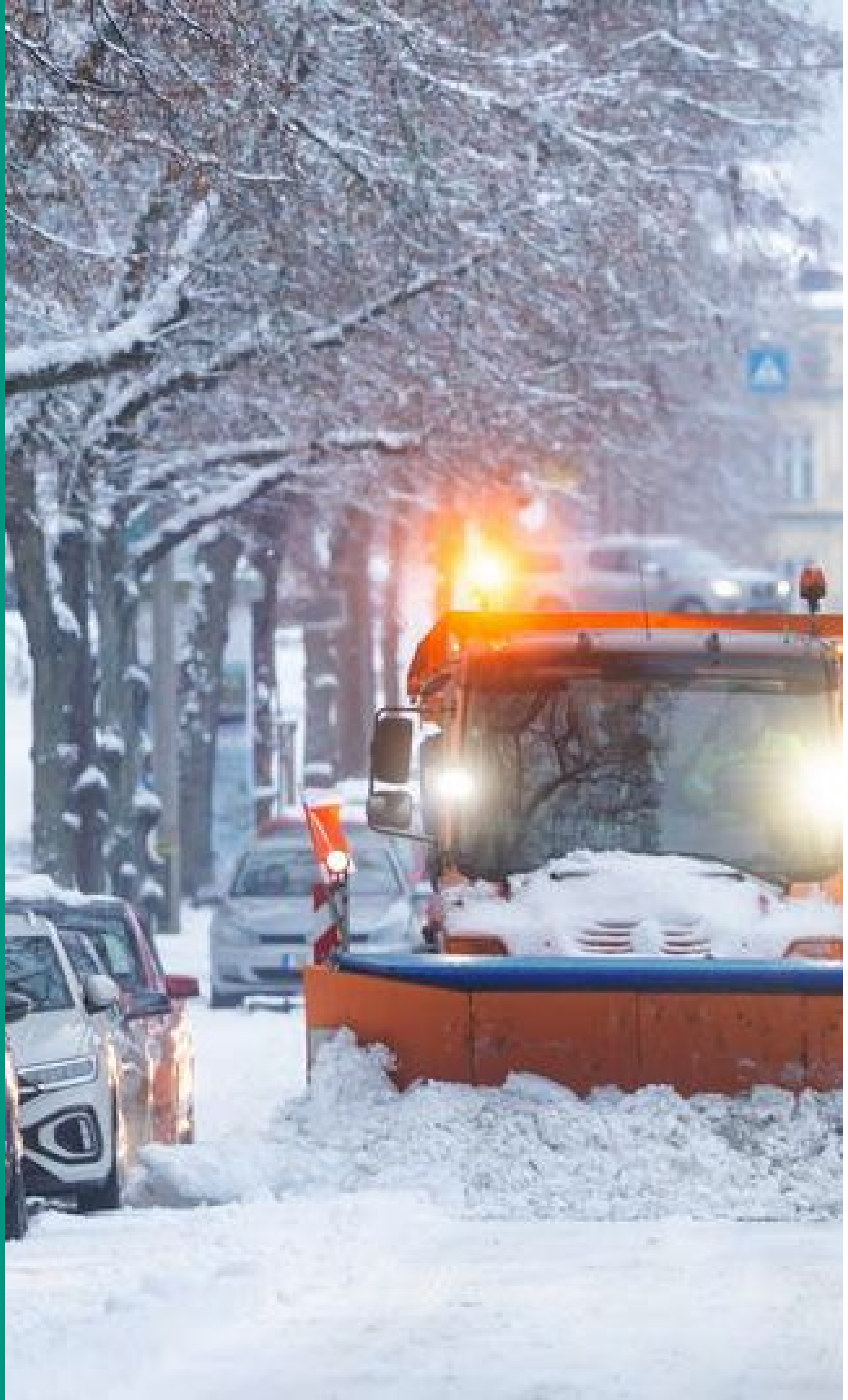
Probabilistic GAN for European Winter Storm Downscaling

Athul Rasheeda Satheesh[1], Kim Stadelmaier[1], Ľuboš Sokol[2], Lea Eisenstein[1], Patrick Ludwig[1], Joaquim Pinto[1], Lukas Braun[2], Alexandros Georgiadis[2]

1) IMKTRO, Karlsruhe Institute of Technology
2) Aon Impact Forecasting



Motivation

Why are we interested in Winter storms?

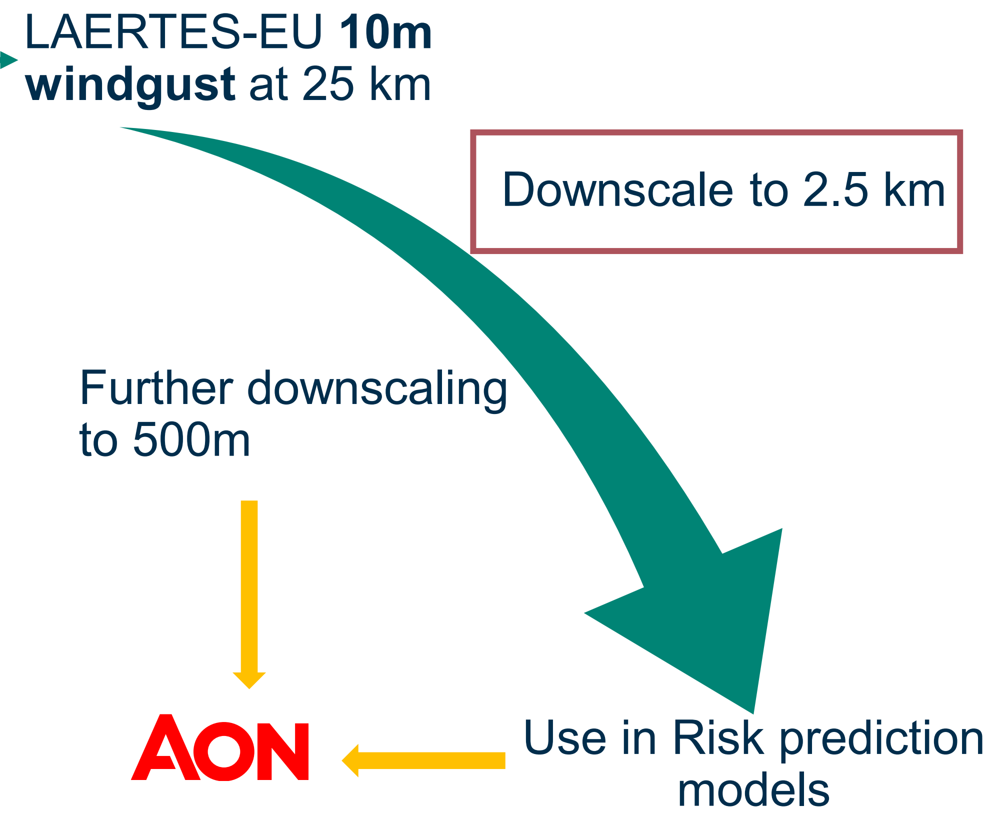
In Europe, winter storms severely affect densely populated regions, mainly urban areas.

Need good risk prediction!

