Downscaling with the Foundation Model AtmoRep

Michael Langguth[1], Christian Lessig[2], Martin G. Schultz[1], and Ilaria Luise[3]

[1] Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, Germany,
[2] ECMWF, formerly Otto von Guericke University Magdeburg, Germany
[3] CERN, IT Innovation, Switzerland

1) Motivation

- High-resolved weather data is of great relevance for industry, society and for risk quantification of hazardous events
- Recent success of deep neural networks (DNN) for statistical downscaling
  - ✓ Computationally cheaper than dynamical downscaling → large ensembles
  - ✓ Accurate and plausible reconstruction of small-scale features
  - x Expensive training for a single application
  - x Pairing of coarse- and high-resolved data often limits data availability
- Foundation Models (FMs) for weather and climate such as ClimaX (*Nguyen et al., 2023*) and **AtmoRep** (*Lessig et al., 2023*) are potentially appealing
  - Self-supervised pre-training on large datasets
  - Scale well with data, compute and network size
  - Wide-range of downstream applications with SOTA performance after fine-tuning

Research question:
Can we push the frontiers of downscaling with Foundation Models?

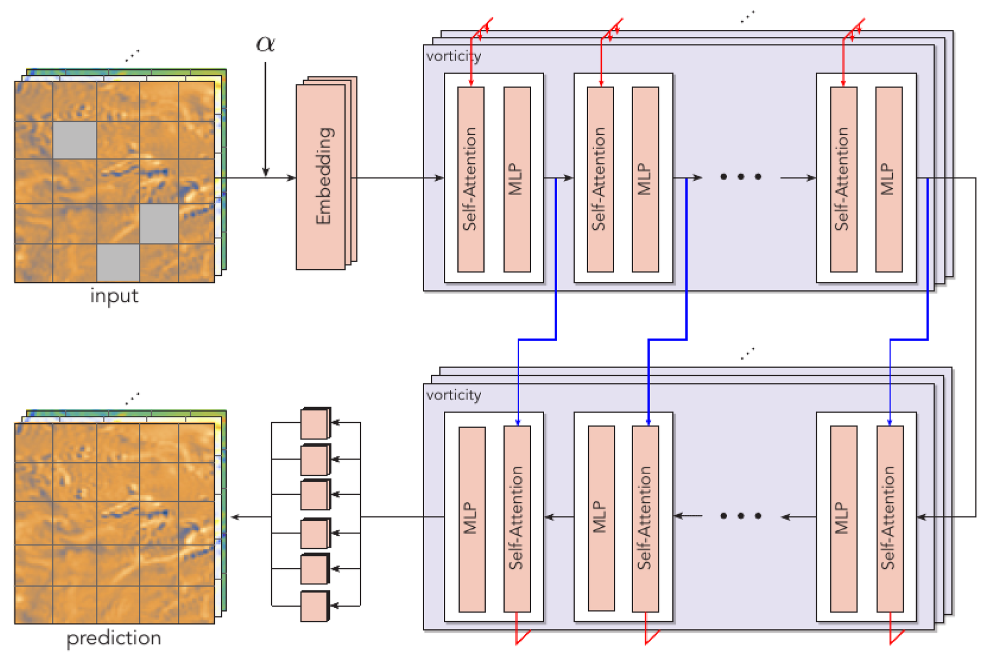
2) The AtmoRep model

- Modelling of the atmosphere as a stochastical system

$$p(\mathbf{y}|\mathbf{x}) \approx p_\theta(\hat{\mathbf{y}}|\hat{\mathbf{x}}, \boldsymbol{\alpha})$$

$\mathbf{y}, \hat{\mathbf{y}}$: dependant (model) state
$\mathbf{x}, \hat{\mathbf{x}}$: conditioning (model) state
$\boldsymbol{\alpha}$: auxiliary information

