GeoSphere Austria

Evaluation AROME 500 m Simulations in the Framework of DE On-Demande-Extremes

Phillip Scheffknecht, Florian Weidle, Christoph Wittmann

Funded by the European Union

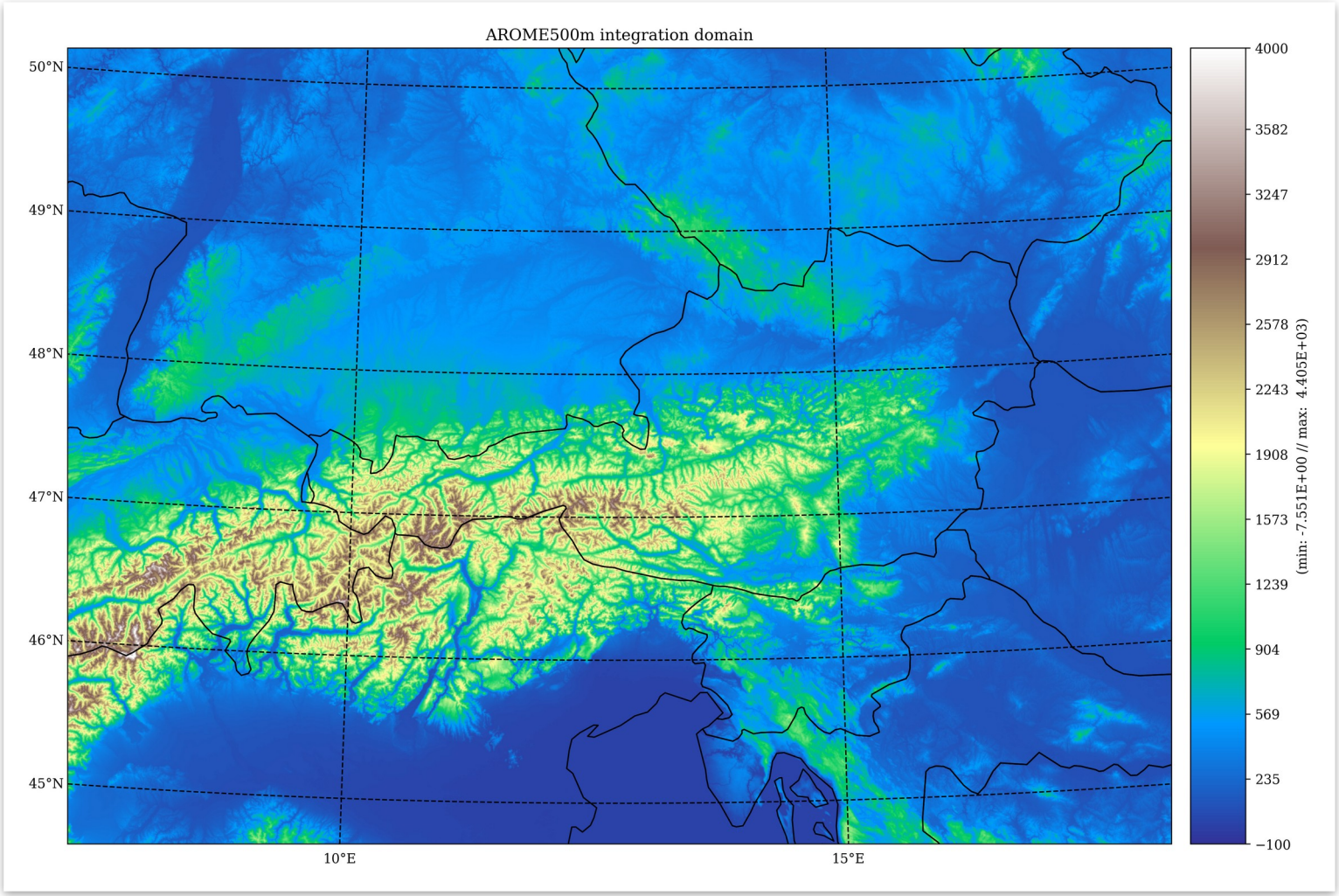
Destination Earth implemented by ECMWF esa EUMETSAT