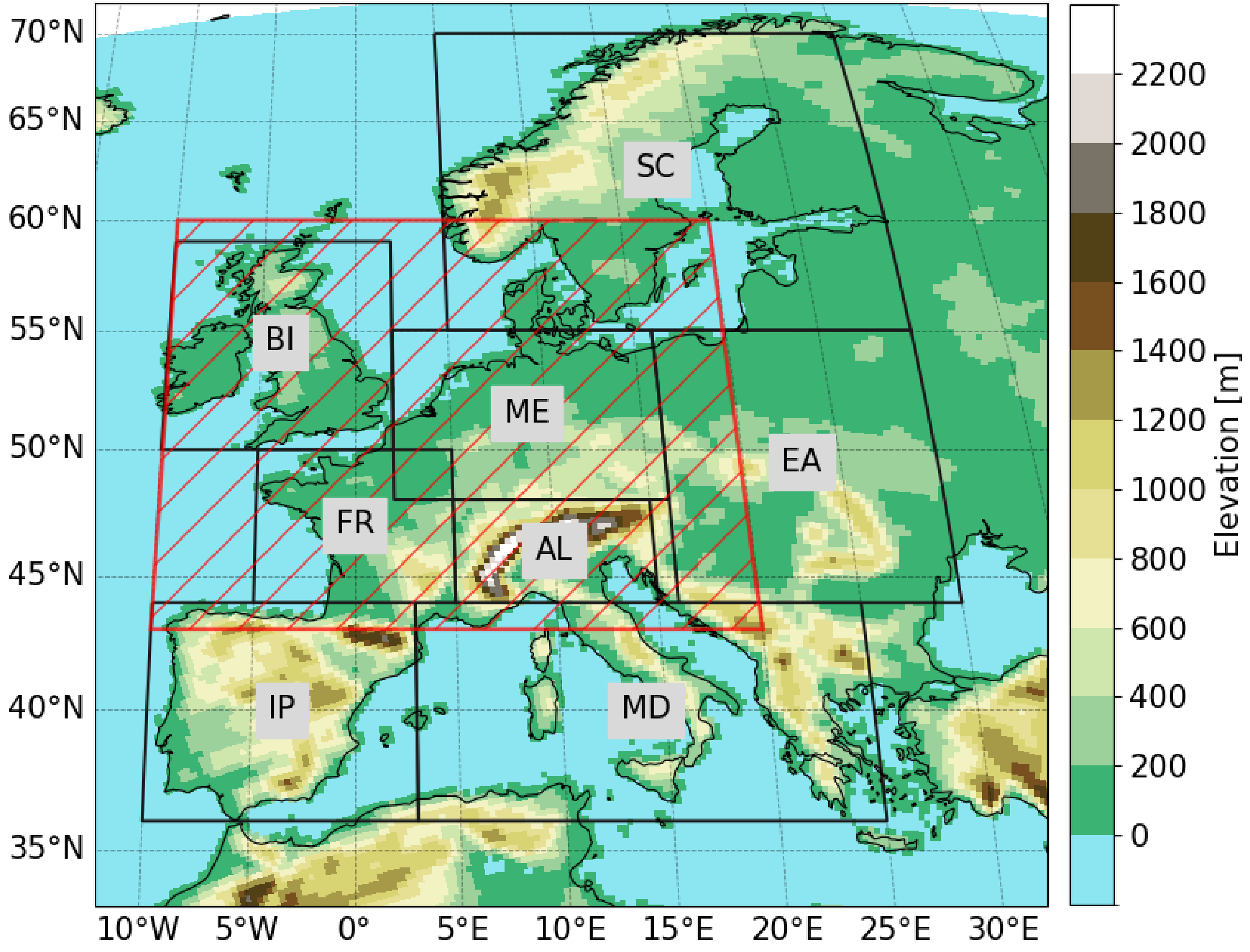
Not enough winter storms in observational data!

The “**LA**rge Ensemble of **R**egional clima**T**e mod**E**l **S**imulations for **EU**rope” (**LAERTES-EU)**

- Stochastic decadal predictions
  - Over 12,000 years of simulated years ✓
  - 5 ensemble members ✓
  - ~25km resolution X



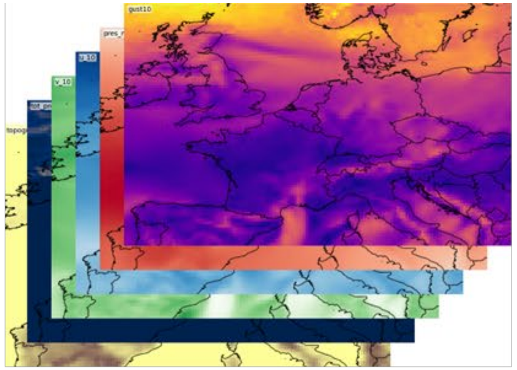
Domain



- *Source: Ehmele et al. 2026*

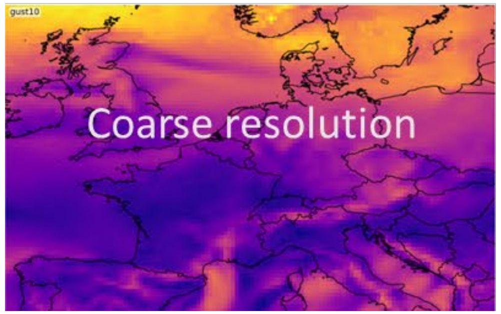
AI-Based (Statistical) Downscaling

Predictors



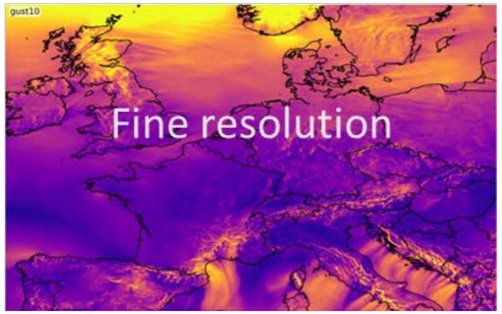
- ❑ 10m wind gust,
- ❑ MSLP + gradient of MSLP
- ❑ 10m u, v
- ❑ Topography
- ❑ Land-sea mask

