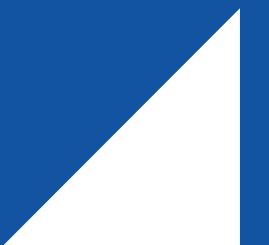


AI-LAM-IT

A km-scale, fully AI-based limited-area model for Italy
in the Anemoi framework

Gabriele Franch, Uladzislau Azhel, Elena Tomasi (FBK) · Luca Monaco (CIMA Research Foundation) · Valerio Capecchi (Consorzio LaMMA) · Edoardo Raparelli (Agenzia ItaliaMeteo) · Eleonora Bergamaschi (CINECA)

28th COSMO General Meeting · WG S/I parallel session · Rome, 14 September 2026



The project in one slide

2



What

The first fully AI-based km-scale weather model for Italy, built entirely in ECMWF's Anemoi framework: a stretched-grid model with a 2 km Italian cutout inside a global ERA5 grid, trained on Italian reanalyses.



Who

A small, operational working group started in January 2026: FBK (AI pipeline and training), CIMA (CHAPTER reanalysis), ItaliaMeteo and LaMMA (ICON-2I analyses, ITINERE reanalysis), CINECA (Leonardo, AI Factory). MeteoSwiss shares know-how and a pre-trained checkpoint.



Why now

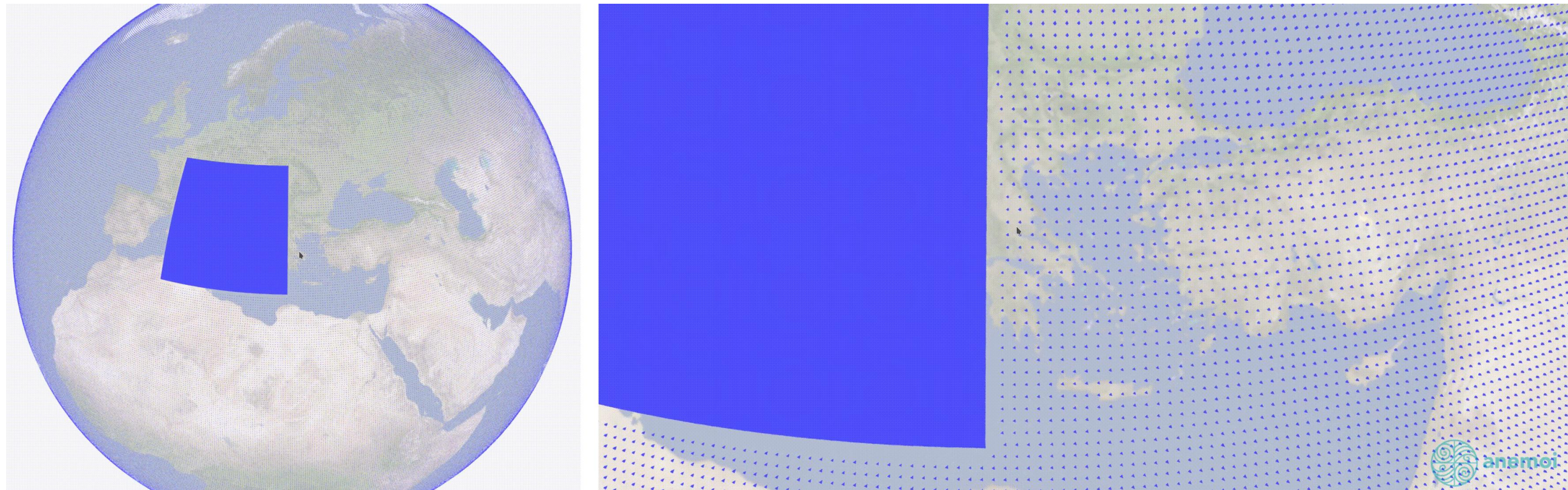
Km-scale Italian reanalyses are becoming available, Anemoi makes regional AI models a configuration problem, and the IT4LIA AI Factory provides the GPUs. AI-LAM-IT is one of the Italian AI Factory use cases.



