Update on Data Assimilation in COSMO

Christoph Schraff

Deutscher Wetterdienst

- Conventional DA: KENDA, based on LETKF / DACE code:
  - "DA resources" diverted to AI (DA + model emulators)
    → very little algorithmic developments (LETKF, EnVar; 4D-EnVar (global), PF)
  - observations (crowd-sourced surface, tower, satellite, DIAL) (DWD, ARPAE, CNMCA, MeteoSwiss)

- AI → AI-Var (*Keller and Potthast, 2024* *(arXiv:2406.00390v1)*)
  (& reanalyses)

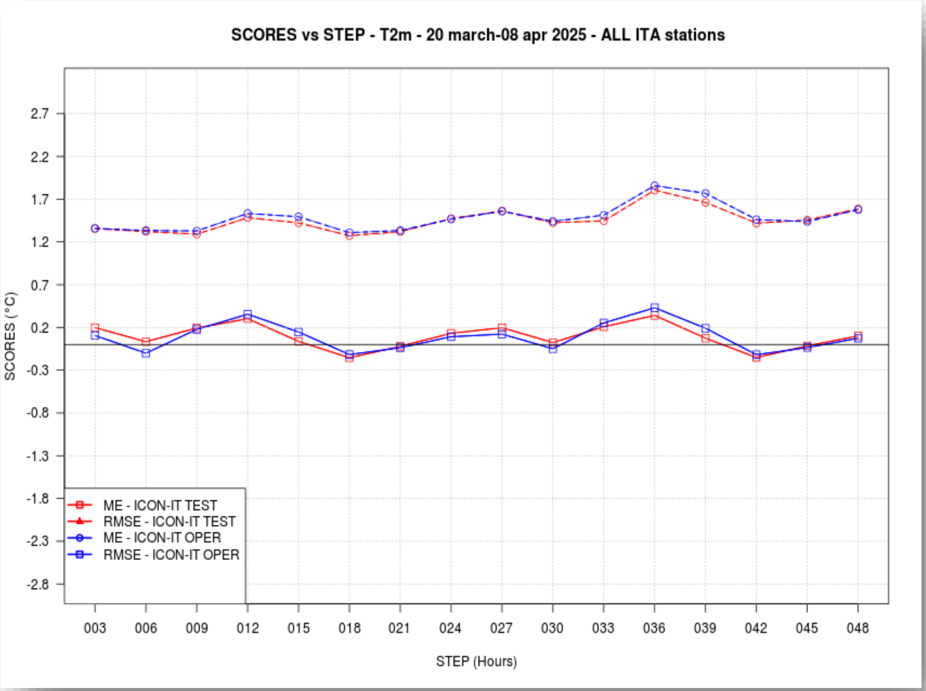
Assimilation of Meteonetwork T2m and wind speed obs

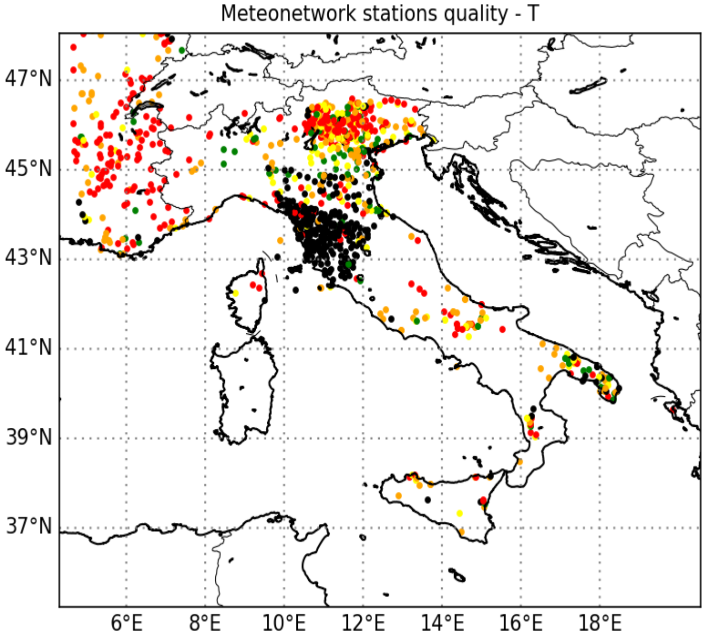
Valerio Cardinali

- Meteonetwork: crowd sourced synoptic stations (approx. 1030)
- Quality check based on bias and stdv of the obs increments

→ 176 stations used for T, 257 for wind speed (**operational in May 2025**)

surface temperature verification vs synop





Green and yellow stations assimilated

2026

- DA of Meteonetwork stations (pressure and rh2m obs)
- DA of surface obs from:
  - MeteoMont (Army Alpine stations): 46 stations
  - Autostrade: 250 stations
  - DPC (Civil Protection): approx. 5500 stations