Input



ML Model

Output



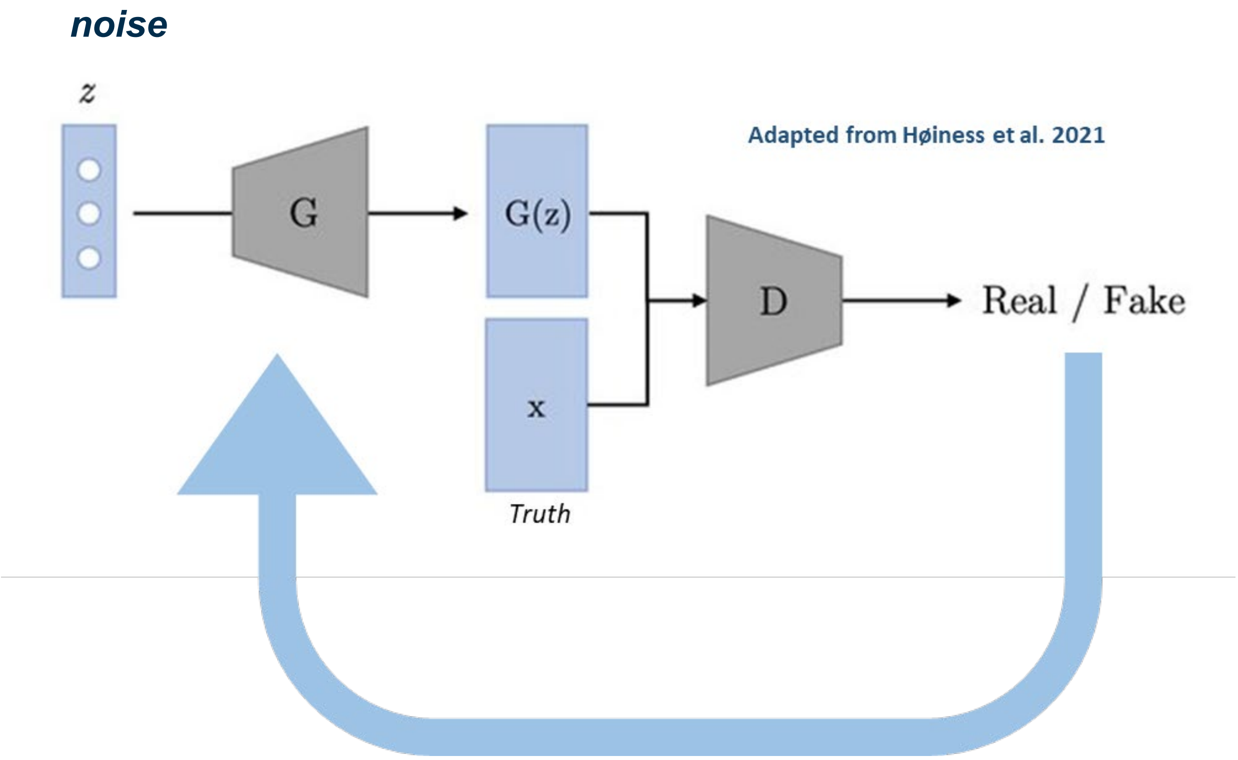
Generative **A**dversarial **N**etwork (**GAN**)

Generative Adversarial Network

Two competing models in **adversarial** setup:

- **Generator(G)**: Generates outputs
- **Discriminator (D)**: Classifies the outputs as Real: **1**/Fake: **0**.

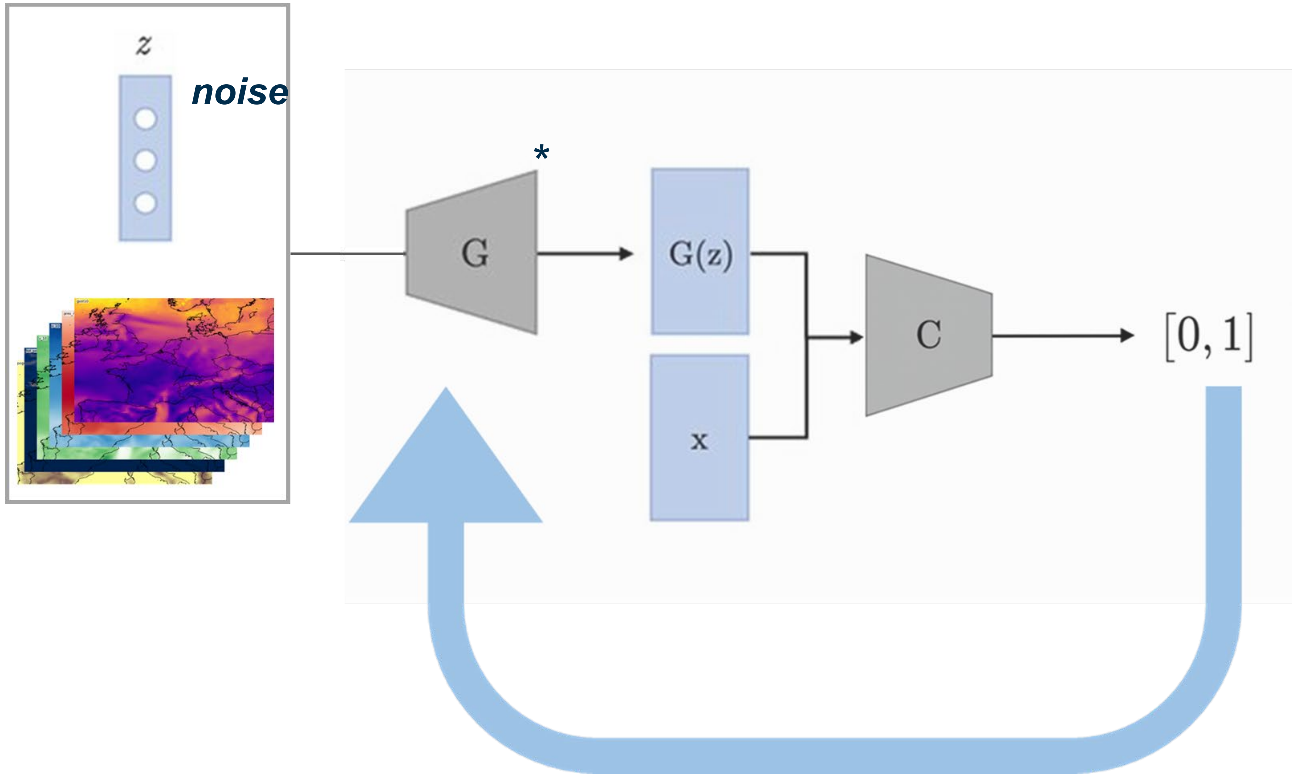
- **Variant Used: Modified Wasserstein GAN**



Wasserstein Generative Adversarial Network

Two competing models in **adversarial** setup:

- ❑ **Generator(G)**: Generates outputs
- ❑ **Critic (C)**: Scores the output between 0 to 1.

