AICON – Data Driven Forecasts

Martin Köhler, DWD

Lead: Jan Keller, Roland Potthast

AICON-global: Tobias Göcke, Florian Prill, Marek Jacob
AICON-LAM: Sabrina Wahl, Hendrik Reich
ICON reanalysis: Arianna Valmassoi, Jan Keller

ICON reanalysis

ICON-DREAM, ICON-FORCE

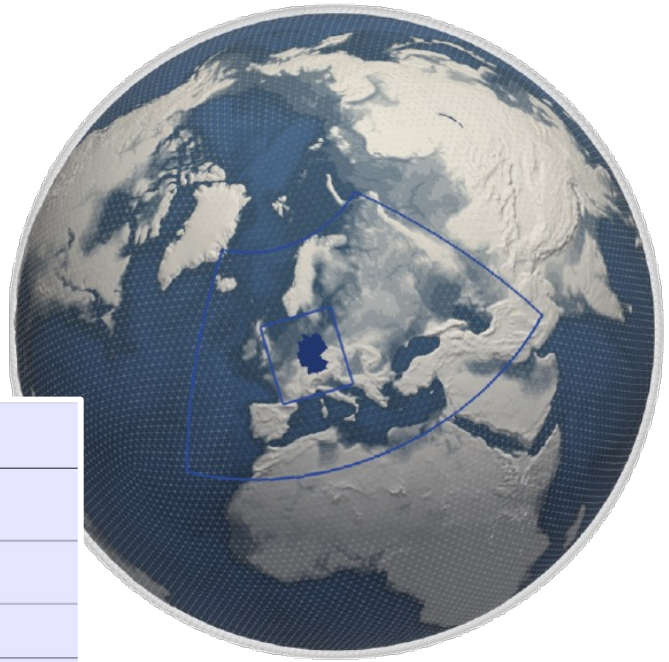
ICON DREAM: a new global reanalysis

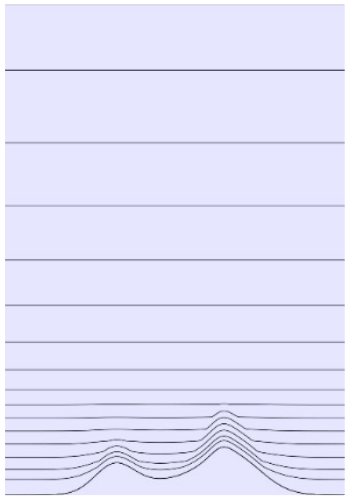
The data-driven AICON-Global is based on the ICON reanalysis dataset

ICON DREAM = Dual resolution Reanalysis for Emulators, Applications and Monitoring

- Global
  - deterministic 13 km
  - ensemble 40km (20mem)
- Europe-Nest
  - deterministic 6.5km
  - ensemble 20km (20mem)
- (current) reanalysis time range: 2010-01 until 2025-04
- storage size of dataset: 6.45 PB
- no release yet; a service is planned for 2026

Operational ICON domains (global/regional)