- ICOS: Integrated Carbon Observation System

  ecosystem, ocean, atmosphere
  (greenhouse gases)

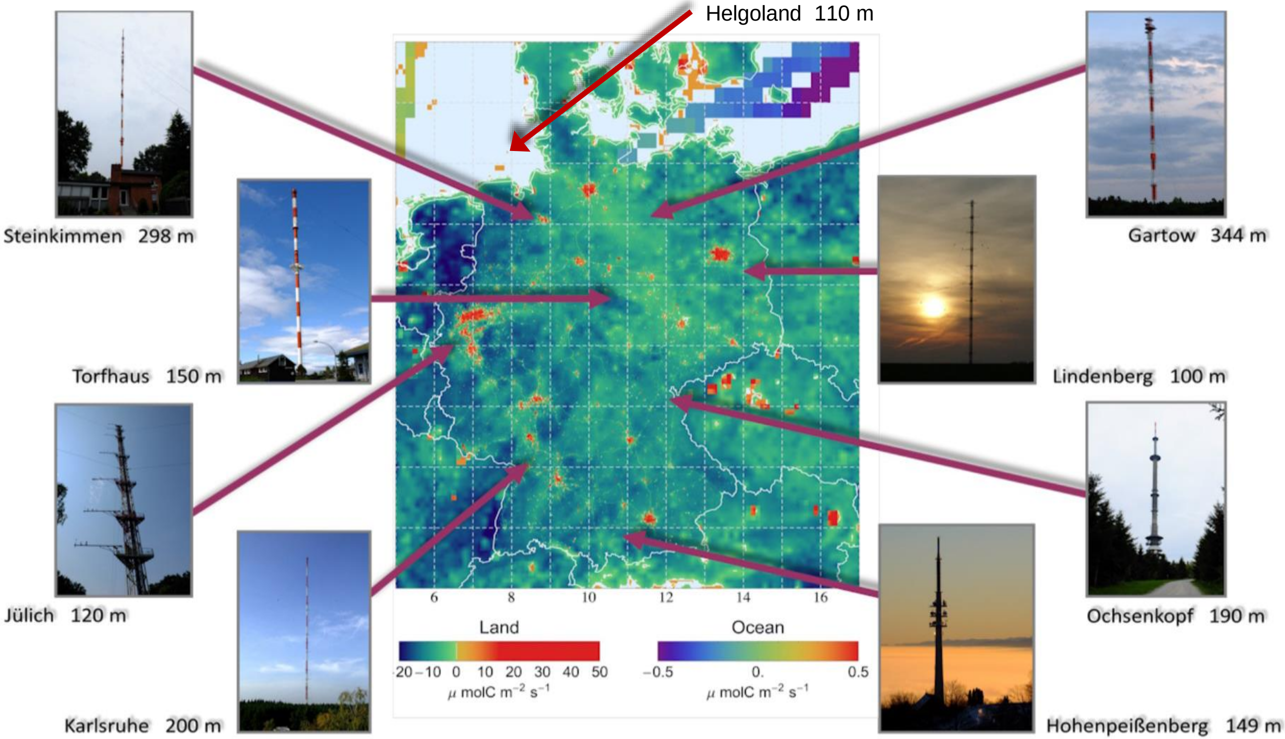
- 9 towers in Germany

  (within ICON-D2: Cabauw NL,
  Ispra IT, 1 CZ, 3 FR
  (no access to data yet))

- heterogeneous architectures,
  max. heights 98 – 341 m
  (min. heights 10 – 60 m)