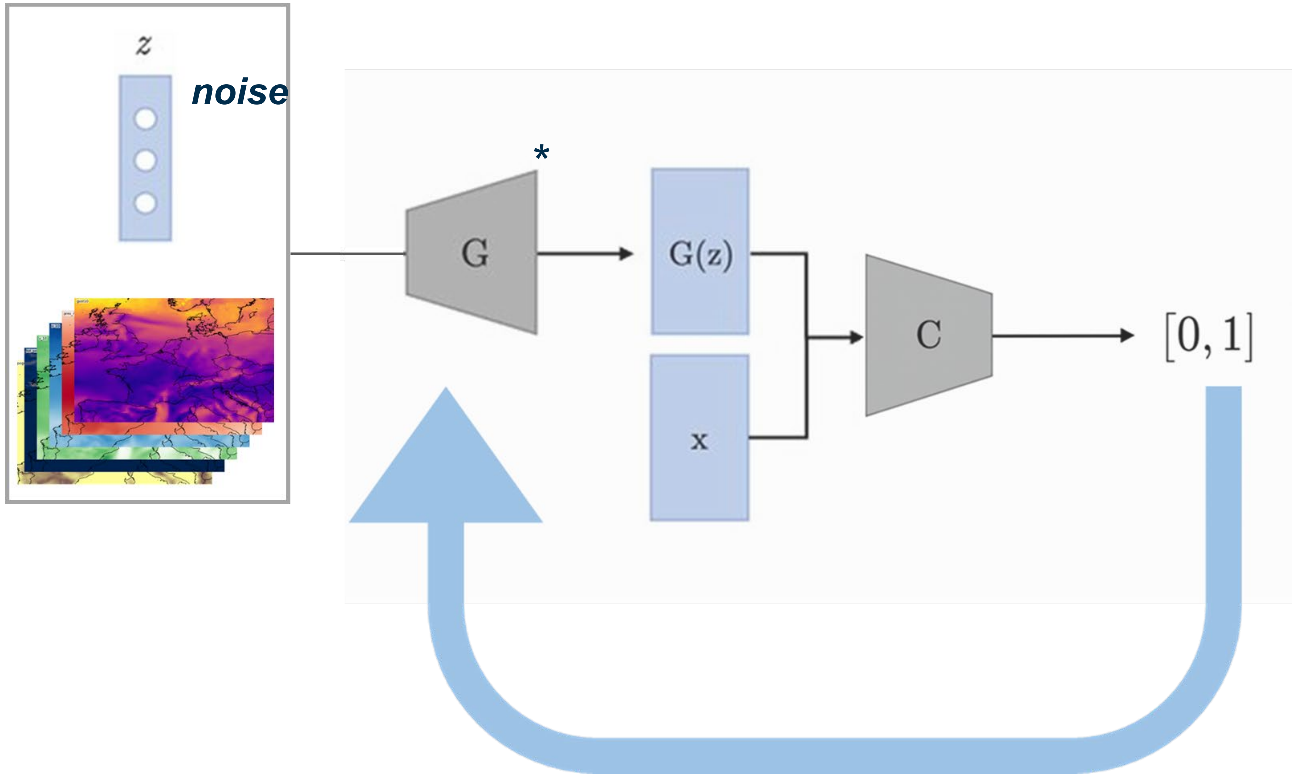


Data and Benchmarking

- ERA5 predictors
- High res. reference: 10m wind gust (x) **ICON simulations**
  - 300 winter storms (225 train, 75 test)

- Benchmarking:
  - Bilinear Interpolation (baseline, only for skillscore)
  - **Linear**
  - **UNET (modified)**
  - **WGAN Deterministic (modified)**

- Final: **Downscale LAERTES-EU wind gusts**



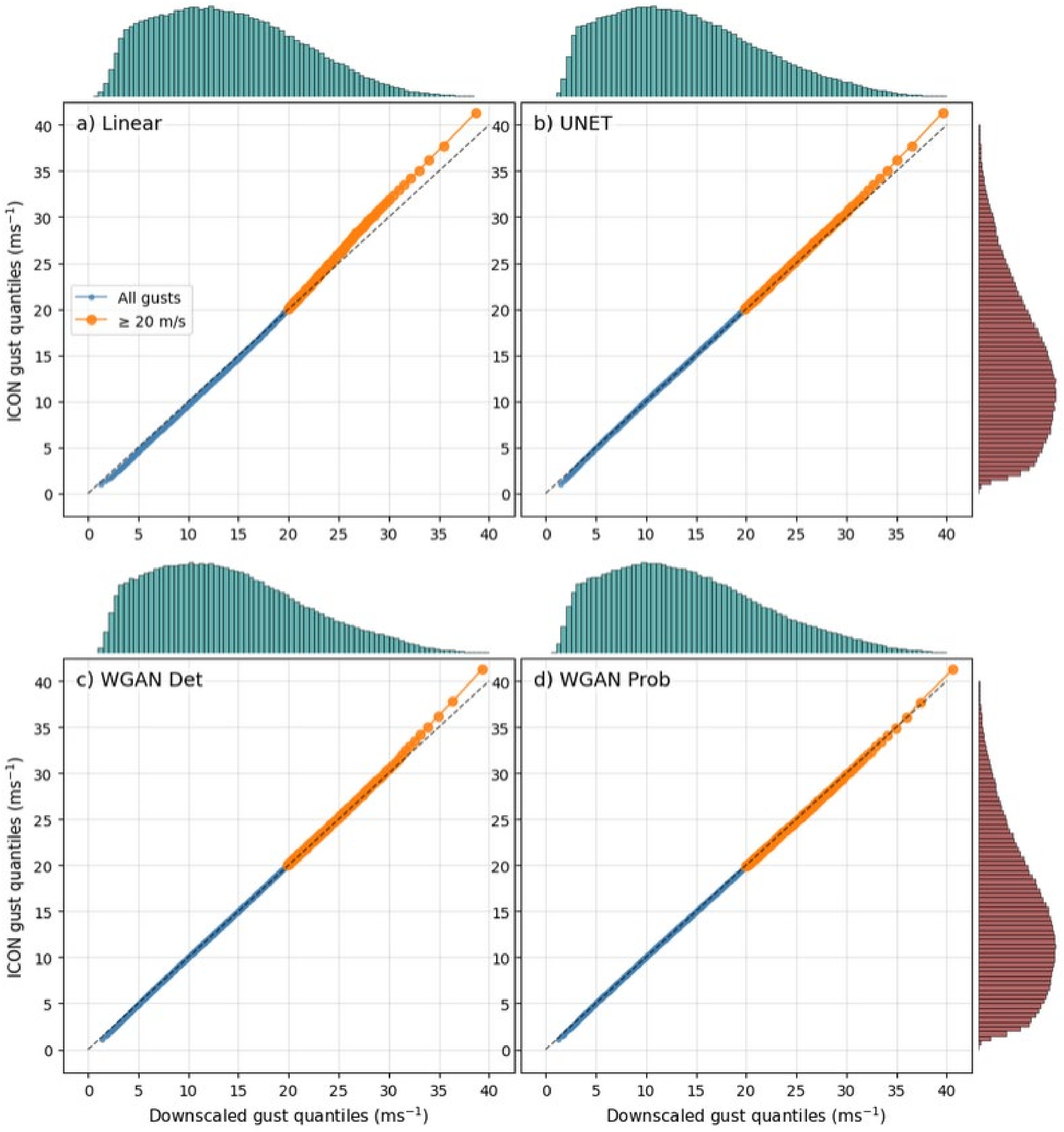
Results I
ERA5

Quantile-Quantile

- 0-20 ms^{-1} (Low - Moderate)
- 20+ ms^{-1} (High - Extreme)

- All models show good skill in Low – Moderate gust regime.

- Large inter - model deviations are observed for High – Extreme gusts.