Stretched grid + transfer learning

3

- One graph, two grids: the Italian km-scale points sit inside a global ERA5 N320 grid (~ 31 km), the hidden mesh is a refined icosahedron denser over Italy. The model forecasts globally, we keep the Italian cutout. No separate boundary model.
- Start from a checkpoint, not from scratch: MeteoSwiss trained a stretched-grid model on the COSMO-2 reanalysis over Switzerland and shared it with us. We first run it as is on our grid, then fine-tune it on Italian data.



Data nodes in the anemoi-graphs inspector. Left: the ICON-2I cutout on the globe. Right: 2 km ICON-2I points against ~ 31 km ERA5 points at the cutout edge.

Data: two Italian km-scale reanalyses

4

CHAPTER (CIMA)

- Dynamical downscaling of ERA5 with WRF at 3 km over the whole of Europe, hourly
- 10 years planned for AI-LAM-IT; produced at LRZ, being transferred to Leonardo and converted to GRIB for Anemoui
- About one year processed so far (Oct 2024 to Apr 2025); the LRZ to CINECA transfer is the bottleneck

ITINERE (Agenzia ItaliaMeteo/LAMMA/UniBo)

- ICON-2I reanalysis for Italy: 2.2 km, 65 levels, 2010-2025 with near-real-time extension
- KENDA / LETKF assimilation of conventional and radar observations, 20-member ensemble
- Produced with Arpae, University of Bologna and LaMMA; in production, first test periods being evaluated

Meanwhile: one year of operational ICON-2I analyses

2.2 km, 761 x 761 grid, 6-hourly, 12 pressure levels + surface, converted to zarr with anemoui-datasets. Used for all the tests shown next, with ERA5 N320 as global boundary.



Variables: same list on every grid

5

67

prognostic (in + out)

- 12 pressure levels (50-1000 hPa, no 600): q, t, u, v, z = 60 fields
- Surface: 2t, 2d, 10u, 10v, sp, msl, skt

1

diagnostic (out only)

- tp, total precipitation, 6 h accumulation
- ReLU bounding in output (never negative); std-only normalisation; ICON values rescaled mm to m

11

forcings (in only)

- Constant: lsm, surface z (from ICON orography)
- Computed by the recipe: cos/sin of latitude, longitude, julian day, local time; insolation

78 inputs x 2 time steps, 68 outputs, 6 h step. Dictated by the MeteoSwiss checkpoint: same variables on the global and regional grid.

Loss weights: z 12, sp 10, t 6, u 0.8, q 0.6, v and 2d 0.5, 10u/10v 0.1, tp 0.025, others 1; upper levels down-weighted. Left out for now: w, cp, tcc, tcw (available in ICON-2I, candidates for training from scratch).



Compute: CINECA and the AI Factory

6

100 k

GPU hours on Leonardo
Booster (800 k std-h)

