A Benchmark Dataset for Meteorological Downscaling

Michael Langguth, Ankit Patnala (Forschungszentrum Jülich), Sebastian Lehner, Markus Dabernig, Konrad Mayer, Irene Schicker (GeoSphere Austria), Paula Harder (Mila)

Motivation

The effective response to climate change necessitates accurate observation and modeling of weather and climate, utilizing **high-resolution data** to comprehend global climate patterns and facilitate the transition to renewable energy sources like **wind and photovoltaic production**, particularly to achieve a fossil fuel-free status by 2040.

Downscaling in climate science, similar to super-resolution in computer vision, offers an efficient method to infer local, high-resolution quantities from the coarser scale variables.

Benchmark datasets like WeatherBench or ClimateBench have helped to improve and unify fields like DL for weather forecasts or climate projections. The field of deep learning for downscaling though is still **lacking a dataset and framework** to compare and benchmark promising architectures. A downscaling benchmark dataset will aid both research and deployment.

Downscaling Task & Data

The benchmark dataset consists of three downscaling tasks for the generation of high-resolved data fields of **2m temperature, 100m-wind, and global horizontal irradiance** generated from a coarse-grained set of predictor variables.

Predictor (input) variables with Δx_{ERA5}=0.25° are obtained from the **ERA5 reanalysis** dataset (Hersbach et al., 2020). Predictands (targets) with Δx_{CREA6}=0.055° come from the **COSMO-REA6 data** (Bollmeyer et al., 2015). The downscaling domain is located in the southern parts of Central Europe, involving the Alpine region and parts of the Mediterranean Sea. As part of the COSMO-REA6 domain, it comprises 144x128 grid points in (rotated) zonal and meridional direction. The netCDF-datafiles can be downloaded via a climetlab-plugin (https://climetlab.readthedocs.io/en/latest/) and are ready-to-use, i.e. no further processing by the user is required.

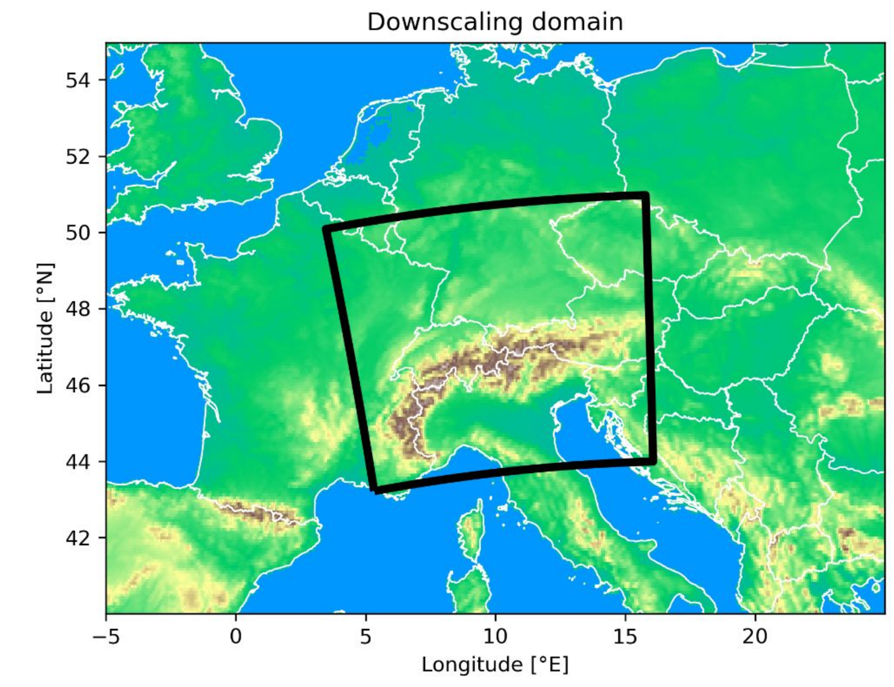

Fig. 1: The downscaling target domain.

Predictors for 2m temperature (ERA5)
- 2m temperature
- temperature at diff. model levels
- 10m (u,v)-wind components
- surface latent
- sensible heat flux
- surface pressure
- static predictors: surface topography and land-sea mask from ERA5 and COSMO REA6

Tab. 1: Predictor variables for the 2m temperature downscaling task.

Baseline neural networks

CNNs, GANs, Vision Transformers and recently diffusion models have been applied for statistical downscaling. Here, we include five deep learning baselines:

- **U-Nets**: Two variants are included, the U-Net by Sha et al., 2020, and the DeepRU by Höhlein et al., 2020. U-Nets are popular CNNs using skip connection to keep high-resolution context.
- **WGANs:** Wasserstein GANs comprise a generator and a critic network and are capable to reconstruct fine-scale features. Here, we choose the WGAN by Harris et al., 2022 and a WGAN using the Sha U-Net as generator.
- **SwinIR:** The transformer-based SwinIR-model by Liang et al., 2021, uses shifted windows to process the tokenized data and was shown to provide SOTA results for single-image super-resolution.

