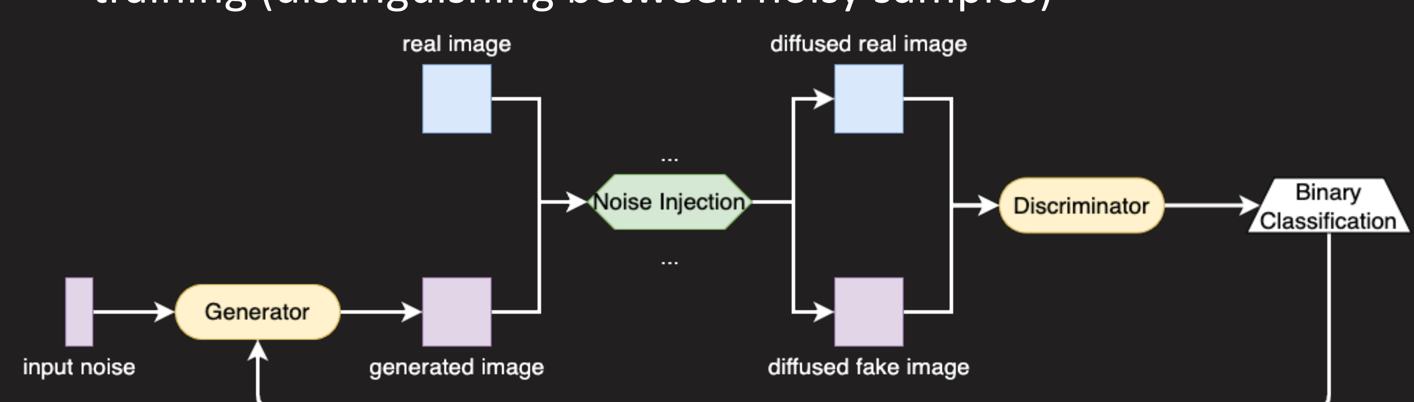
- **Standard GAN**: A baseline adversarial network generator vs discriminator



- **WGAN-GP**: Similar structure as **Standard GAN** but with Wasserstein loss and gradient penalty for better training stability
- **U-net Diffusion Model**: Uses diffusion processes (denoising) to generate high-resolution outputs



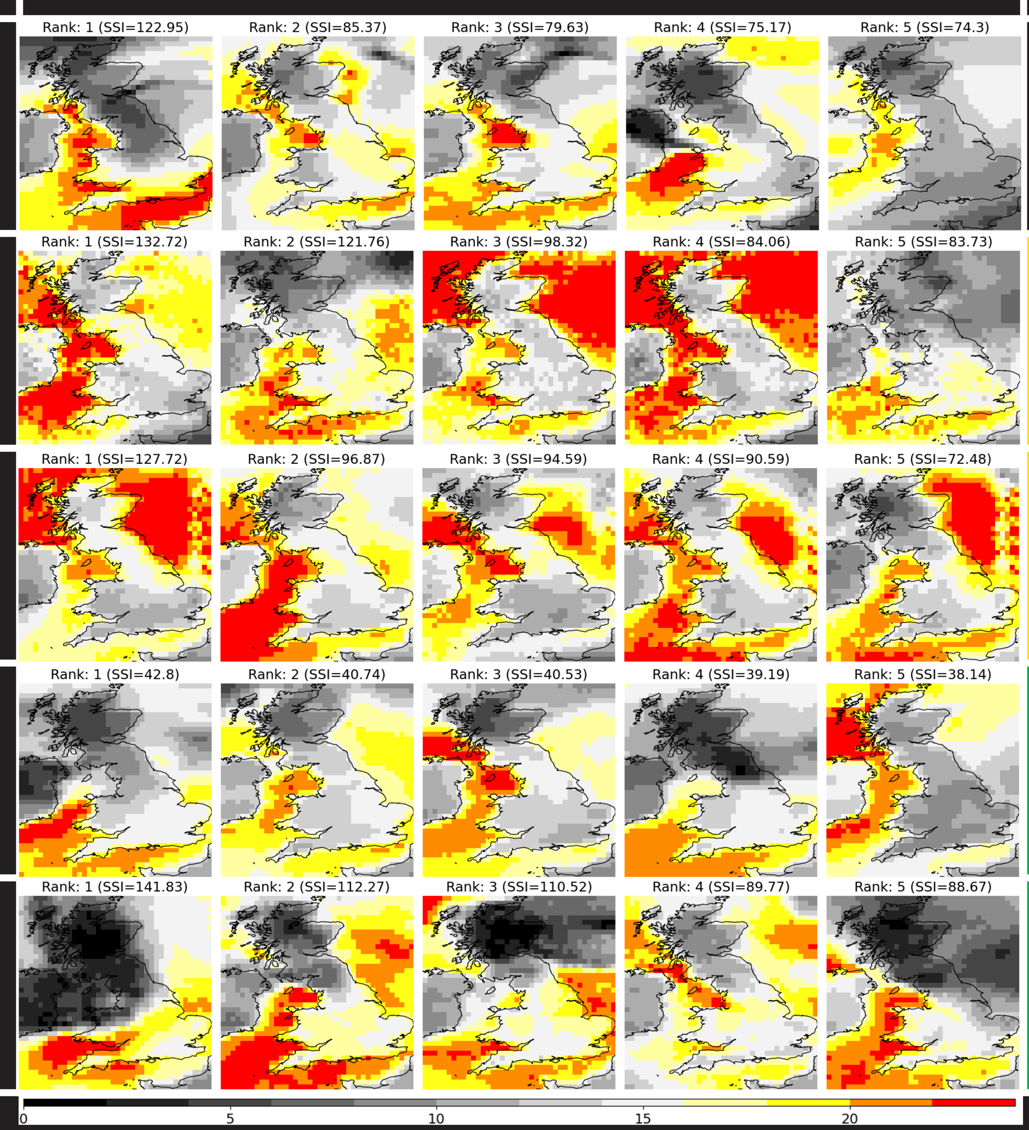
- **Diffusion-GAN**: Combines diffusion processes with adversarial training (distinguishing between noisy samples)