- T, RH on 2 – 6 levels,
  wind on 1 – 5 levels

- plus 2 towers at former
  nuclear power plants

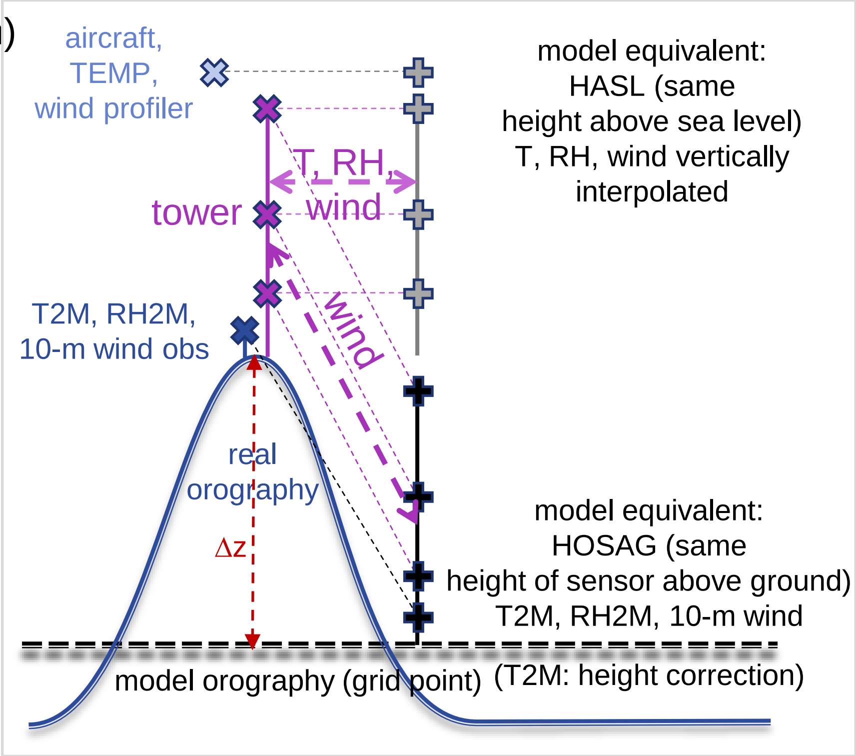


PBL profiler obs: ICOS towers

- 2 mountain station ($\Delta z \cong 100m$), 2 hill stations ($\Delta z \cong 30m$)
- tower T, RH: upper-air “HASL” operator always better
- tower wind: 2 towers, surface “HOSAG” operator is better for lower level (60m, 93m), “HASL” operator better for all other levels

→ need to look at each individual tower / level

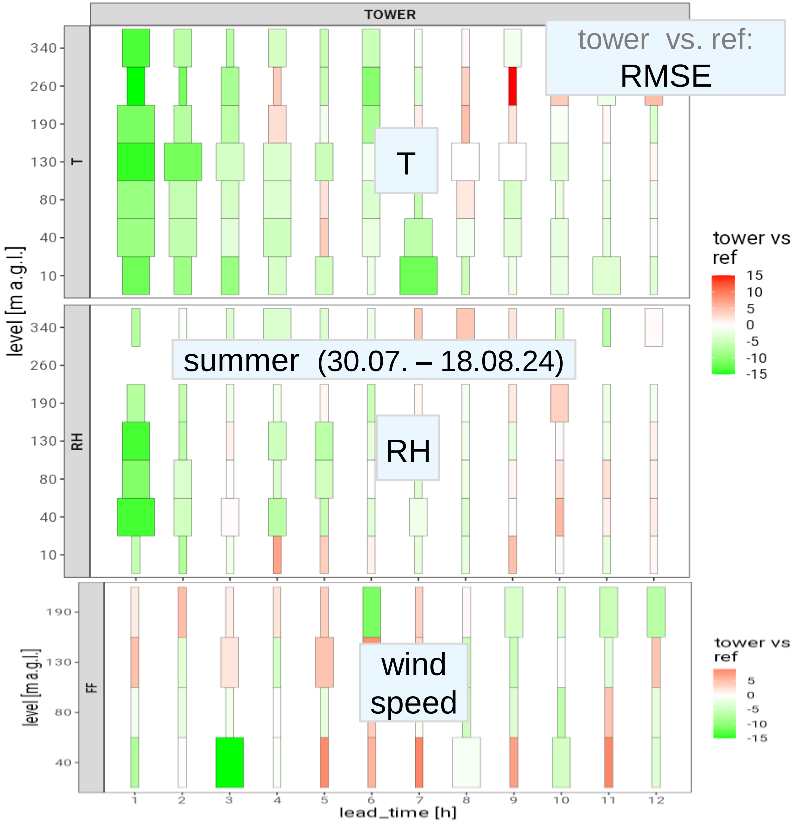
- ICOS towers: wind speed is underestimated at many towers / levels (→ wind blacklisted)
- assimilated data:
  - T : 34 obs levels from 9 towers
  - RH: 28 obs levels from 9 towers
  - wind: 8 obs levels from 4 + 2 towers



PBL profiler obs: ICOS towers

- impact experiments
  (2 winter + 2 summer periods, 14 – 20 days, weak advection
  - verification against towers:
    large positive impact vs. tower T + RH in first hour(s)
    decreasing rapidly, but pertaining up to 3 / 6 / 12 / 24 hrs

    little impact on tower wind
    (only 8 obs levels from 6 towers)
  - difficult to see impact vs. other obs
    → impact is local
- operational since Feb. 2025