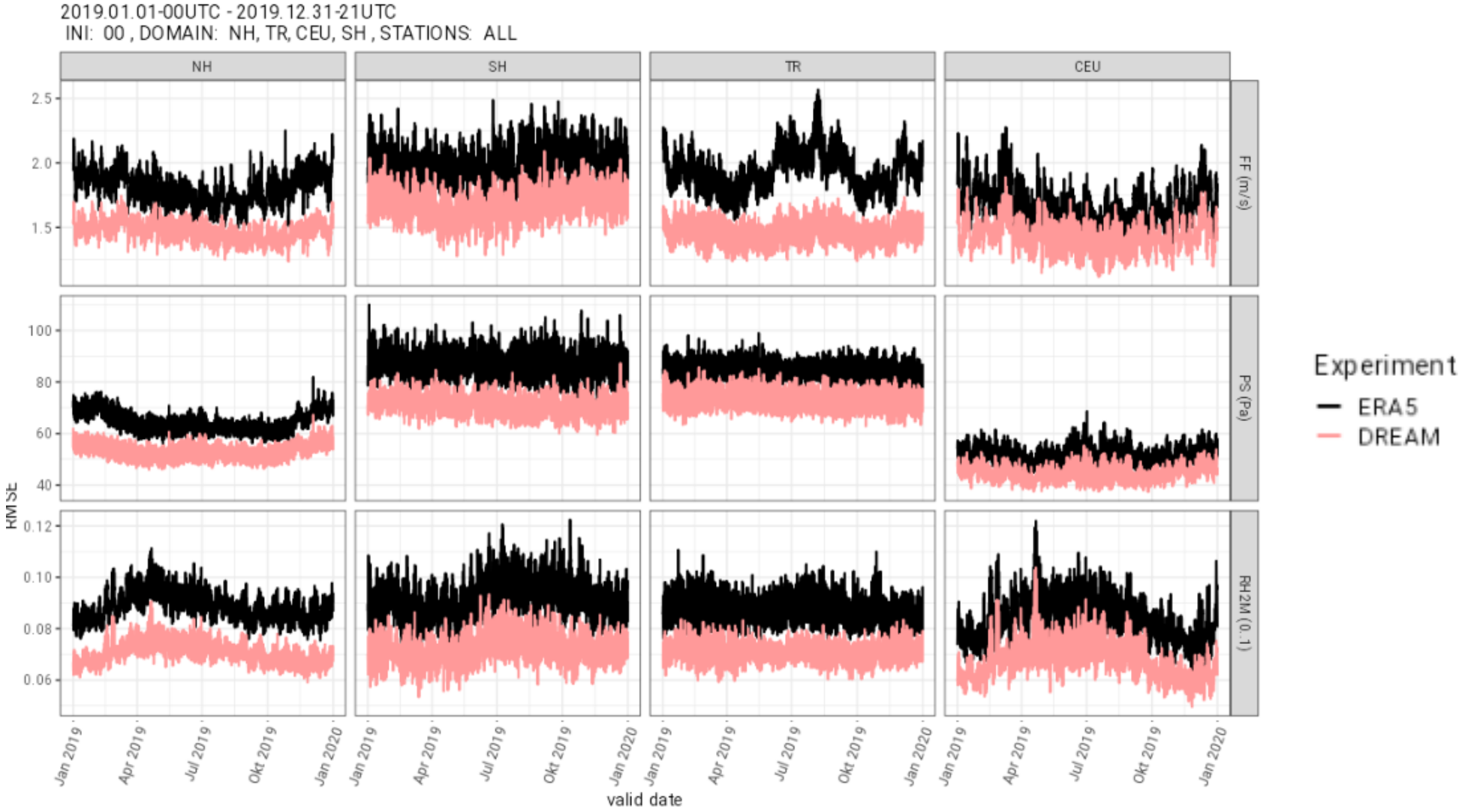




ICON Smooth Level Vertical (SLEVE) z-coordinate;
Note: ERA5 uses (hybrid sigma-) pressure levels!

ICON-DREAM: A new dual resolution reanalysis from DWD. 6th WCRP International Conference on Reanalysis. (2024). https://confit.atlas.jp/guide/event/icr6/subject/OR2-01/detail
(A. Valmassoi et al., DWD)

ICON DREAM: a new global reanalysis



ICON-DREAM: A new dual resolution reanalysis from DWD. 6th WCRP International Conference on Reanalysis. (2024). https://confit.atlas.jp/guide/event/icr6/subject/OR2-01/detail
(A. Valmassoi et al., DWD)

ICON regional reanalyse ICON-FORCE

Fine-scale **O**bservation-based **R**eanalysis for **C**entral **E**urope

- 2 km horizontal resolution
- 20 ensemble members
- hourly data assimilation (also radar)
- hourly output or more
- production starting soon

AICON
design

The Anemoi Framework

Collaborative European initiative

- Anemoi plays a key role in the development of multiple ML-powered weather models:
  AIFS (Artificial Intelligence Forecasting System, ECMWF),
  Bris (MetNorway, extends the AIFS) and **AICON** (DWD).
- **DWD abandoned its in-house development in June 2024 in favor of the shared Anemoi codebase.**
- The Anemoi Framework received the EMS Technology Achievement Award 2025.
- Development related to EUMETNET E-AI: Artificial Intelligence and Machine Learning in Weather, Climate and Environmental Application

AICON Model Architecture

Reduced-level Zarr dataset: 29.5 TB

prognostic

```
PS, P[:], T[:], U[:], V[:], QV[:], T_2M, U_10M, V_10M, QV_S, RELHUM_2M, T_G, ALB_RAD,
H_SNOW, SMI[0/1], T_SO[0/1],
```

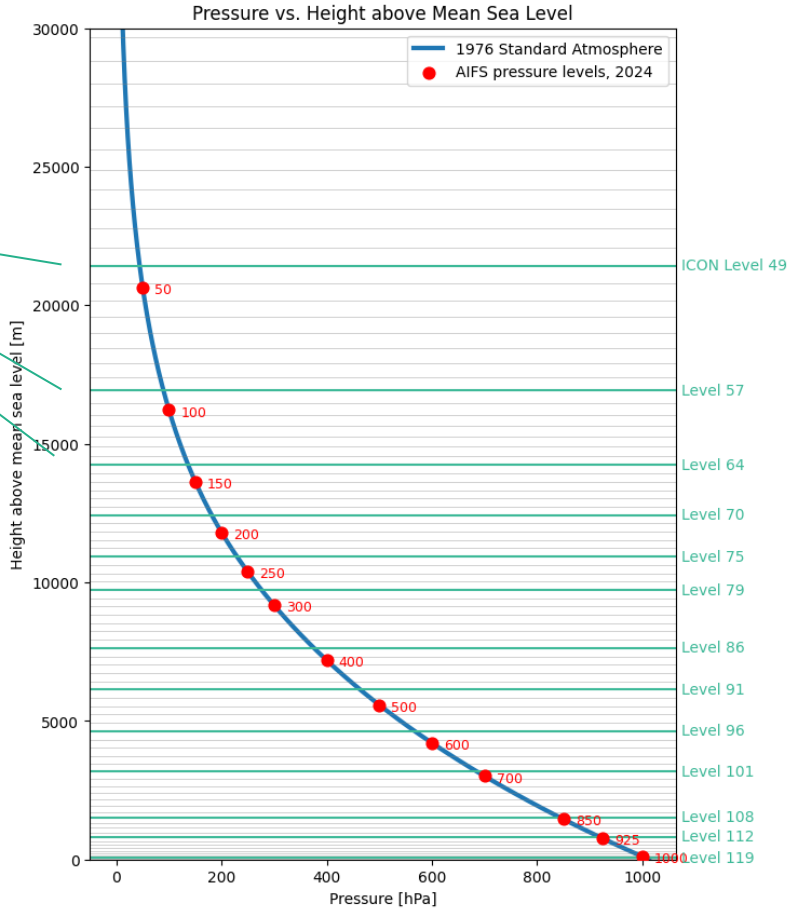
forcings:

```
HSURF, FR_LAND, Z0, FR_LAKE, EMIS_RAD, SSO_STDH, cos/sin_latitude/longitude,
cos/sin_julian_day, cos/sin_local_time, insolation
```