References

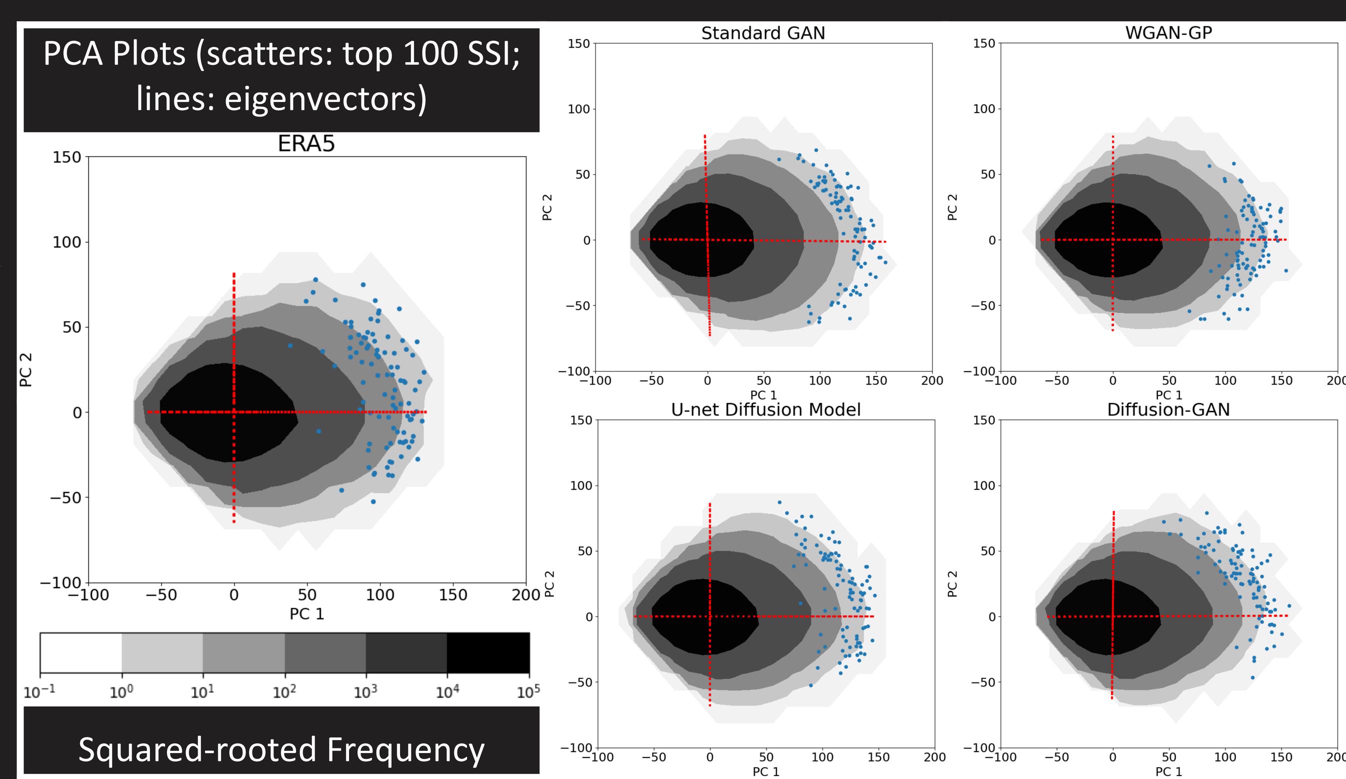
- Besombes, C., and Coauthors, 2021: Producing realistic climate data with generative adversarial networks. Nonlinear Processes in Geophysics, 28(3).
- Brochet, C., and Coauthors, 2023: Multivariate Emulation of Kilometer-Scale Numberical Weather Predictions with Generative Adversarial Networks: A Proof of Concept. Artificial Intelligence for the Earth Systems, 2(4).