All-sky (SEVIRI →) **MTG FCI** IR WV + VIS data

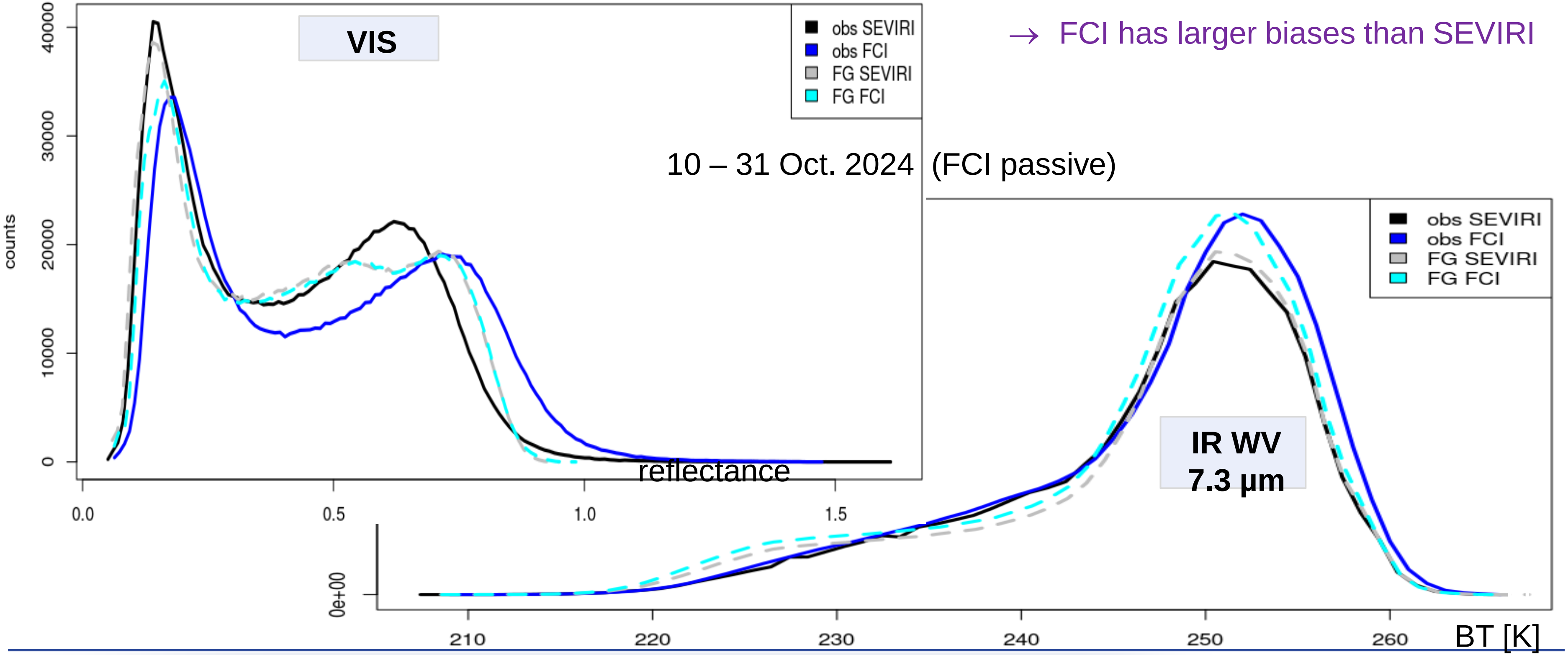
Annika Schomburg a.o.

Deutscher Wetterdienst DWD

- monitoring: comparison to SEVIRI: 2D images, time series of RMSE + bias, histograms

→ FCI has larger biases than SEVIRI

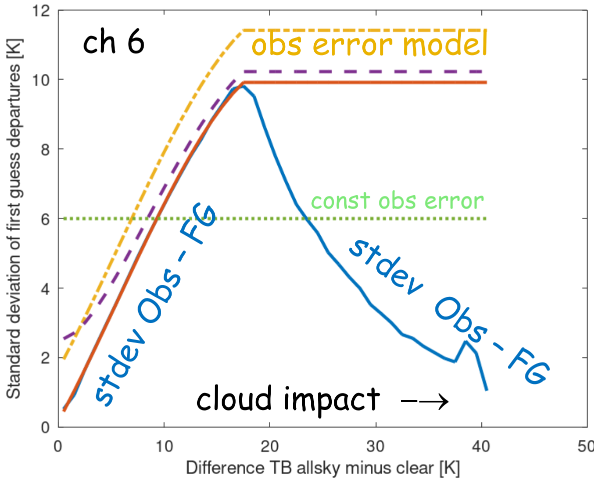
10 – 31 Oct. 2024 (FCI passive)