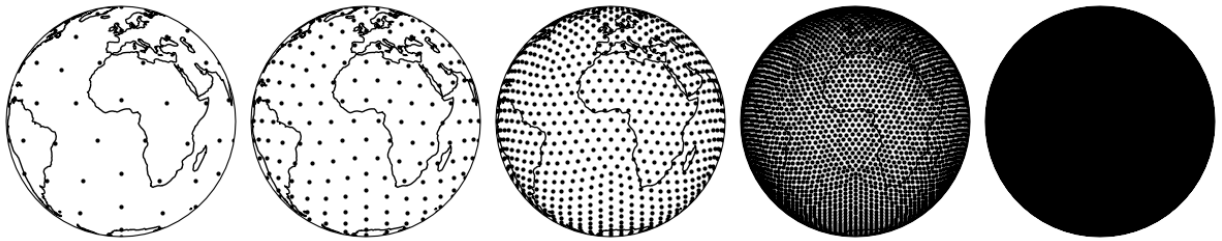
diagnostic:

```
TOT_PREC
```

13 ICON levels (top-down ordering):
49, 57, 64, 70, 75, 79, 86, 91, 96, 101, 108, 112, 119



ICON data locations are based on cell centers of a triangular grid. Grid generation inherently defines hierarchical decompositions:



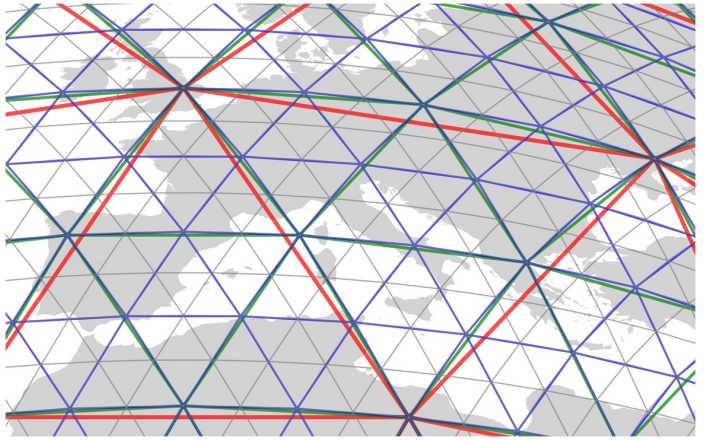
AICON Model Architecture (cont'd)

GraphCast-like encoder-processor-decoder architecture.