Results

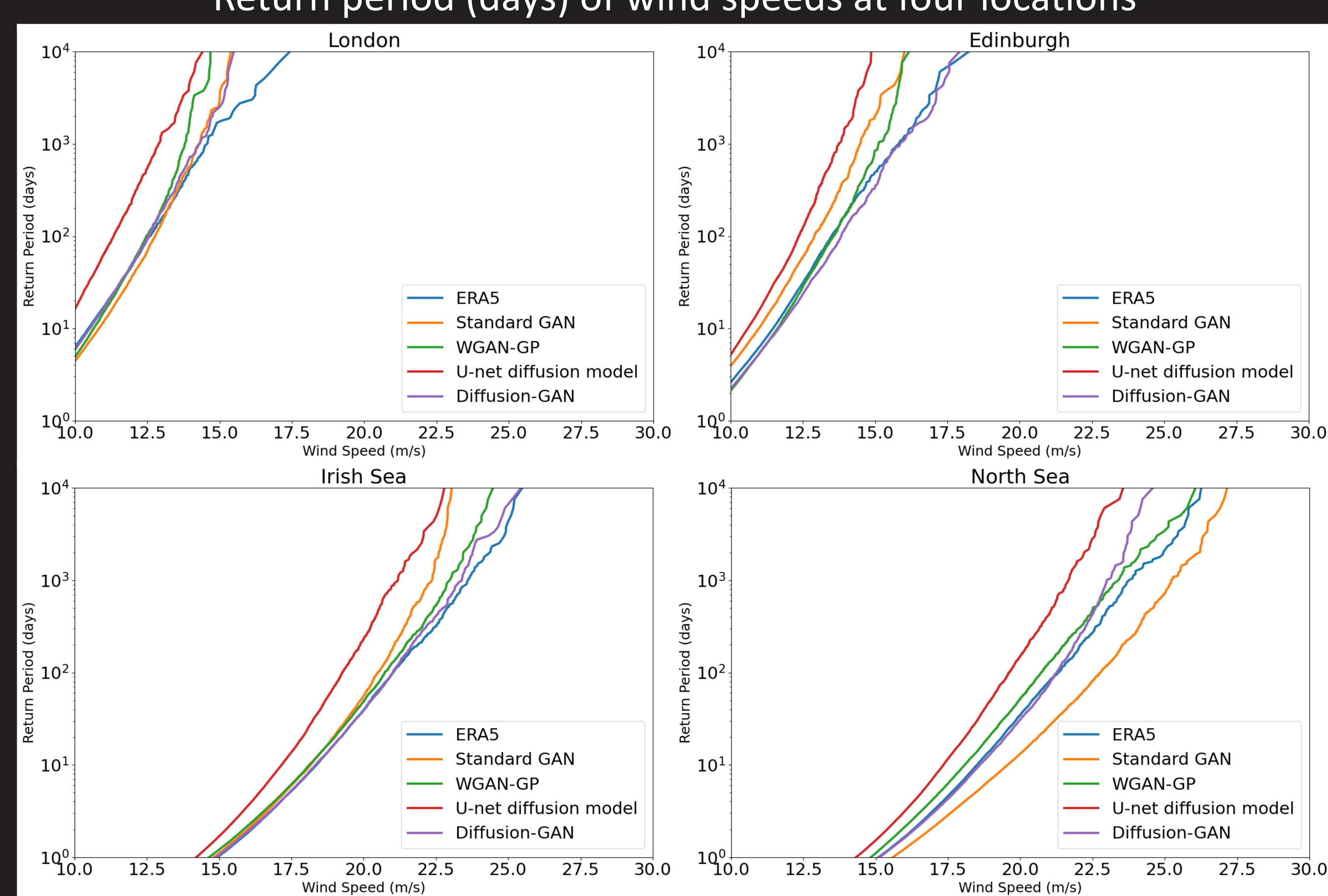


	Extreme Cases (Highest SSI)	Spatial pattern	Image quality	Storm intensity
ERA5				
Standard GAN		Over-extension of intense winds	Less smooth boundary	Over-estimate
WGAN-GP		Mismatch areas of strongest wind	Random Impulse Noise	
U-net Diffusion Model				Under-estimate
Diffusion-GAN				Over-estimate

- Similar shape and density of contours
- More dispersed distribution of the top 100 SSI cases → more variability in the generated extremes
- **Standard GAN** & **WGAN-GP**: Extremes with more negative PC2 (more storms in the Scottish coast)



Return period (days) of wind speeds at four locations



- **U-net diffusion** model over-estimates return periods across the displayed range
- Other models slightly over-estimates return periods at the rarer tail-end (>10^3 days)
- **Standard GAN** underestimates at the North Sea
- **WGAN-GP** & **Diffusion-GAN**: Consistent in capturing reasonable distributions

Conclusion

Trade-offs between stability, variability, and the ability to represent extremes!

- **Standard GAN**: Struggles in replicating image quality and extreme events
- **WGAN-GP**: Captures intensity well but sometimes misrepresents extremes
- **U-net Diffusion Model**: Good visual quality but underestimates intensity
- **Diffusion-GAN**: Best overall but overestimates most extreme intensity

Future work

- Incorporating multiple meteorological variables and temporal dimensions
- Expand approach to other meteorological hazards and regions
- Identify the strengths (which regions/scenarios) of different models
- Develop an ensemble approach for targeted optimization and applications