Distribution

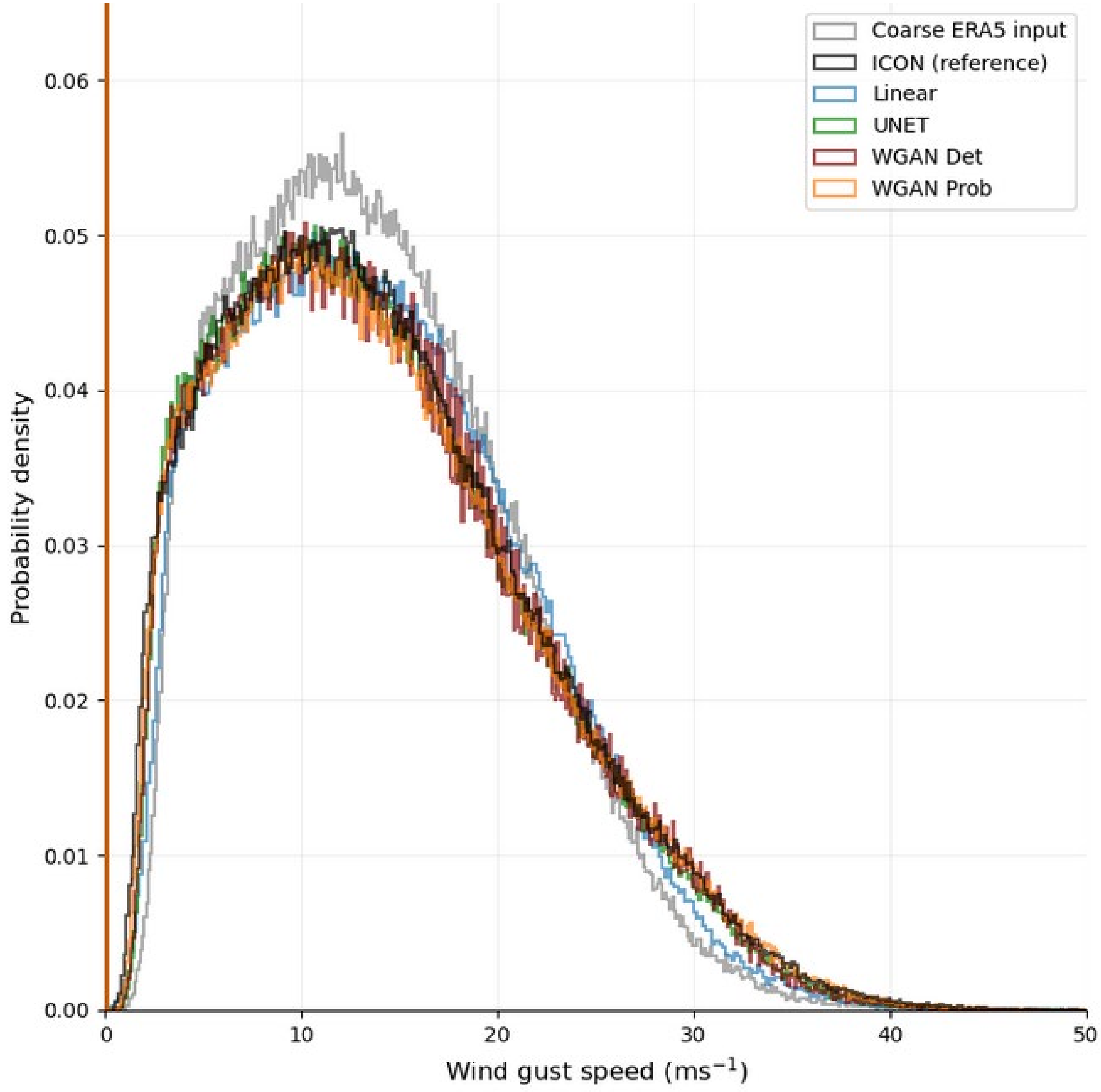
- 0-20 ms^{-1} (Low - Moderate)
- 20+ ms^{-1} (High - Extreme)

- All models show good skill in Low – Moderate gust regime.

- Large inter - model deviations are observed for High – Extreme gusts.

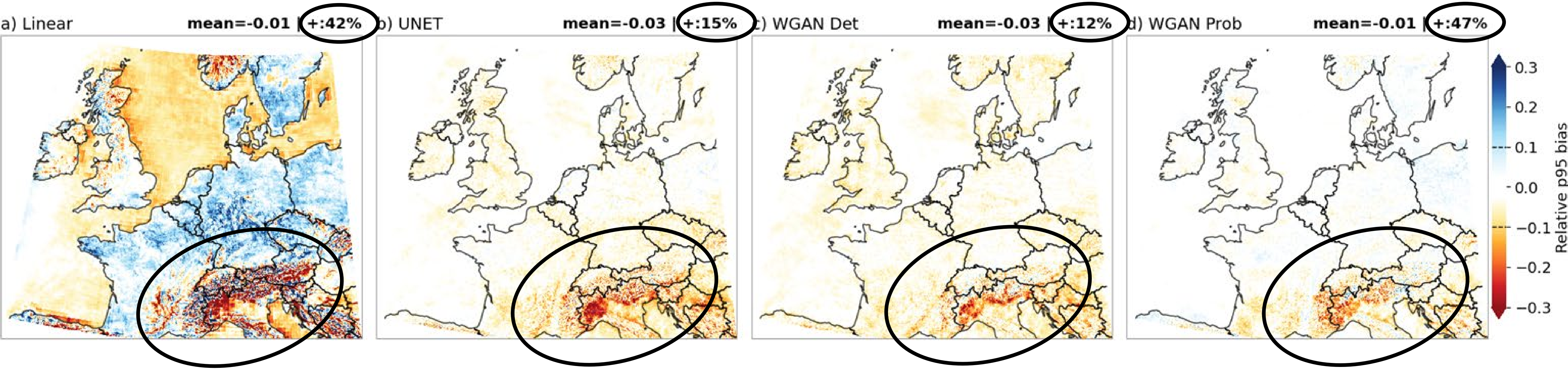
- Linear model closely resembles coarse ERA5.

- High-Extreme gusts cause most winter-storm related damages.