- **The network**



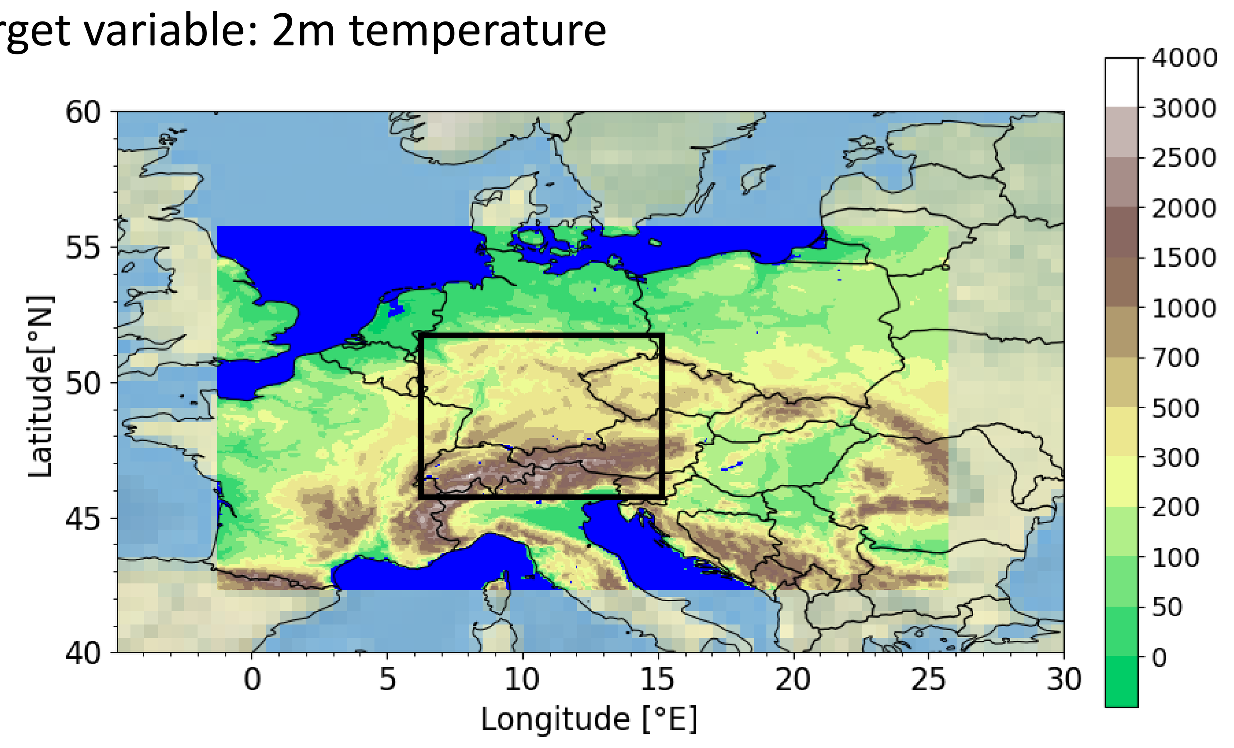
- Trained with ERA5 reanalysis data (1979 – 2017)
- Variables: $T, q_v, \mathbf{v}_h, w$ on model levels [96, 105, 114, 123, 137] + $precip$
- Modular AtmoRep configuration with two-step training (singleformer → multiformer)

Want to learn more about AtmoRep?
Attend the AtmoRep talk: ***ESSI 1.1, 16:30 – 16:50 CEST***

3) Downscaling with AtmoRep

The downscaling task

- Downscale ERA5-data ($\Delta x_{ERA5} = 0.25°$) to (re-projected) COSMO-REA6 data ($\Delta x_{CREA6} = 0.0625°$) → downscaling factor: 4
- Target downscaling domain (216 x 432 grid points) over Central Europe (Fig. 2) Paired training data between 1995 and 2017
- Test data from 2018
- Target variable: 2m temperature