encoder: maps onto a hidden state

processor: contains latent information

decoder: takes final hidden state

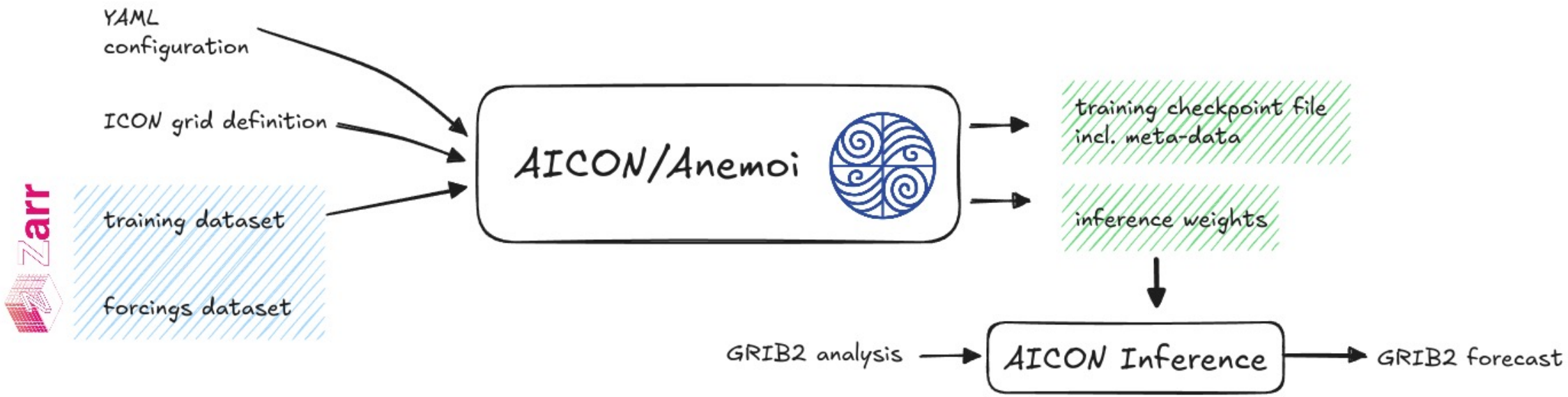


Processor: Different levels of refinement facilitate communication at different scales.

- Graph construction directly based on ICON's triangular meshes.
- Graph-Transformer GNN: Message passing is done via a multi-head attention mechanism. Each node's new features become a weighted average of its neighbors' features.

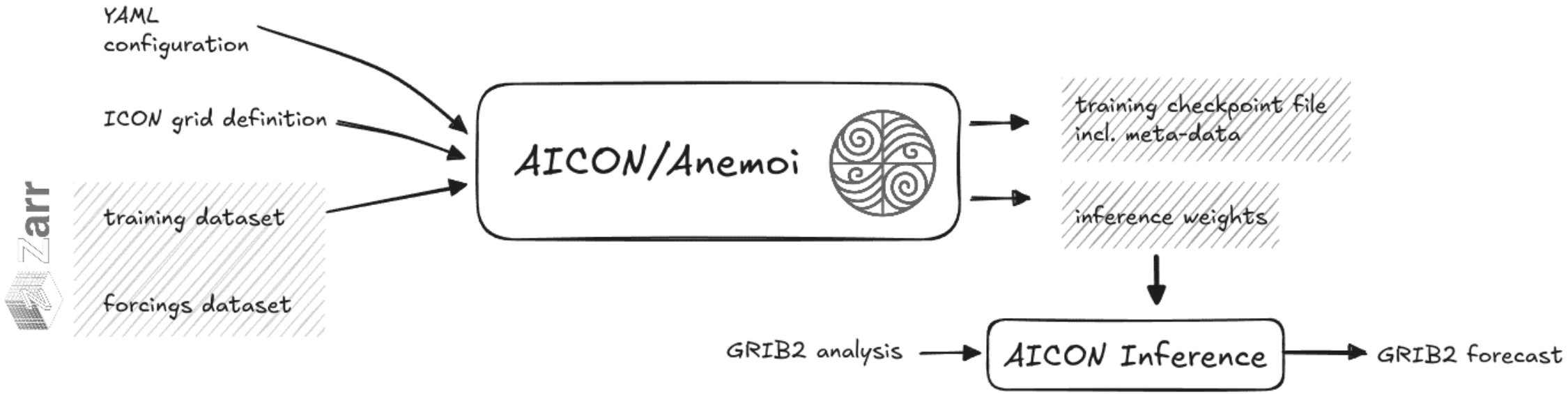
Ex.: Attention graph, relative importance of a neighbor to a target node.

Training Environment



Transfer learning between mesh resolutions: organization of a cost-efficient training schedule as long as higher resolution training graphs preserve the essential structural and statistical properties: 53 km → 26 km → 13 km

Training Environment



Training on HoreKa A100-40G

HoreKa supercomputer at KIT (Photo: S. Raffeiner, KIT/SCC)