110 k

std-h on Leonardo DCGP for
data preparation

50 TB

WORK storage for raw
reanalyses

5 TB

FAST storage for training zarr

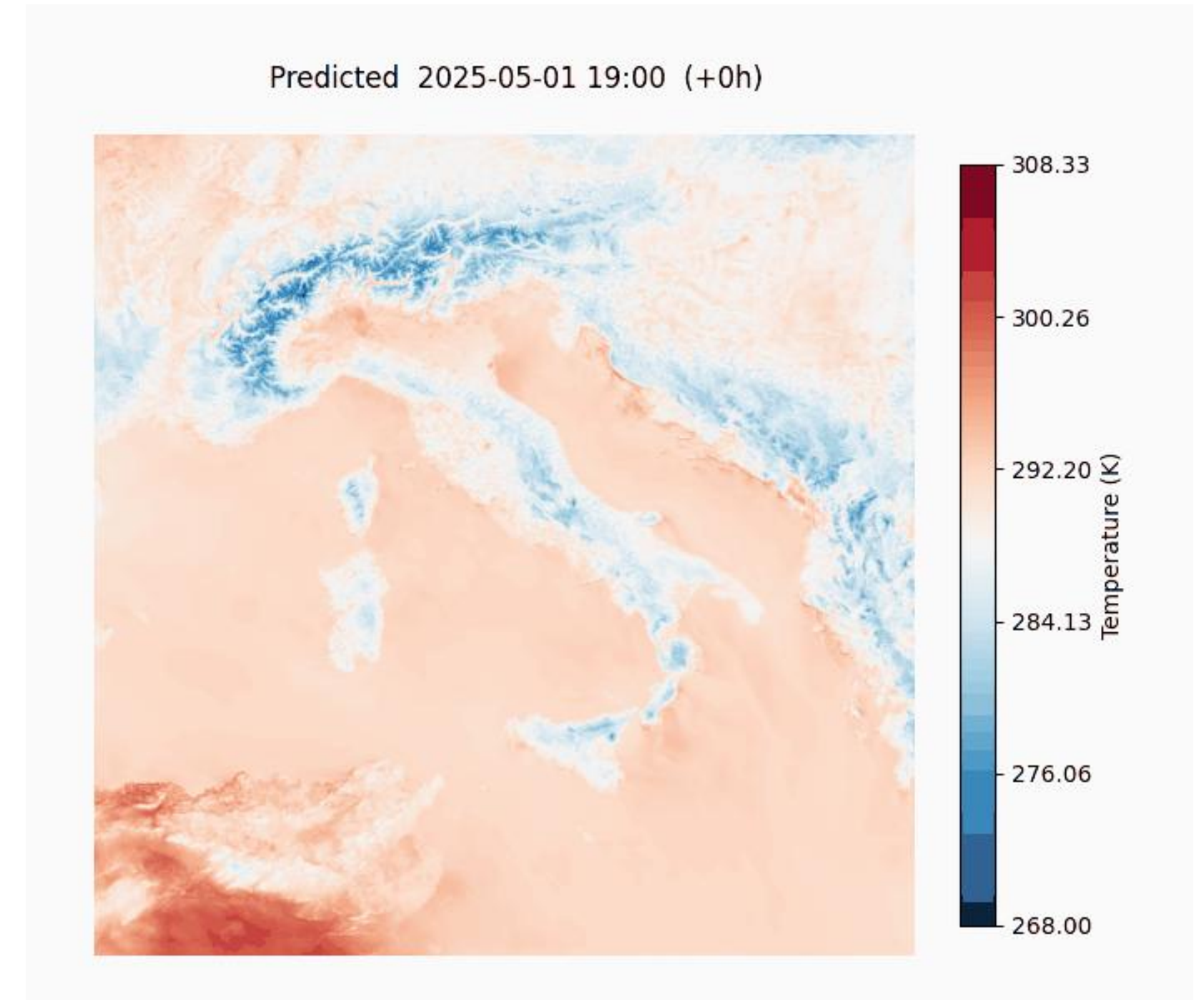
- Project AILAMIT on Leonardo, opened in February 2026 for one year under the AI Factory resources for partner projects.
- Training runs on 8 nodes x 4 A100 (32 GPUs), model sharded over 8 GPUs; dataset and graph creation on CPU nodes.
- CINECA also hosts the data hub: reanalyses land on Leonardo WORK, training zarr on FAST; extra storage being requested for the CHAPTER transfer.



So far: data, graph, and first tests

7

- Datasets: ICON-2I analyses and ERA5 N320 as anemoi zarr; variable list aligned with the MeteoSwiss checkpoint; precipitation re-accumulated to 6 h.
- Graph: stretched icosahedral mesh (level 10 over Italy, 7 elsewhere), KNN encoder/decoder edges, ~1.06 M data nodes.
- Zero-shot inference: the MeteoSwiss checkpoint run unchanged on our grid with anemoi-inference. 60 h forecasts look physically plausible (right), enough to justify fine-tuning.
- Fine-tuning on the ICON-2I year: warm start, 4-step rollout, 5 % of the loss weight on the Italian cutout, 32 GPUs. Training runs; evaluation against ICON-2I analyses in progress.