Model setup for downscaling with AtmoRep

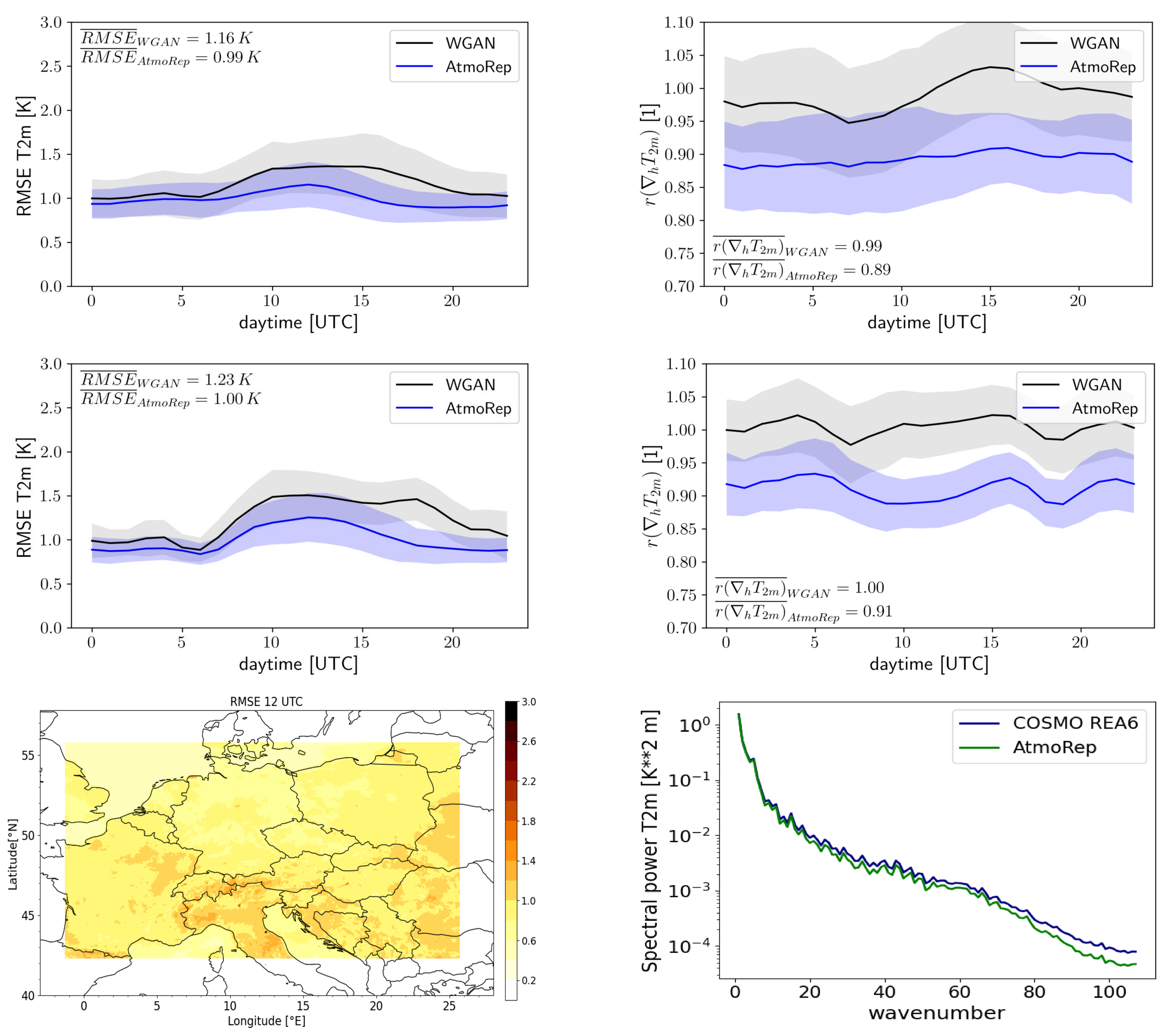
- Use three-field AtmoRep-configuration ($[\mathrm{u}, \mathrm{v}, T]$) with data from model level 137 only
- Tail network appended to AtmoRep core model:
  - Increased token size at the beginning of tail network
  - Further embedding layer with doubled embedding dimension and updated positional encoding
  - Six transformer layers for downscaling
  - Ensemble tail for probabilistic output
- Trainable core model and tail network (→ very deep network)

Competing model

- Train a competing Wasserstein GAN to contextualize AtmoRep results
- Generator model of WGAN: U-Net by *Sha et. al, 2020*, standard ConvNet for critic model
- Smaller target domain (96x120 grid points) due to shift invariance of CNNs
- Training data:
  - T from model levels [96, 105, 114, 123, 137]
  - $\mathbf{v}_h$ from model level [137]

4) Results

Point-wise evaluation | **Evaluation of spatial variability**



5) Conclusion and outlook

- Downscaling with AtmoRep has the potential to outperform existing solutions
- RMSE of initial AtmoRep downscaling model matches that of WGAN trained with comprehensive set of predictors (not shown)
- Very deep neural network (1.85B parameters) → barely utilizes FM benefits

Next steps:

- Improve set-up of architecture (e.g. diffusion model) with prob. downscaling
- Analyse added value of FM approach
- Multivariate downscaling and application to arbitrary region

The project is funded by the European Union's Horizon 2020 research and innovation programme and the EuroHPC Joint Undertaking under Grant Agreement 955513 and co-funded by the German Ministry of Education and Research (BMBF) under funding reference 16HPC029.
The authors have used computation resources of the supercomputers JUWELS and Juwels Booster at JSC.

References:
[1] Lessig et al., 2023, *arXiv preprint arXiv:2308.13280*
[2] Nguyen et al., 2023, *arXiv preprint arXiv:2301.10343*
[3] Sha et al., 2020, J. Appl. Meteorol. Clim, 59.12

AtmoRep-paper

MAELSTROM report