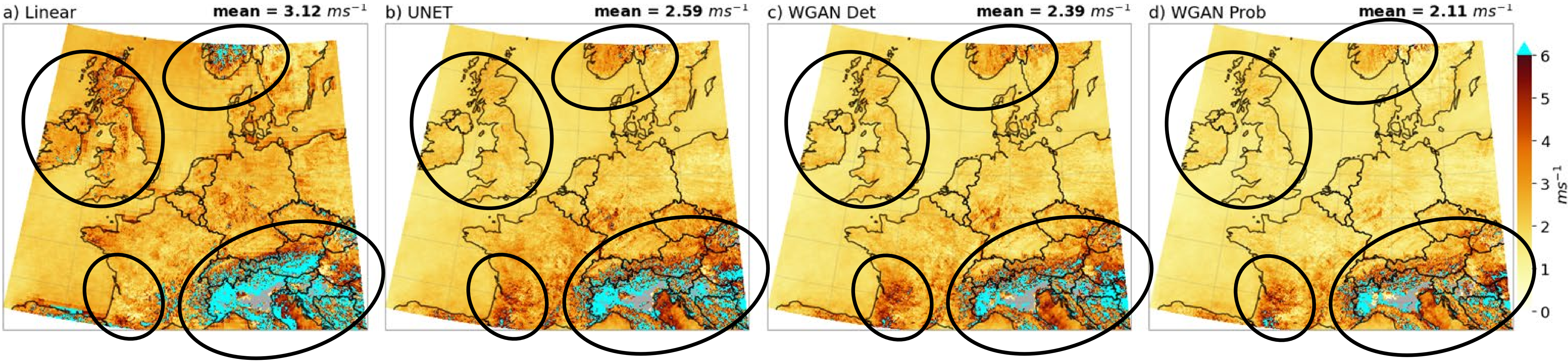
KIT

Relative Bias: 95th percentile



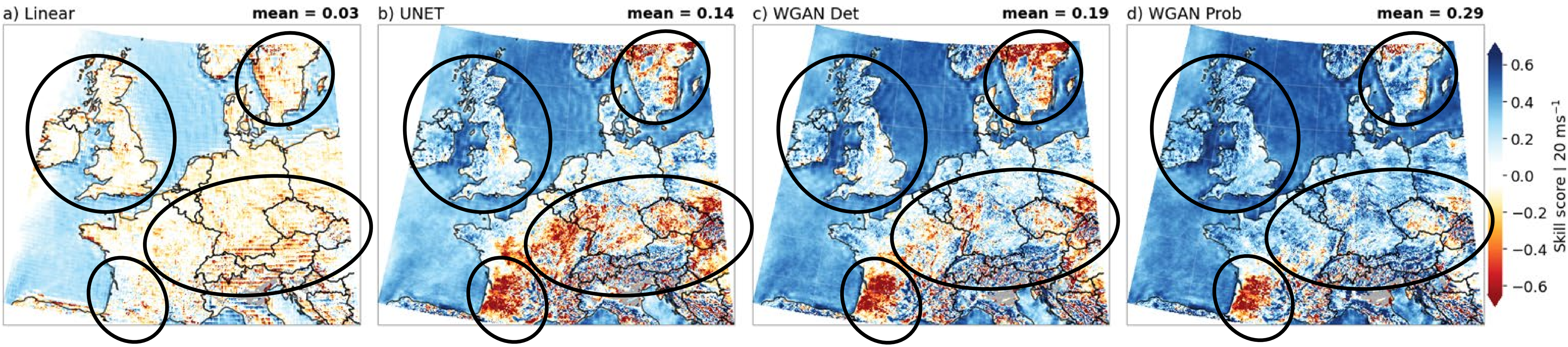
- Clear negative bias over the Alps in all models.
- High positive bias over land in Linear.
- Very low negative bias in UNET and WGAN Det.
- Mix of very low positive and negative bias in WGAN Prob.

Absolute Error: High-Extreme gusts



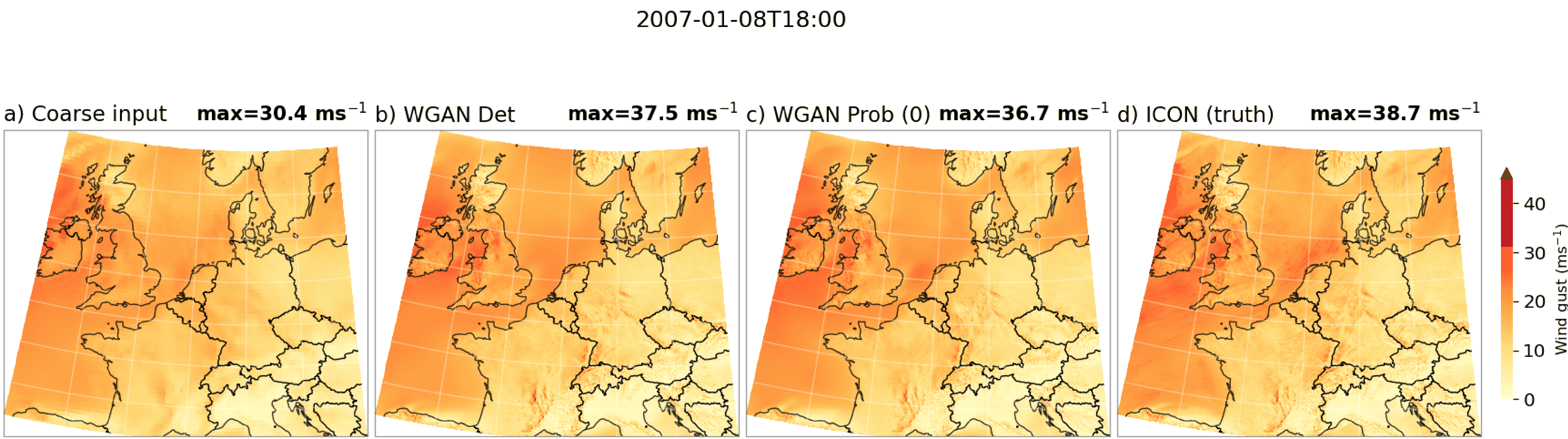
- Highest errors over the Alps.
- High errors in southern Norway.
- Linear Model shows high error over the UK.
- Generative models show high error in southern France, Linear model performs better.
- WGAN Prob shows the lowest over all error.

Skill Score: High-Extreme gusts

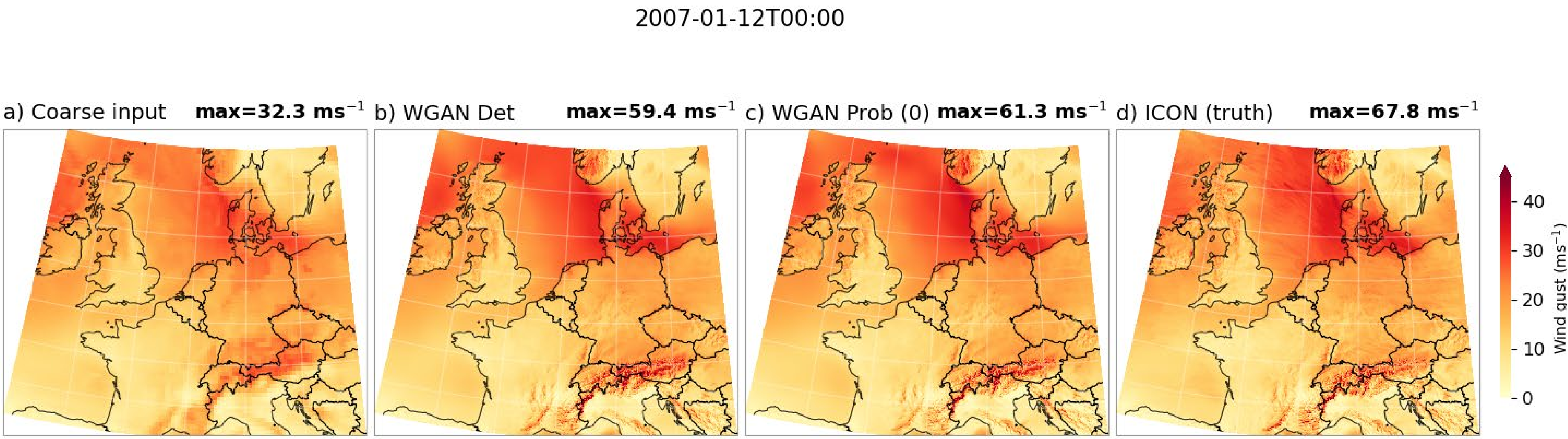


- Shows skill of each model relative to **Bilinear Interpolation (baseline)**.
- Very low skill in Generative models over southern France, Linear model shows slight positive skill.
- Positive skill all over the UK in WGAN Prob, no skill in Linear model.
- Positive skill over southern Scandinavia in WGAN Prob., UNET and WGAN Det struggle over southern Sweden.
- WGAN Prob shows good performance over mainland CE, no skill in Linear, UNET and WGAN Det. show high regional variabilities.