Link to the github-repo!

Classical competitor

As a classical competitor, we choose **SAMOS**, a model based on Standardized Anomaly Model Output Statistics. SAMOS uses non-homogeneous Gaussian regression with standardized anomalies from a spatial climatology (Dabernig, 2017). This removes seasonal and location-specific data characteristics, allowing its application to new data points as required for downscaling.

Evaluation

We suggest a two-step approach:

1. A script-based and task-specific model evaluation with various plots for analysis, e.g. energy spectra and conditional quantile plots.
2. An interactive Jupyter-Notebook for model inter-comparison, e.g. in terms of skill scores.

Example

Scores for 2m downscaling:

- RMSE
- Bias
- Mean Error of Standard Deviation
- Integrated Quadratic Distance

First Results

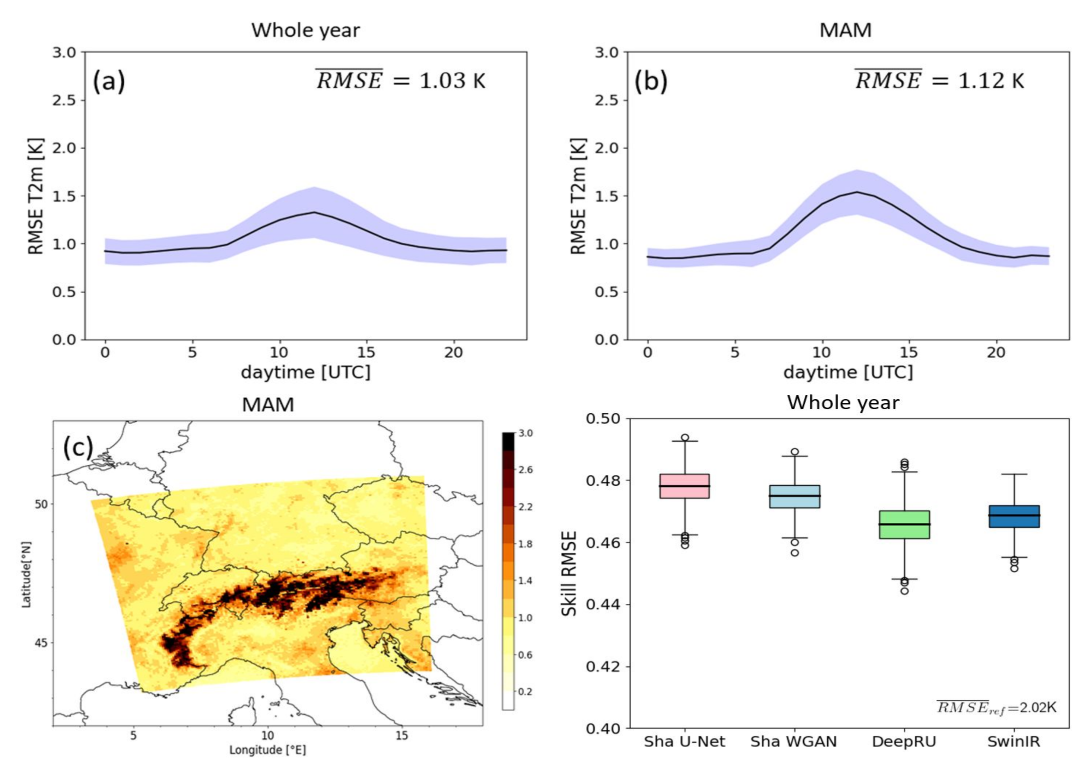

Fig. 2: Exemplary evaluation results of the Sha WGAN model for the 2m temperature downscaling task (a-c). (d) shows an intercomparison between different baseline models.

Conclusion and next steps

We present the first **comprehensive benchmark dataset** for downscaling temperature, wind and solar irradiance. We provide a ready-to-use dataset and a set of baseline models, involving DL approaches and a classical competitor. A comprehensive evaluation framework enables intercomparison between model solutions. Within the next weeks, the dataset, the baseline models and the evaluation framework will be published alongside with a paper submission. For the future, we plan an extension to a probabilistic framework including downscaling of precipitation on kilometre-scale.

References:

[1] Bollmeyer et al., 2015, QJRMS, 141.686
[2] Dabernig, et al., 2017, QJRMS, 143.703
[3] Harris et al., 2022, JAMES, 14.10
[4] Hersbach et al., 2020, QJRMS, 146.730
[5] Höhlein et al.,2020, Met. App., 27.6
[6] Liang et al. 2021, IEEE ICCV
[7] Sha et al., 2020, App. Met. Clim, 59

Part of the work 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.