Experiment tracking MLFlow

`mlflow.dwd.de`

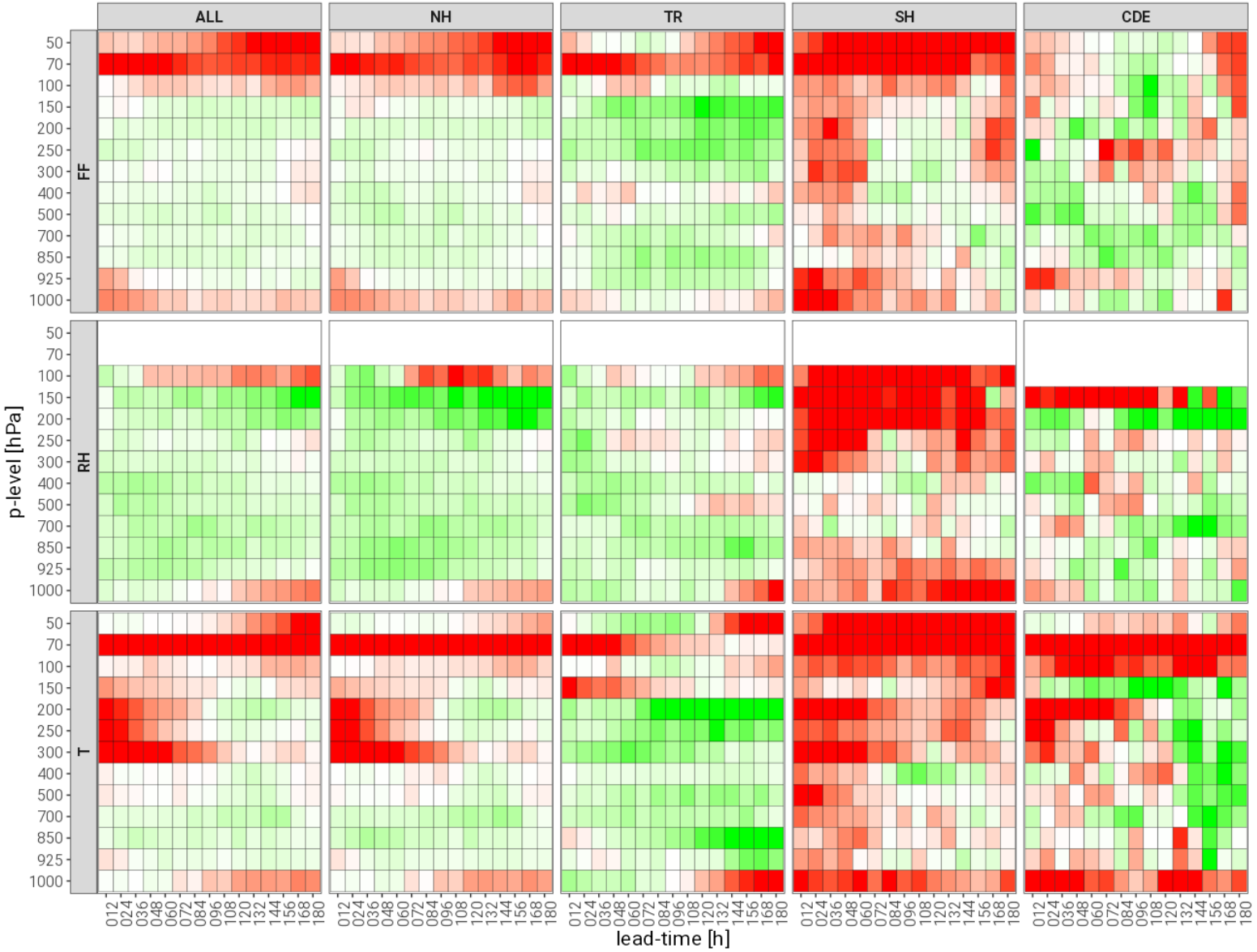
AICON
scores

AICON vs ICON upper air RMSE

Wind

Relative Humidity

Temperature



RMSE

green: AICON better

Aug 2025

AICON vs. AIFS

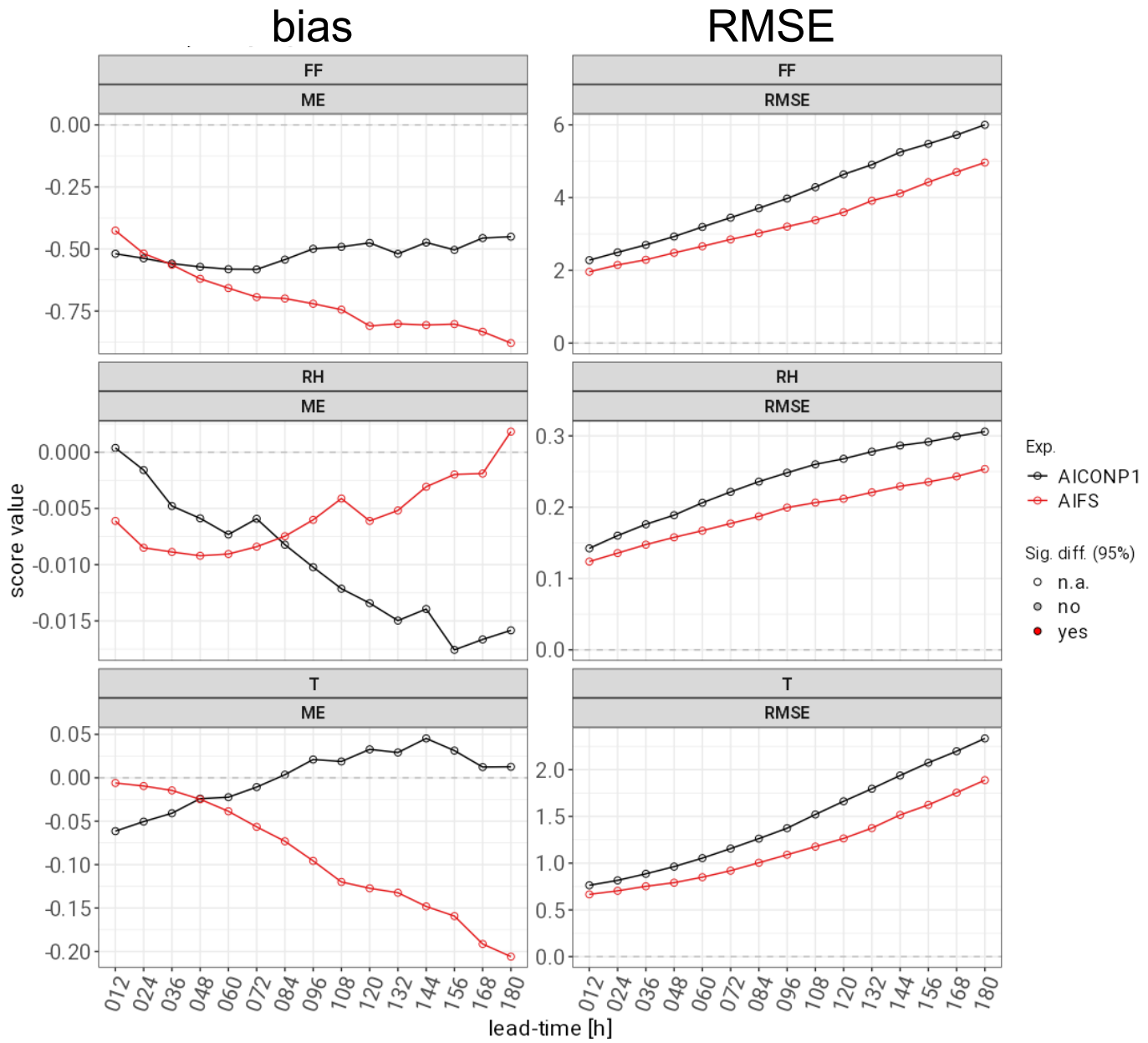
500 hPa

AICON
AIFS

Wind

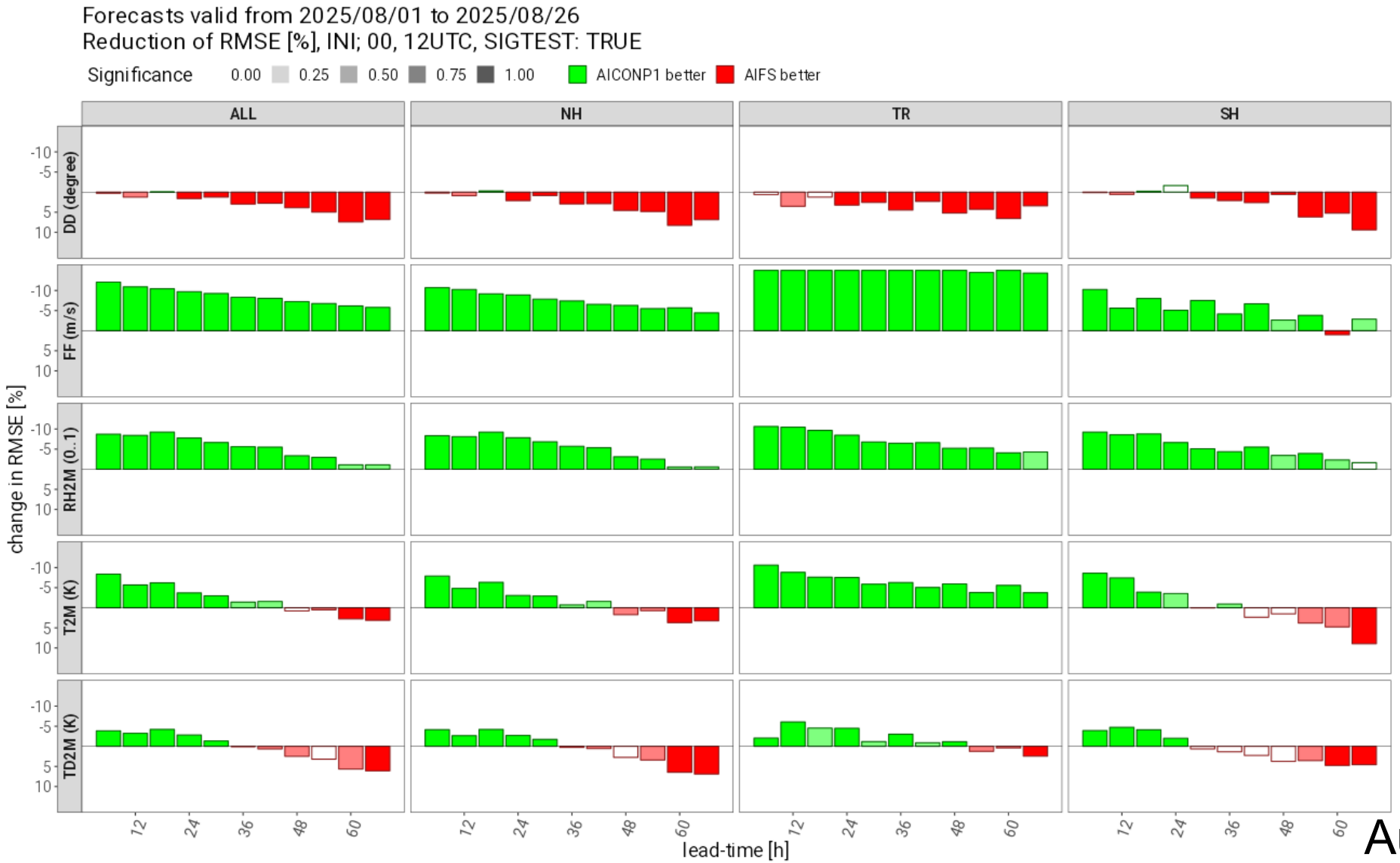
Humidity

Temperature



Aug 2025

AICON vs AIFS surface



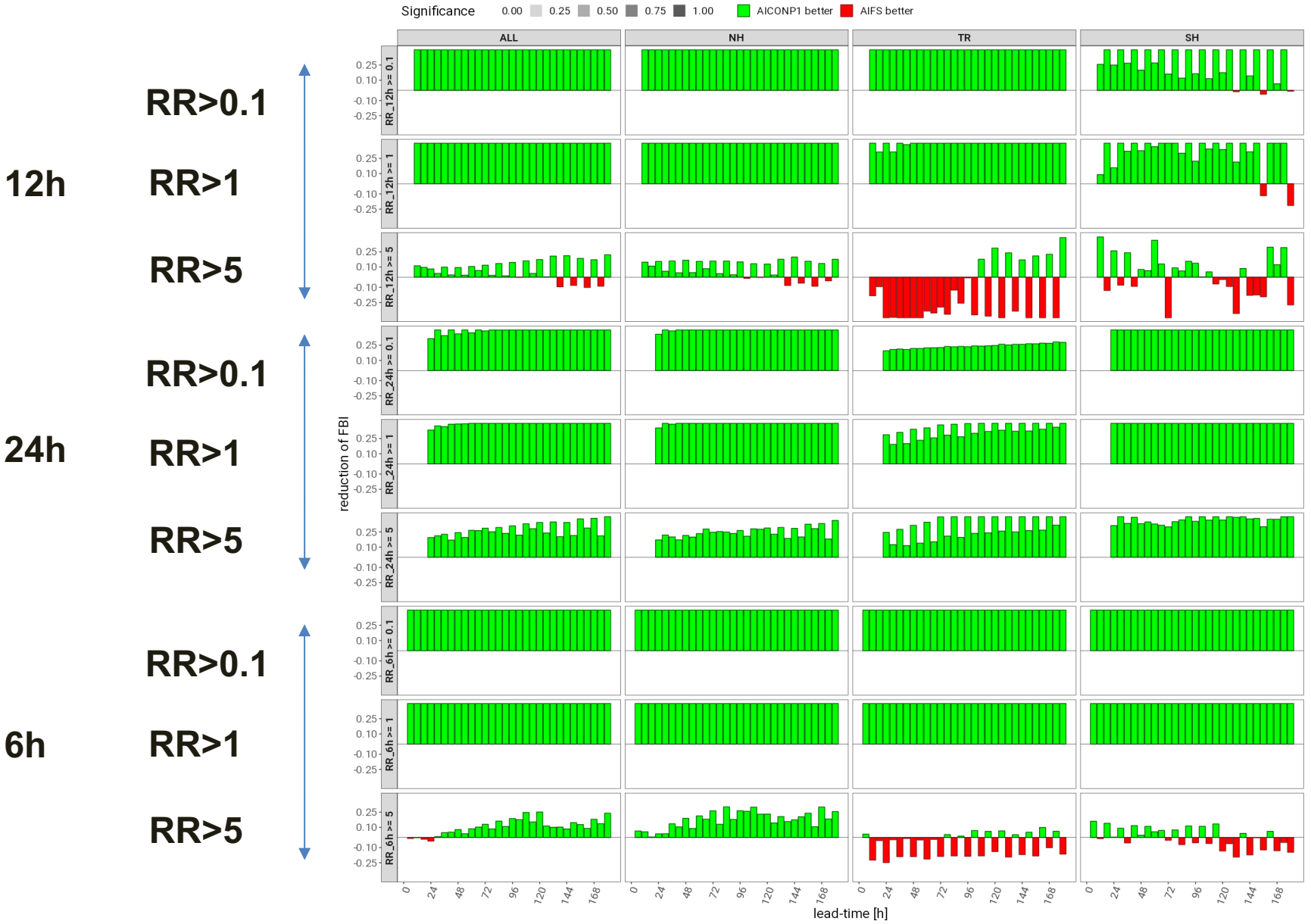
Wind

Humidity

Temperature

Aug 2025

AICON vs AIFS precipication



RR>0.1
12h RR>1
RR>5

RR>0.1
24h RR>1
RR>5

RR>0.1
6h RR>1
RR>5

Aug 2025

other DWD ai/ml applications