Example Storm 1 in ERA5: Kyrill



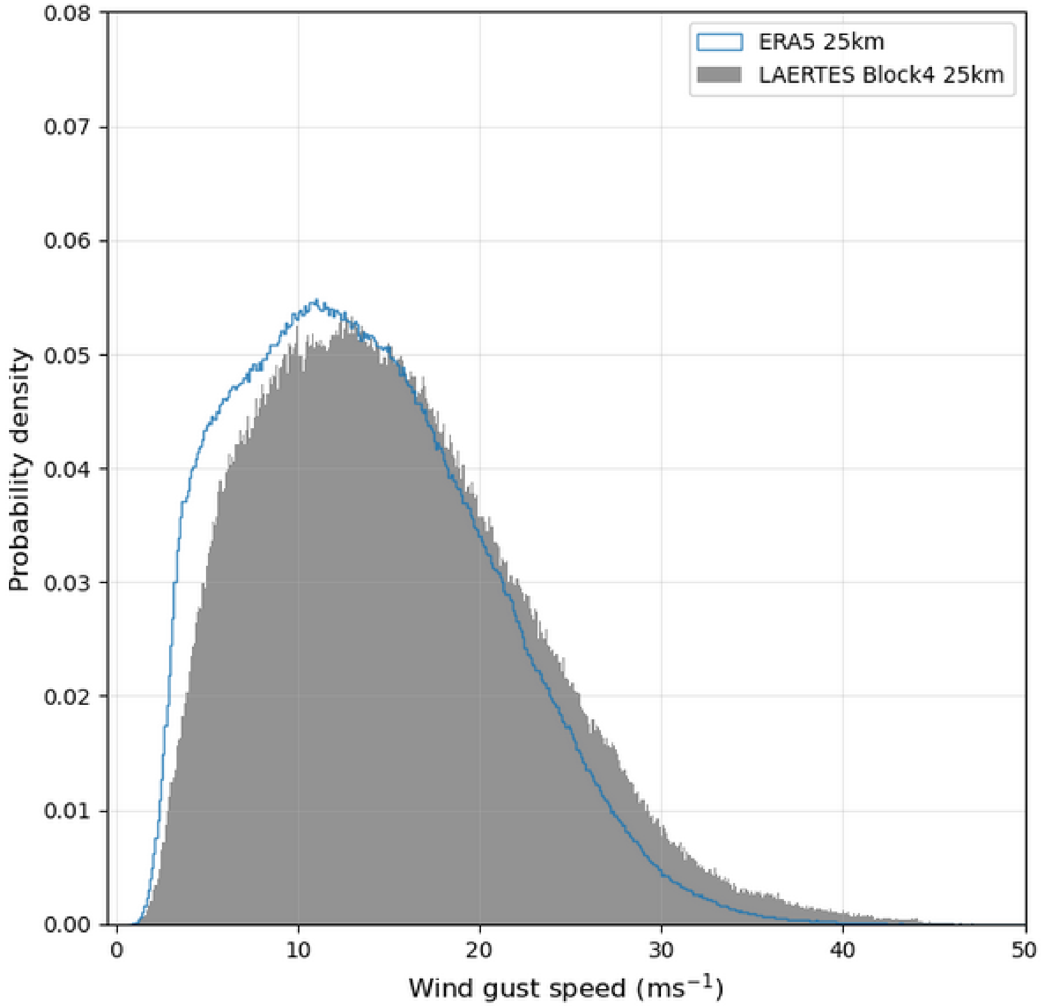
Example Storm 1 in ERA5: Kyrill



Results II
LAERTES-EU

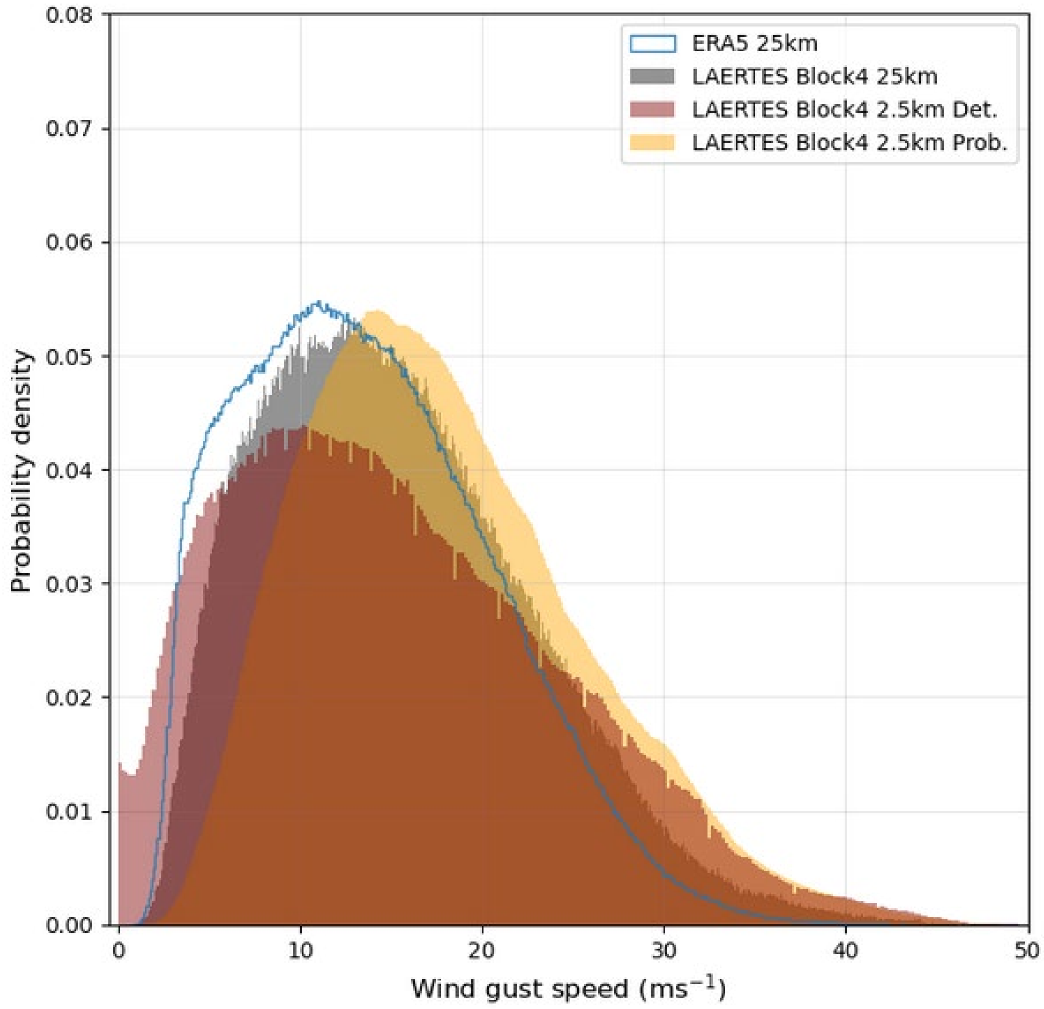
Distribution

- ERA5 and LAERTES have very similar distributions.
- LAERTES is slightly right-shifted:
  - Slightly more high-extreme gusts, fewer low gusts.