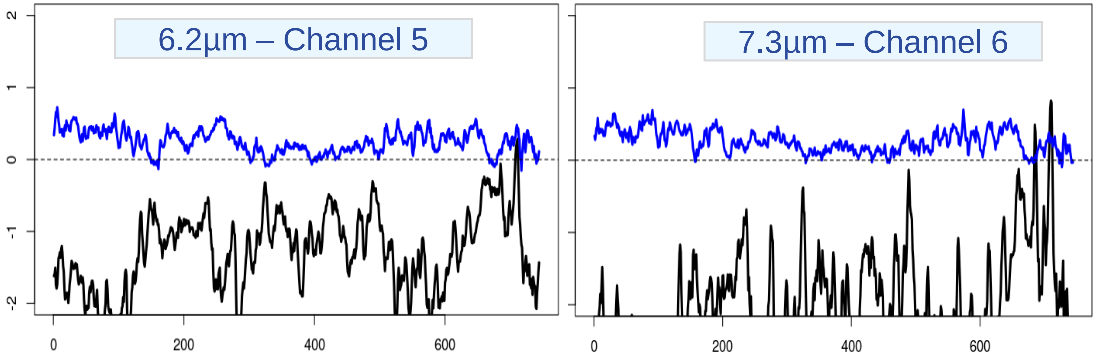
all-sky (cloudy) SEVIRI **IR WV** radiances (operational)

Deutscher Wetterdienst

- so far – no bias correction for SEVIRI …
  - obs error of 2 K assigned to clear-sky obs = 2 K >> stdev(O – FG)
  - → low weight given to clear-sky data (→ water vapour) (while cloud info is used well)



- for clear-sky, need bias correction
  idea: derive BC from clear-sky … (SEVIRI data passive, 4 weeks August 2022)

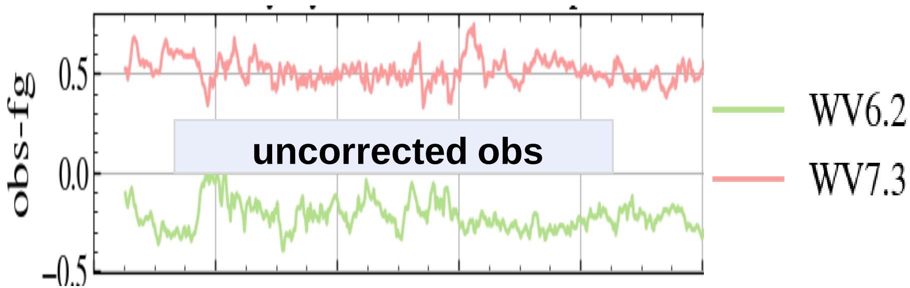


→ *average clear-sky bias: + 0.25 K* → *average clear-sky bias: + 0.27K*

… and apply all-sky

→ need bias correction based on clear-sky cases, also in order to use clear-sky WV info well

- clear-sky screening for bias correction by symmetric cloud impact check: $FG_{all-sky} – FG_{clear}$, $Obs – FG_{clear}$ (for SEVIRI also: NWC-SAF cloud mask): implementation, tuning, evaluation *(by Mahdiyeh Mousavi)*



time series 8 – 28 Oct. 2024 (passive FCI monitoring exp.)

→ bias (Obs_{orig} – FG) and (Obs_{bcor} – FG) found sensitive to symmetric cloud check in experiment, $|Obs – FG_{clear}|$ < threshold implies FG check!

→ this needs to be replaced by NWC-SAF cloud mask for clear-sky screening (for QC & BC)

