Towards ensemble data assimilation of satellite radiances in Machine Learning Limited Area Models

Marcello Grenzi [1], Thomas Gastaldo [2], Virginia Poli [2], Chiara Marsigli [2], Tijana Janjic [3] and Alberto Carrassi [1]

[1] Department of Physics and Astronomy, University of Bologna, Italy

[2] Arpae Emilia-Romagna, Bologna, Italy [3] Catholic University of Eichstätt-Ingolstadt, Ingolstadt, Germany

5th ECMWF-ESA Machine Learning Workshop

INTRODUCTION

1) Need to exploit satellite radiances better in convective-scale assimilation

Satellite radiances assimilation is crucial for Numerical Weather Predictions (NWP) in global models. However, radiances are under-used in Limited Area Models (LAM), especially at microwaves.

2) Machine Learning Weather Predictions (MLWP) complementing or replacing NWP

Several global MLWP models exist, but with limited resolution. Recent attempts to build high-resolution Machine Learning Limited Area Models (ML LAM) up to convective-scale [1].

Based on these aspects, we first show the implementation of **ensemble-based assimilation of microwave and infrared water vapour channels** in a operational NWP LAM.

- Microwaves: low spatio-temporal resolution, good cloud penetration.
- Infrared: high spatio-temporal resolution, no cloud penetration.

This motivates the development towards radiances assimilation in a ML LAM. Methodology, preliminary results, main scientific challenges, and future outlooks are described.

1. RADIANCES ASSIMILATION IN NWP LAM

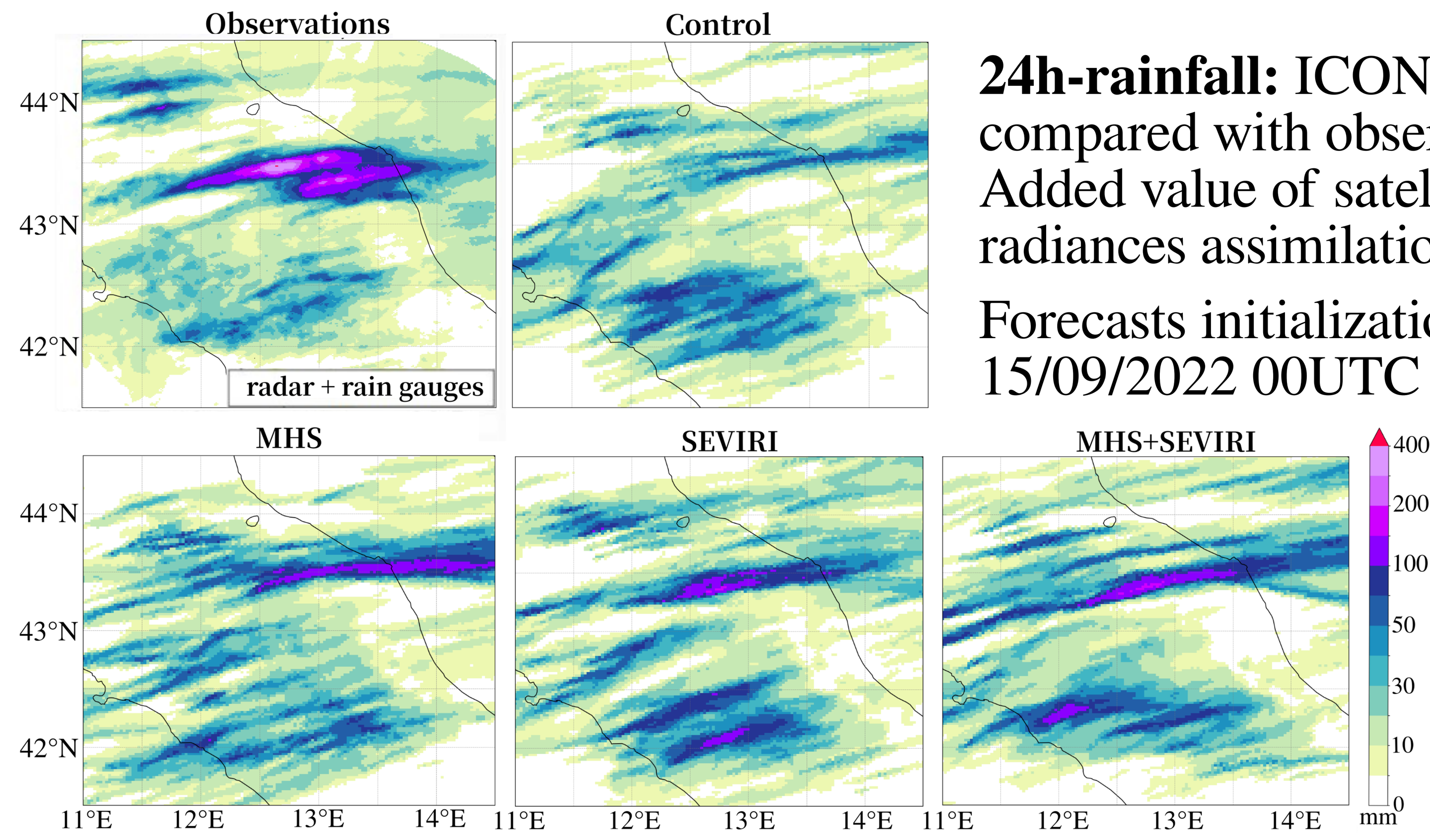
ICON: ***ICOsahedral Non-hydrostatic model*** *at 2.2km resolution*