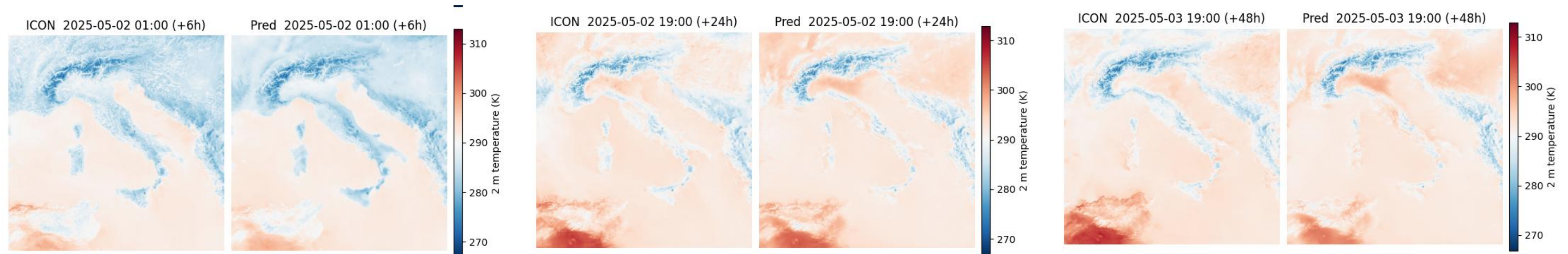


Zero-shot 2 m temperature, 6 h steps to +60 h (animated in slideshow mode).