All-sky (SEVIRI →) **MTG FCI** IR WV + VIS data

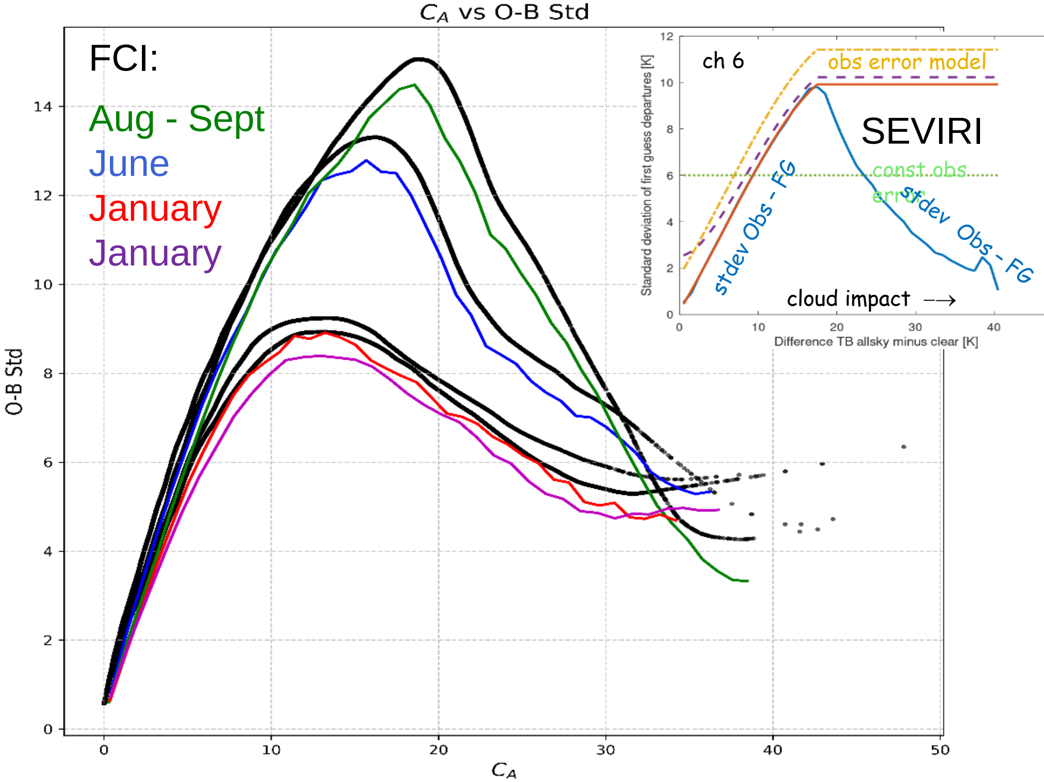
Annika Schomburg

Further steps:

- re-train observation error model (based on observation impact) for FCI

  → O – B for FCI very dependent on season(
  or probably more on cloud type
  (stratiform, convective)

- run and evaluate
  active assimilation experiment
  (i.e. FCI active, SEVIRI passive)



Satellite DA: MHS *(Marcello Grenzi)*

Testing the assimilation of **humidity-sensitive microwave channels** from the Microwave Humidity Sounder (**MHS**)

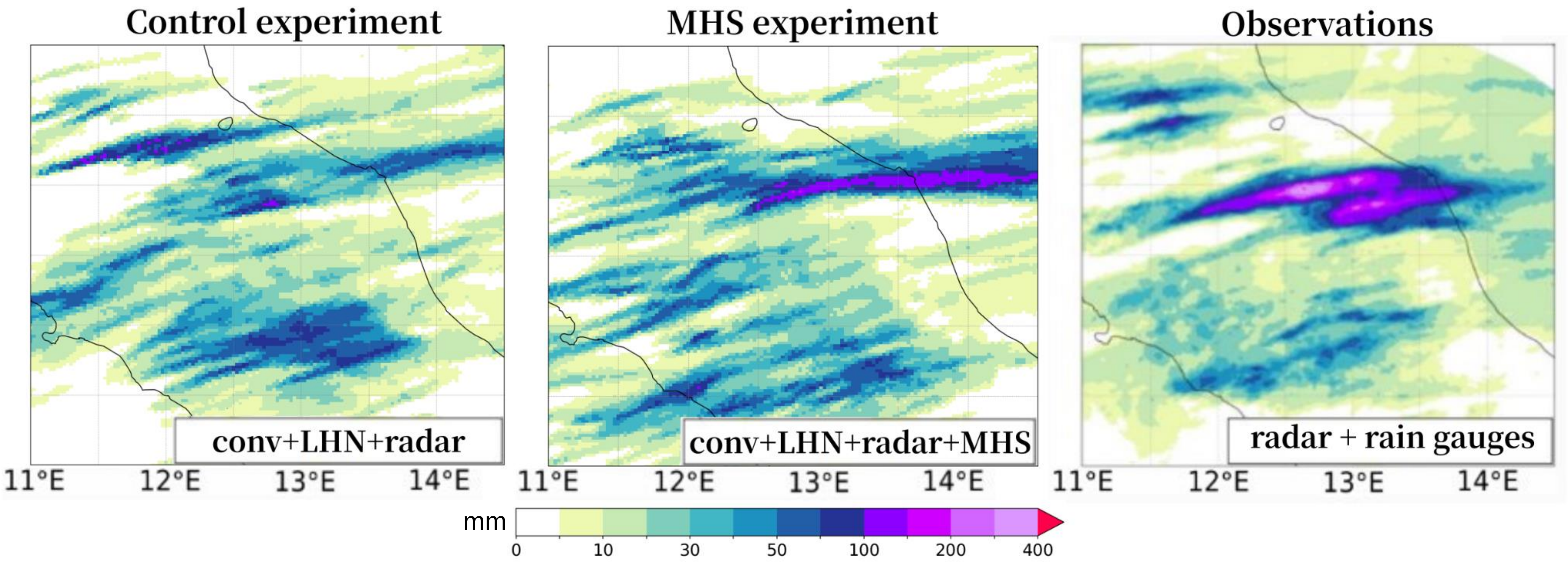
- 3 channels peaking at different levels (lowest channel only over sea - experiments ongoing over land too)
- currently **clear-sky** only (cloud detection - *Buehler et al. 2007*)
- No interchannel correlations and no bias correction applied
- Model & KENDA setup same as the operational system

5-day experiment in September 2022 (severe convection event):

- **Control** (operationally assimilated obs.)
- **MHS** (operational obs. + MHS)

Currently experiments with the addition of **SEVIRI all-sky** data are also running: evaluate the benefit of combining infrared and microwave radiances.