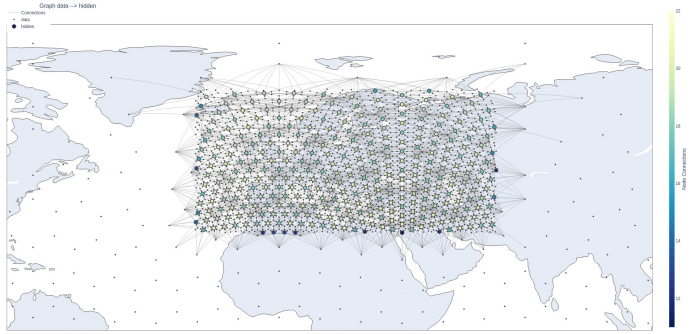
AICON Limited Area Model

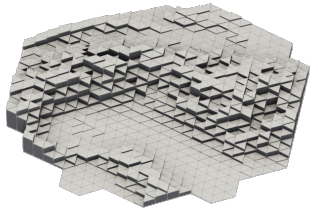
Work in progress (Sabrina Wahl et al., DWD)

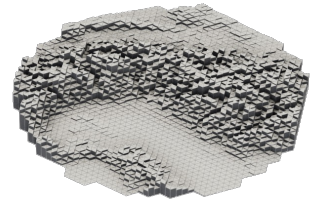
Merge global and regional input datasets.

- Regional reanalysis ~6.5 km EU nest.
- Overlay the regional dataset t+0h with the global dataset t+3h.
- Global input at boundaries and inside the domain, encoder edges from the sets of global and regional data nodes.
- Processor: multi-mesh over the LAM region.