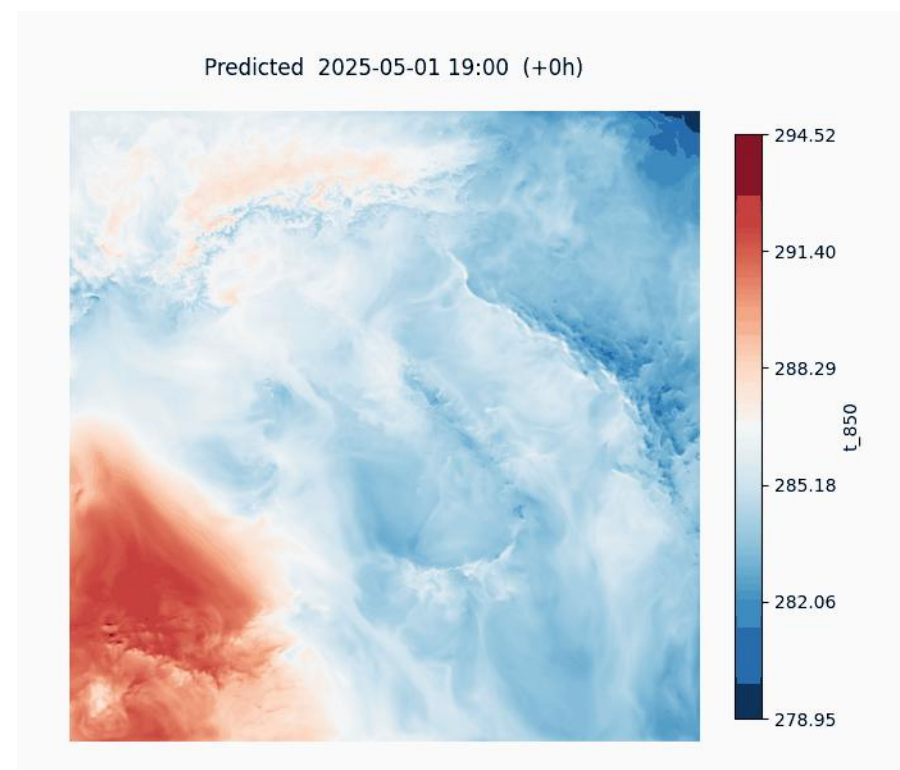


Zero-shot forecast vs ICON-2I analysis

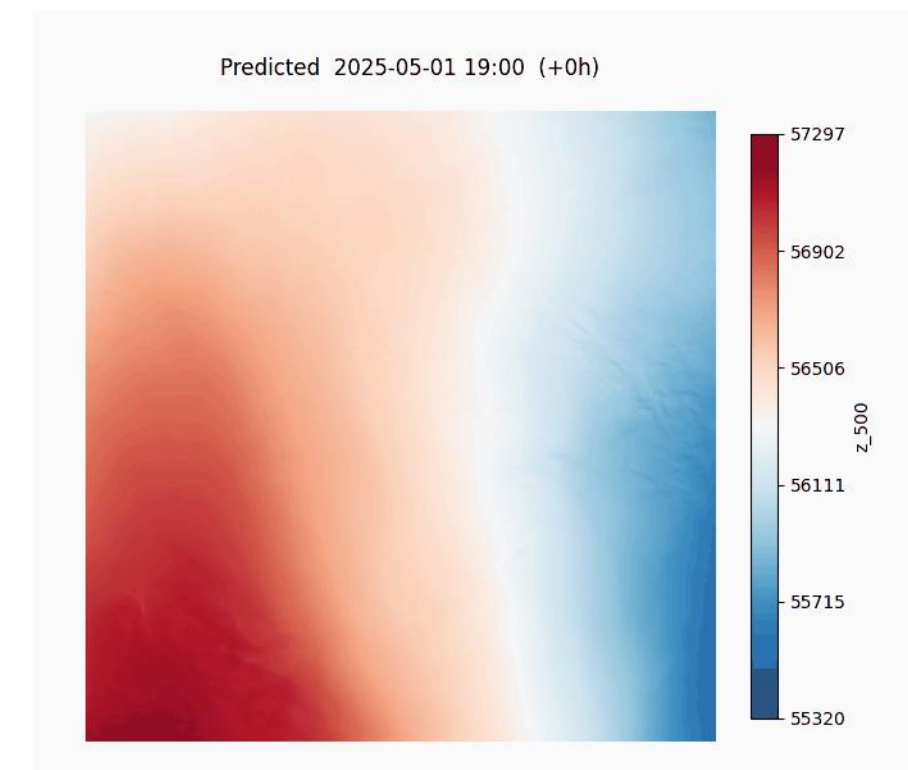
8



2 m temperature: ICON-2I analysis (left of each pair) vs zero-shot forecast (right) at +6 h, +24 h and +48 h, case 2025-05-01 18 UTC.



Temperature at 850 hPa (animated)



Geopotential at 500 hPa (animated)



Lessons learned (the expensive ones)

9

1

Pin the Anemoi version

A checkpoint only loads with the library versions it was trained with. We rebuilt that environment from the checkpoint metadata and repacked the state dict for Lightning.

2

Use N320 for ERA5

Node area weights on the lat-lon grid were wrong; N320 (the IFS / AIFS grid) fixed them and matches the operational data stream.

3

Area-weights bug

anemoi-graphs SphericalAreaWeights drops empty Voronoi regions and mis-assigns every area after the gap. We ship a patched class with the configs.

4

Copy the original config

Imputer order, normalisers, bounding, embeddings: every deviation from the checkpoint's config cost accuracy. Start from an exact copy.

5

Check what loaded

A setup mismatch silently re-initialised some layers, which then over-fitted the small dataset. Read the loaded-keys report every time.

6

6 h precipitation

AIFS-style checkpoints predict 6 h accumulations; ICON and ERA5 carry 1 h. Accumulate in the recipe and patch the GRIB metadata Anemoi relies on.



Next steps

10

1 Bring in the reanalyses

Finish the CHAPTER transfer and conversion; ingest the first ITINERE periods as they are released. Same variable list, same recipes.

2 Train on Italy

Fine-tune the stretched-grid model on CHAPTER and ITINERE instead of one year of analyses; longer rollouts; hourly steps where the data allow.

3 Evaluate properly

Scores against ICON-2I analyses and station observations, case studies on recent floods, comparison with the operational ICON-2I forecast.

4 Share

Recipes, graph configs and checkpoints through the IT4LIA AI Factory; we are keen to compare notes with other COSMO / ICON AI efforts.



Thank you

Gabriele Franch, franch@fbk.eu
Fondazione Bruno Kessler, Trento

Most of the hands-on work: Uladzislau Azhel (FBK).
With thanks to MeteoSwiss (Daniele Nerini, Francesco Zanetta, Alberto Pennino) for the checkpoint and the advice, to Renata Pelosini (ItaliaMeteo) for starting the project, and to CINECA for the Leonardo resources under the IT4LIA AI Factory.