KENDA: *Kilometre-Scale Ensemble Data Assimilation* system, based on a Local Ensemble Transform Kalman Filter (**LETKF**), 40 members, 1h-cycle

Severe convection case study in Marche, Central Italy (12-16 Sept. 2022) [2]

Assimilated observations	control	MHS	SEVIRI	MHS+SEVIRI
Conventional (SYNOP, AIREP, TEMP)	✓	✓	✓	✓
Radar volumes (reflectivity and radial winds)	✓	✓	✓	✓
Latent Heat Nudging (LHN)	✓	✓	✓	✓
Microwave Humidity Sounder (MHS) chan. 3, 4, 5 in *clear-sky* conditions	✗	✓	✗	✓
SEVIRI chan. 5, 6 in *all-sky* conditions	✗	✗	✓	✓

Operational setup of ItaliaMeteo and Arpae

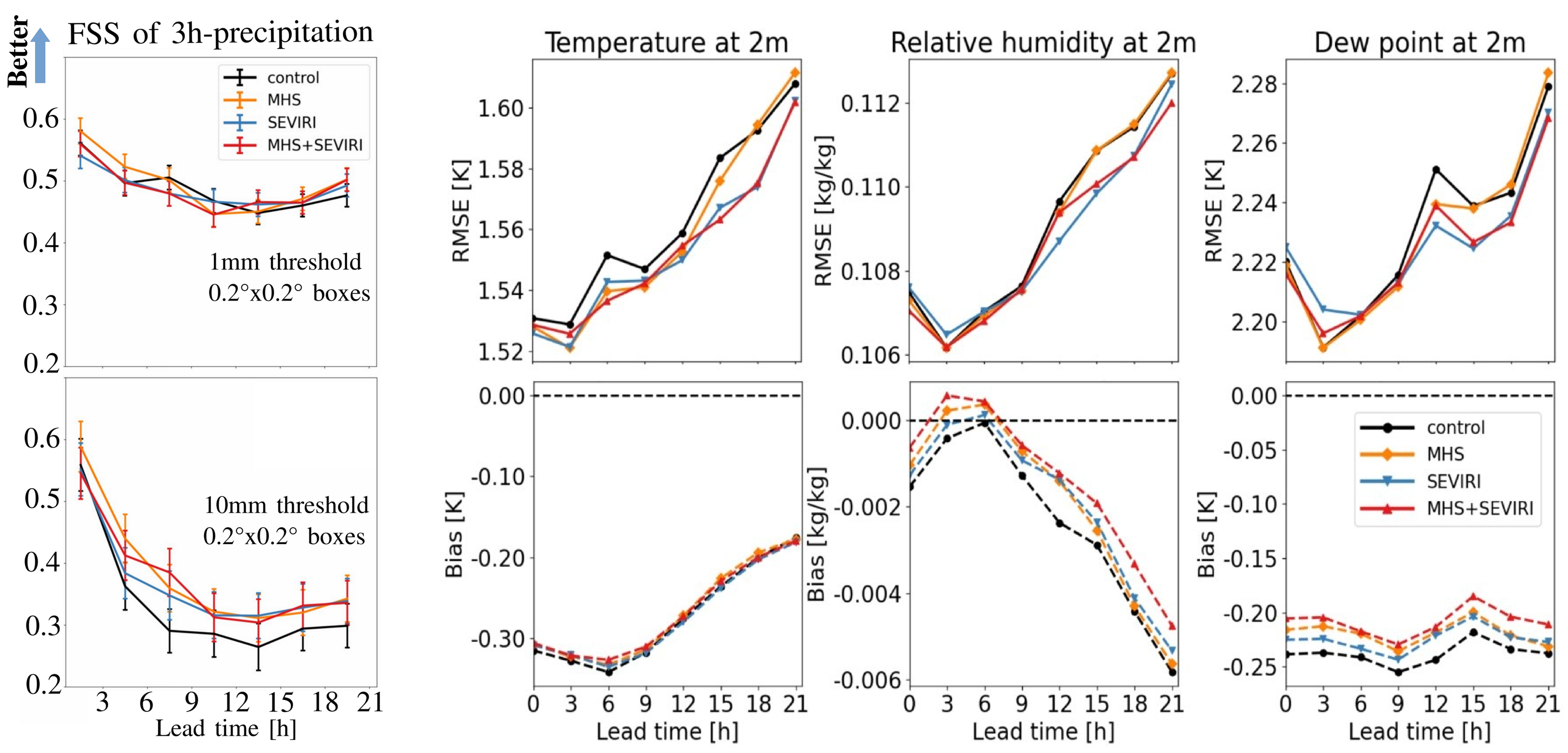


24h-rainfall: ICON forecasts compared with observations. Added value of satellite radiances assimilation.

Forecasts initialization: 15/09/2022 00UTC

Bottom left column: **precipitation** verification (Fractions Skill Score, FSS) against radar-adjusted data.

Bottom right: **surface variables** verification (RMSE and bias) against ground stations data.

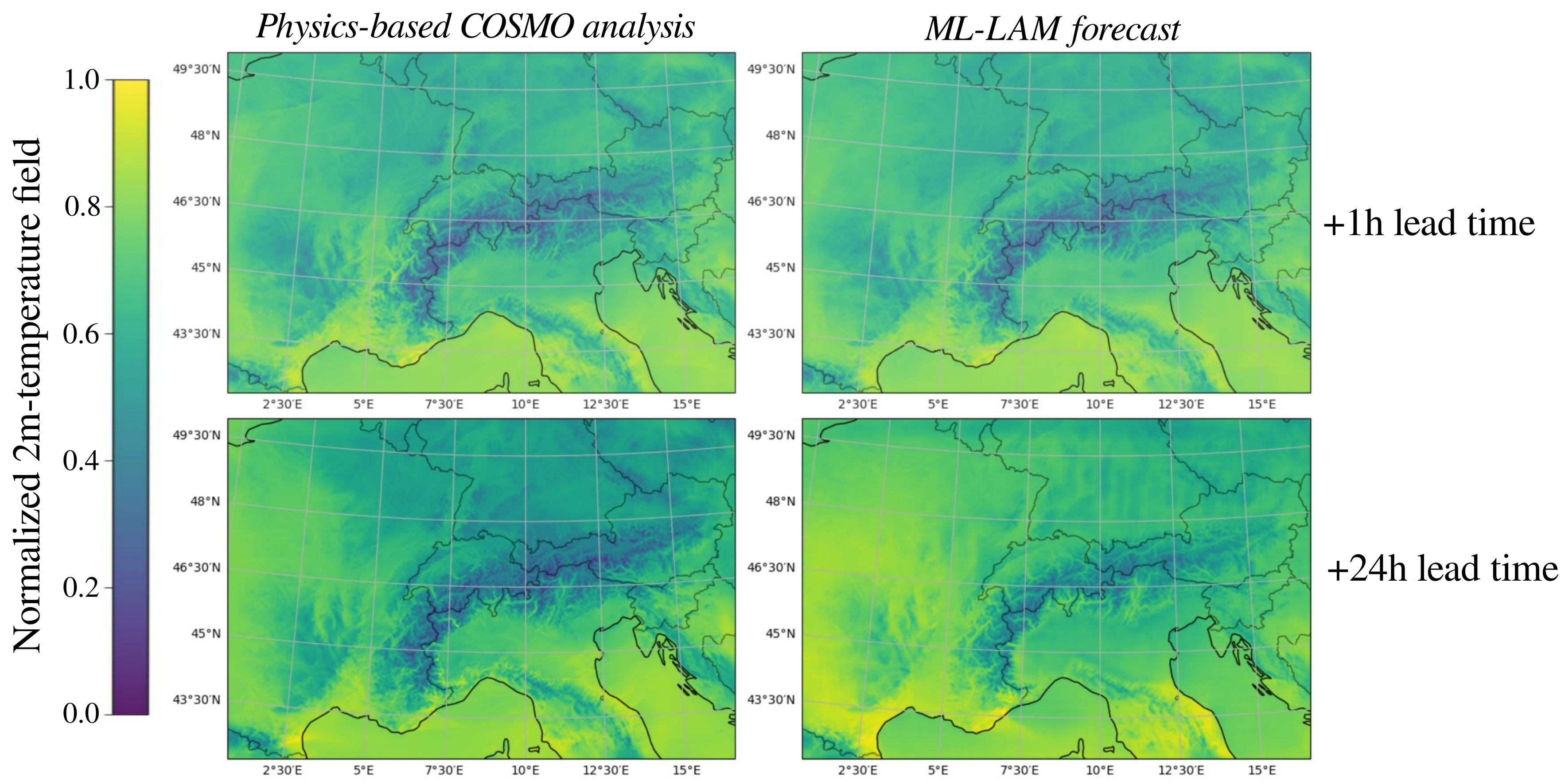


2. PRELIMINARY TESTS WITH ML LAM

ML LAM setting by Adamov et al., 2025 [1]