MHS assimilation: daily accumulated rainfall



24h rainfall of 15 Sept. 2022 (forecast initialized at 00UTC), zoom over the area affected by the flood. MHS assimilation improves the forecast.

PBL profiler obs: DIAL (Differential Absorption Lidar)

Jens Pruschke

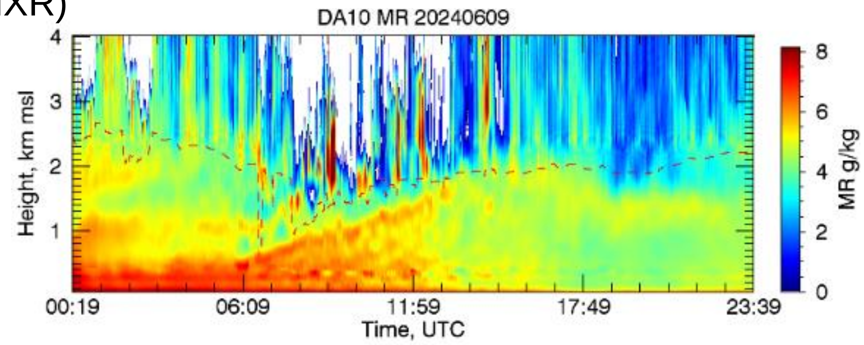
Deutscher Wetterdienst

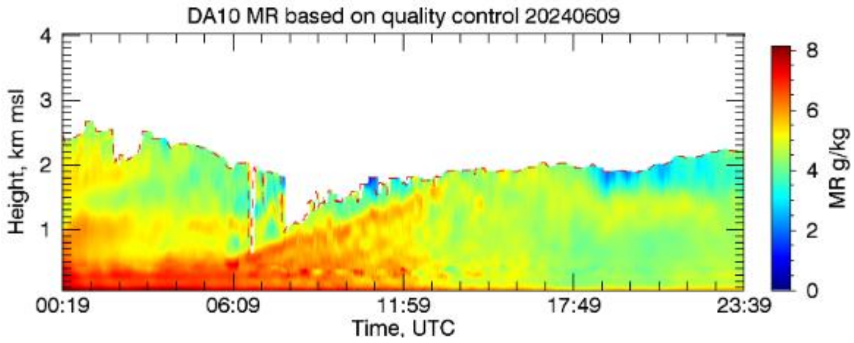
DIAL → vertical profiles of water vapour mixing ratio (MXR) in lowest 3 – 4 km (or up to low cloud base), resolution 10 m (30 m?), every 20 min

- DWD plans to buy **10 – 14 DIAL** instruments for locations across Germany … in addition to wind lidars

- 1 testbed DIAL located at Lindenberg observatory (BUFR data), maturity for DA tbd

- DACE observation operator for MXR implemented (in MEC), obs error set to 3g/kg





PBL profiler obs: DIAL (Differential Absorption Lidar)