… different from “stretched-grid model approach” where global grid points that overlap with the regional model are removed (Nipen et al. 2024: https://arxiv.org/abs/2409.01891).







R03B07, 13 km grid ICON, alpine region, R03B08, 6.5 km ASTER orography

AI-based data assimilation (AIDA)

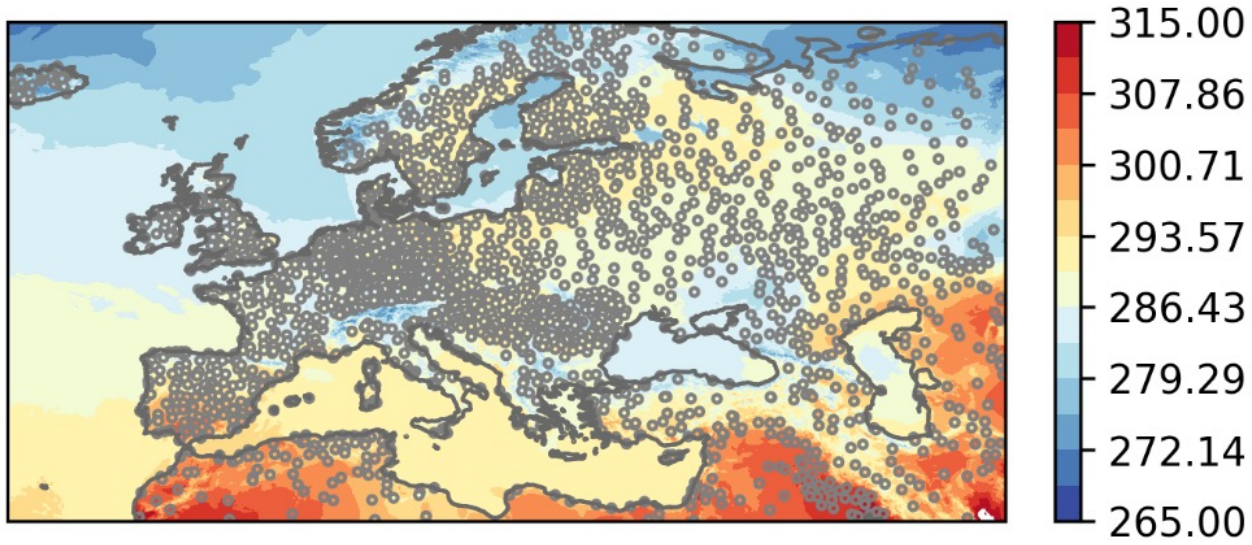
"From observation to prediction in seconds"

Fully AI-integrated NWV process chain requires an AI data assimilation approach

- AI-Var as a data assimilation meth

  Keller and Potthast (2024)

2m Temperatur in Europa am
10.05.2023, 12:00 Uhr UTC



ICON-EU Vorhersage
für 570674 Gitterpunkte
und Messwerte von
2802 Wetterstationen

$$l = (\hat{\mathbf{x}} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\hat{\mathbf{x}} - \mathbf{x}^b) + (\hat{\mathbf{y}} - \mathbf{y})^T \mathbf{R}^{-1} (\hat{\mathbf{y}} - \mathbf{y})$$

Keller, Jan & Potthast, Roland (2024). AI-based data assimilation: Learning the functional of analysis estimation.

Preparing for Operational Rollout

Plans

- Q3 2025 : AICON (13 km global) technically operational for validation by the DWD forecast center ✓
- Q2 2026 AICON-LAM (6.5 km EU) in NWP operation (technically operational)
- Q1 2027 AICON-LAM (2 km DE) in NWP operation (technically operational)