- Trained on operational analyses dataset over Switzerland generated with COSMO model and KENDA system
- GNN with encoder-processor-decoder architecture
- MSE loss function, 8 vertical levels, 2.2km resolution

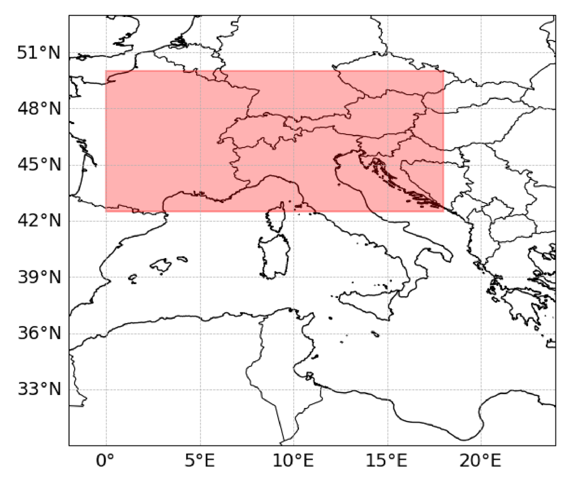
We implemented this open-source ML LAM on Leonardo supercomputer in Bologna. 24h forecast takes 4/5 minutes on A100 GPU.



Next step: running the model *over Italy* in severe convection conditions.

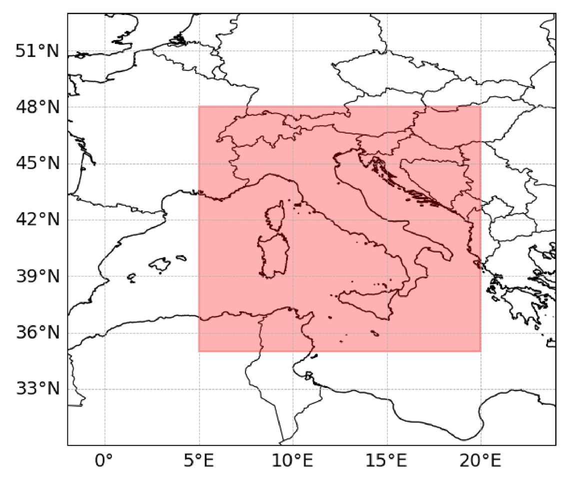
COSMO analysis data from Arpae Emilia-Romagna used as initial conditions:

- 2018-08-25 to 2025-09-30
- Archived in GRIB1 format, need conversion to **Zarr**



1) First use pre-trained Swiss model weights

2) Check model performance and if fine-tuning is needed



3. TOWARDS RADIANCES ASSIMILATION IN ML LAM

Can we successfully perform ensemble-based assimilation of radiances using this ML LAM as forecast model?

Challenges:

Few model vertical levels: detrimental for radiative transfer computations.

→ Evaluate the performance of RTTOV forward operator applied on ML LAM fields with low number of model levels.

Under-dispersion: missing upscale evolution of small perturbations, detrimental for ensemble forecasting.

→ Test boundary conditions perturbations to increase LETKF ensemble spread.

Fields smoothing: quadratic loss function training

→ Evaluate the power spectrum of ML LAM on 1h lead time forecasts (i.e. analysis cycle frequency)

CONCLUSIONS

Key results

- Assimilation of water vapour satellite channels: improving forecast of a severe convection event in an operational **NWP LAM**[3].
- Successfully running a convective-scale **ML LAM** trained on Switzerland[1].

Current work

- Evaluate the performance of **ML LAM on Italy** in **severe convection** conditions.
- Check for possible fine-tuning over Italy.

Future outlooks

- Address challenges towards ensemble assimilation of radiances in ML LAM.

References

[1] Adamov, S. et al., *Building Machine Learning Limited Area Models: Kilometer-Scale Weather Forecasting in Realistic Settings*, arXiv preprint, 2025.
[2] Grenzi, M., *Towards the assimilation of satellite radiances in the convection-permitting ICON model*, NCC, 2025.
[3] Grenzi, M. et al., *Microwave radiances assimilation for severe convection forecasting in an operational limited-area model*, QJRMS, 2026 (submitted).

Contacts
marcello.grenzi3@unibo.it