Jens Pruschke

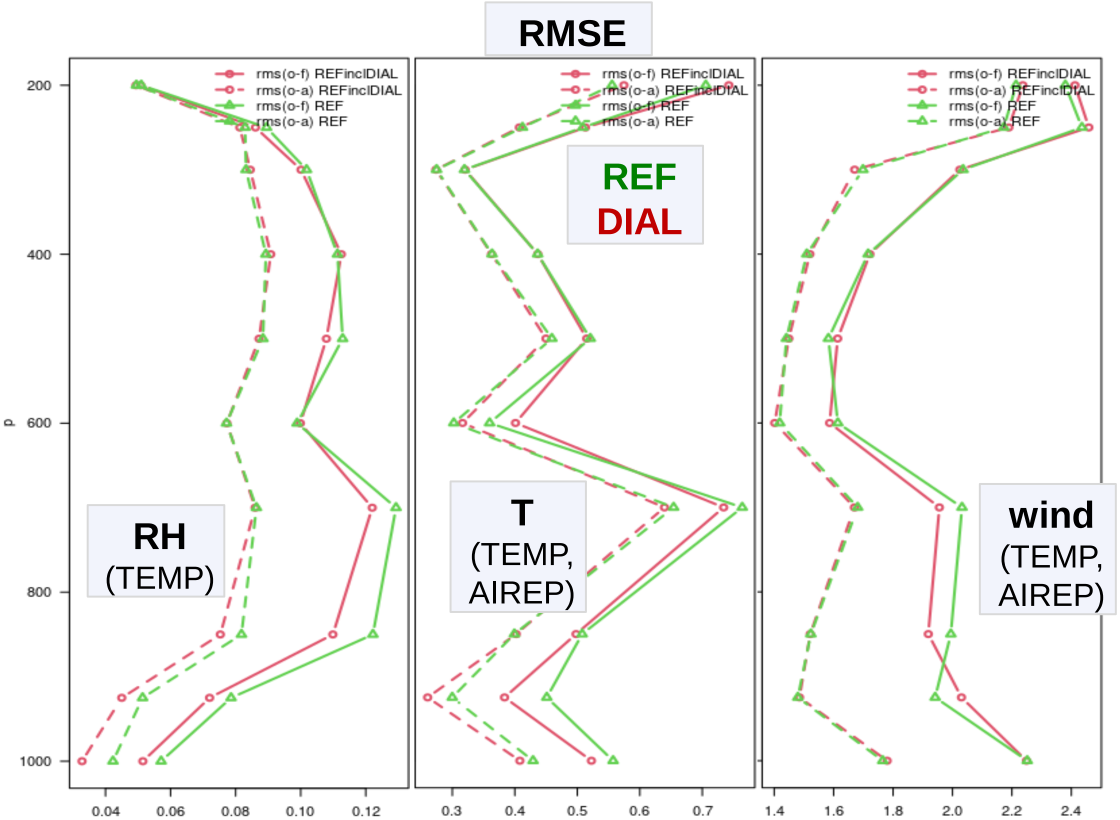
- first DA experiments over 8 days
  (4 – 11 June 2024)

- verification: obs_err_stats
  around Lindenberg (± 0.1°)

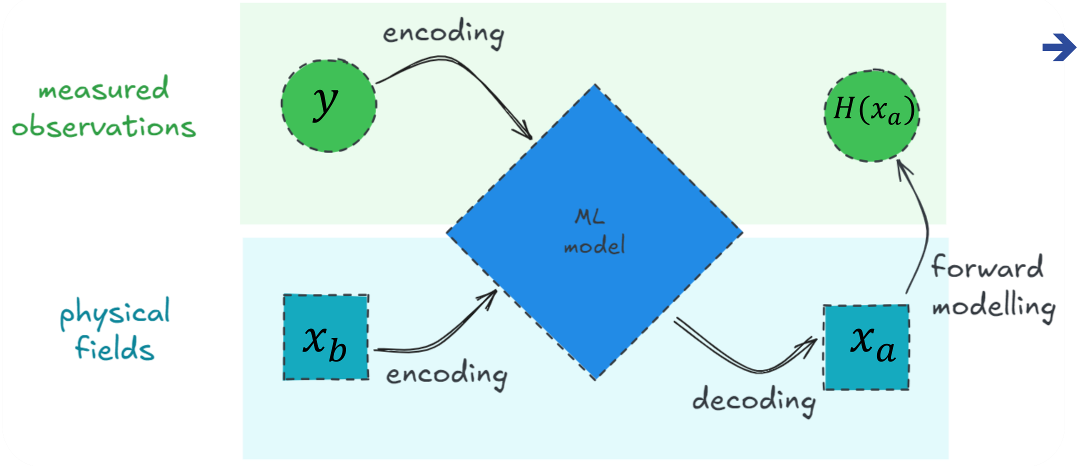
  → positive impact on humidity
  (& temperature, wind)

- next steps: further experiments with
  - height dependent observation error
  - profile thinning
  - longer periods (summer/winter)

  forecast verification



Data Assimilation with AI



There is no target data (*labels*) for training!

The model *learns* to estimate the analysis from minimising the loss function.

THE METHOD

- The classical data assimilation functional is used as a loss function to optimize the model weights:

$$J[x] = \alpha \|x - x_b\|^2_{B^{-1}} + \|H[x] - y\|^2_{R^{-1}}$$

α: Tikhonov regulator
$x_a = \text{argmin}(J)$: analysis state
x_b: background state
B: model space covariance
R: observation space covariance
H: forward operator
y: observation

- reference: J. Keller and R. Potthast, arXiv:2406.00390

Data Assimilation with AI

ACHIEVEMENTS

- ➔ global DA of 2m temperature
- ➔ reading and processing of NWP data and observations
- ➔ NN forward operators for in-situ data
- ➔ post-processing of cloud diagnostics with SEVIRI data
- ➔ basis for FRAIM (AI framework at DWD, with modularized AIDA)

IN DEVEOPMENT

- ➔ DA of upper air observations for smaller domains
- ➔ running RTTOV as forward operator inside a NN (for DA of satellite data)
- ➔ combination of different NWP forecasts (SYNCAST)

FUTURE PLANS (a.o.)

- ➔ implementation of more forward operators (radar, …)