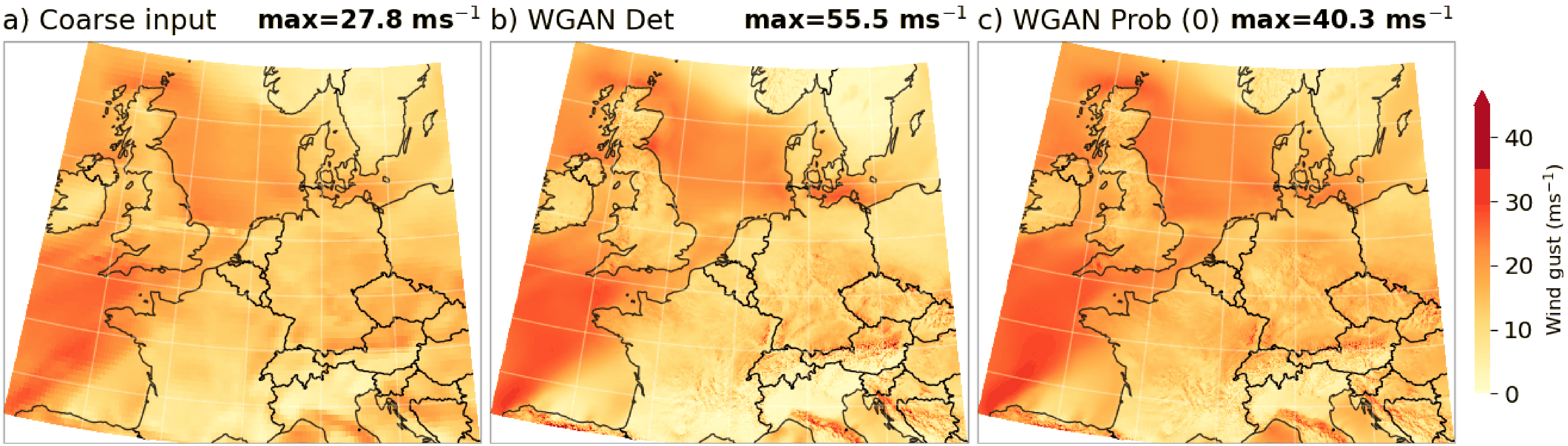


Distribution

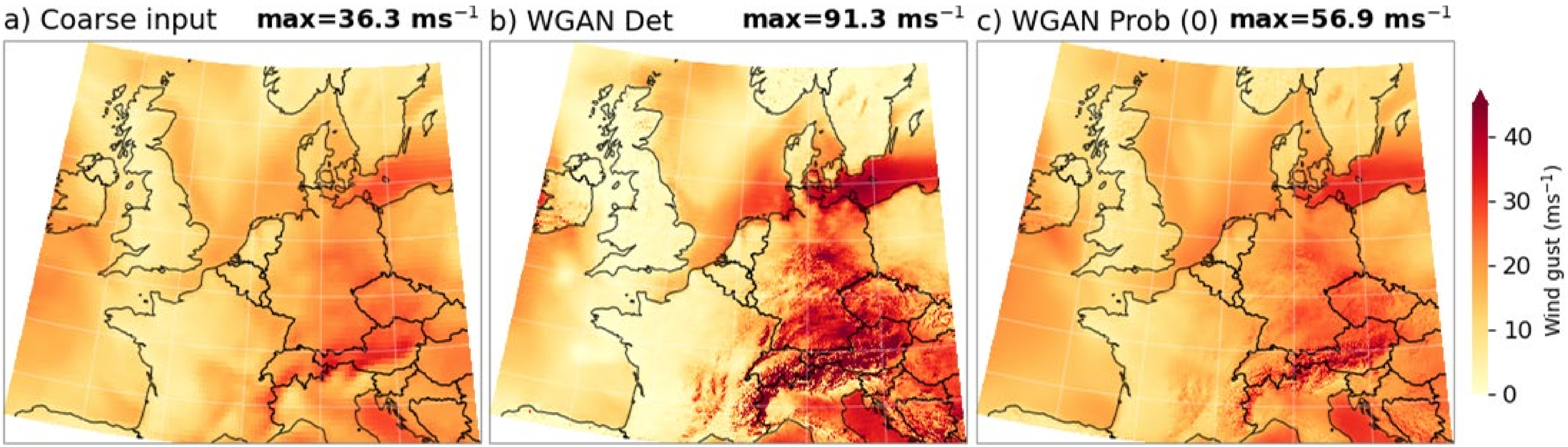
- ERA5 and LAERTES have very similar distributions.
- LAERTES is slightly right-shifted:
  - Slightly more high-extreme gusts, fewer low gusts.
- WGAN Det. severely underestimates moderate gusts, overestimates very low gusts and extreme gusts.
- WGAN Prob. follows similar distribution as LAERTES but slightly right-shifted:
  - more extreme gusts and fewer low gusts



Example Storm 1 in LAERTES-EU



Example Storm 1 in LAERTES-EU



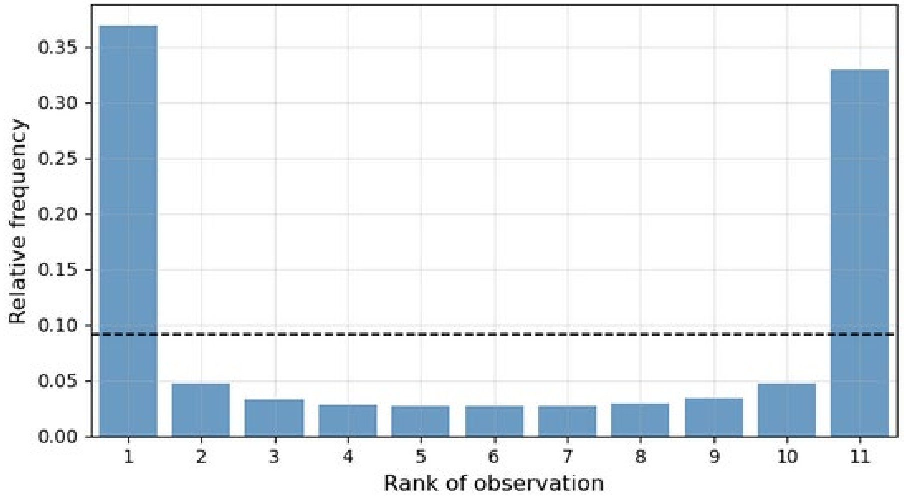
Summary

- A Probabilistic Wasserstein GAN is trained on ERA5-ICON simulation pairs of 300 Winter storms.
  - Realistically downscales storms to 2.5km resolution, capturing important features and high gusts.
  - Outperforms even established generative models like UNET.

- WGAN Prob downscales LAERTES-EU storms a bit conservatively compared to deterministic counterpart.

Outlook

- WGAN Prob is far too underdispersed: Ens members are too clustered. More work is being done to update the model to mitigate this.



Architecture

1. UNET: (Think of this as bias-correction/ post-processing wrt. ICON sims)
    1. 4 downsample blocks of double convolution (reduce 16x)
    2. Bottleneck layer: 1-strided double convolution (no shape change)
    3. 4 upsample blocks (increase 16x)

Encode-Decoder

2. Upsampling to 100x
    1. 3 layers of 1-strided convolution (1024,512,256)
    2. One layer of downscale 2x (Transpose Conv)
        1. Near. Neighbour Upsample LSM 2x
        2. Concatenate with 2x LSM
    3. One layer of downscale 5x (Transpose Conv)
        1. Near. Neighbour Upsample LSM 10x
        2. Concatenate with 10x LSM