Overview

1. AROME 500 m – Setup and Tested Cases
2. Methodology for Gridded Verification
3. Results in Detail
4. Summary

Case Selection

- Selection based on station data
- 65 days with observed rr > thres selected from 2016 – 2023
- 48 h forecasts with AROME 500m
- Inits one day before and on the day of the observed peak
- Overall approx. 130 model runs performed on ECMWF ATOS
- Fixed domain covering the (eastern) Alpine region



GeoSphere Austria

AROME 500m and AROME-Aut

AROME 500m setup (Deode prototype v.03):

- AROME cy48t3_deode
- dx, dy= 1km, linear truncation
- 1728 x 1250 grid points
- 90 Levels (lowest 5m)
- Time step: 30s
- Initialization: IFS HRES (3D + surface)
- Coupling: IFS HRES (hourly)
- Modified dyn setup w.r.t. AROME-Aut (e.g. comad, p/c nsiter=2, no slhd, tuning spectral nudging)

AROME-Austria setup:

- AROME cy43t2
- dx, dy= 2.5km, / linear truncation
- 600 x 432 grid points
- 90 Levels (lowest 5m)
- Time step: 60s
- Initialization: 3DVar, OI
- Coupling: IFS HRES (hourly), time

```
NAMDYN:
  LADVF: True
  LQMPD: False
  LQMT: False
  LQMVD: False
  LRHDI_LASTITERPC: True
  NITMP: 4
  NSITER: 2
  NSPDLAG: 3
  NSVDLAG: 3
  NTLAG: 3
  NVLAG: 3
  NWLAG: 3
  RDAMPDIV: 20.0
  RDAMPPD: 20.0
  RDAMPQ: 0.0
  RDAMPT: 0.0
  RDAMPVD: 20.0
  RDAMPVOR: 20.0
  REPS1: 0.0
  REPS2: 0.0
  REPSM1: 0.0
  REPSM2: 0.0
  REPSP1: 0.0
  SDRED: 1.0
  SIPR: 60000.0
  SITR: 350.0
  SITRA: 50.0
  SLHDA0: 0.25
  SLHDD00: 6.5e-05
  VESL: 0.05
  XIDT: 0.0
  ZSLHDP1: 1.7
  ZSLHDP3: 0.6
```

```
NAMDYNA:
  LCOMADH: True
  LCOMADV: False
  LCOMAD_GFL: True
  LCOMAD_SP: True
  LCOMAD_SPD: True
  LCOMAD_SVD: True
  LCOMAD_T: True
  LCOMAD_W: True
  LGWADV: True
  LNESC: True
  LPC_CHEAP: True
  LPC_FULL: True
  LRDBBC: False
  LSETTLS: False
  LSETTLST: True
  LSLHD_GFL: False
  LSLHD_OLD: False
  LSLHD_SPD: False
  LSLHD_SVD: False
  LSLHD_T: False
  LSLHD_W: False
  ND4SYS: 2
  NPDVAR: 2
  NVDVAR: 4
  SLHDEPSH: 0.08
  SLHDKMAX: 6
```

```
NEMELBC0B:
  NEFRSPCPL: 1
  NEK0: 220
  NEK1: 255
  NEN1: 4
  NEN2: 8
  SPNUDDIV: 0.3
  SPNUDQ: 0.
  SPNUDT: 0.3
  SPNUDVOR: 0.3
  TEFRCL: ${namelist.tefrcl}
```

```
gridtype = "linear"
name = "AUSTRIA_CASES"
nimax = 1717
njmax = 1239
tstep = 30
xdx = 500.0
xdy = 500.0
xlat0 = 47.5
xlatcen = 47.5
xlon0 = 12.75
xloncen = 12.75
```

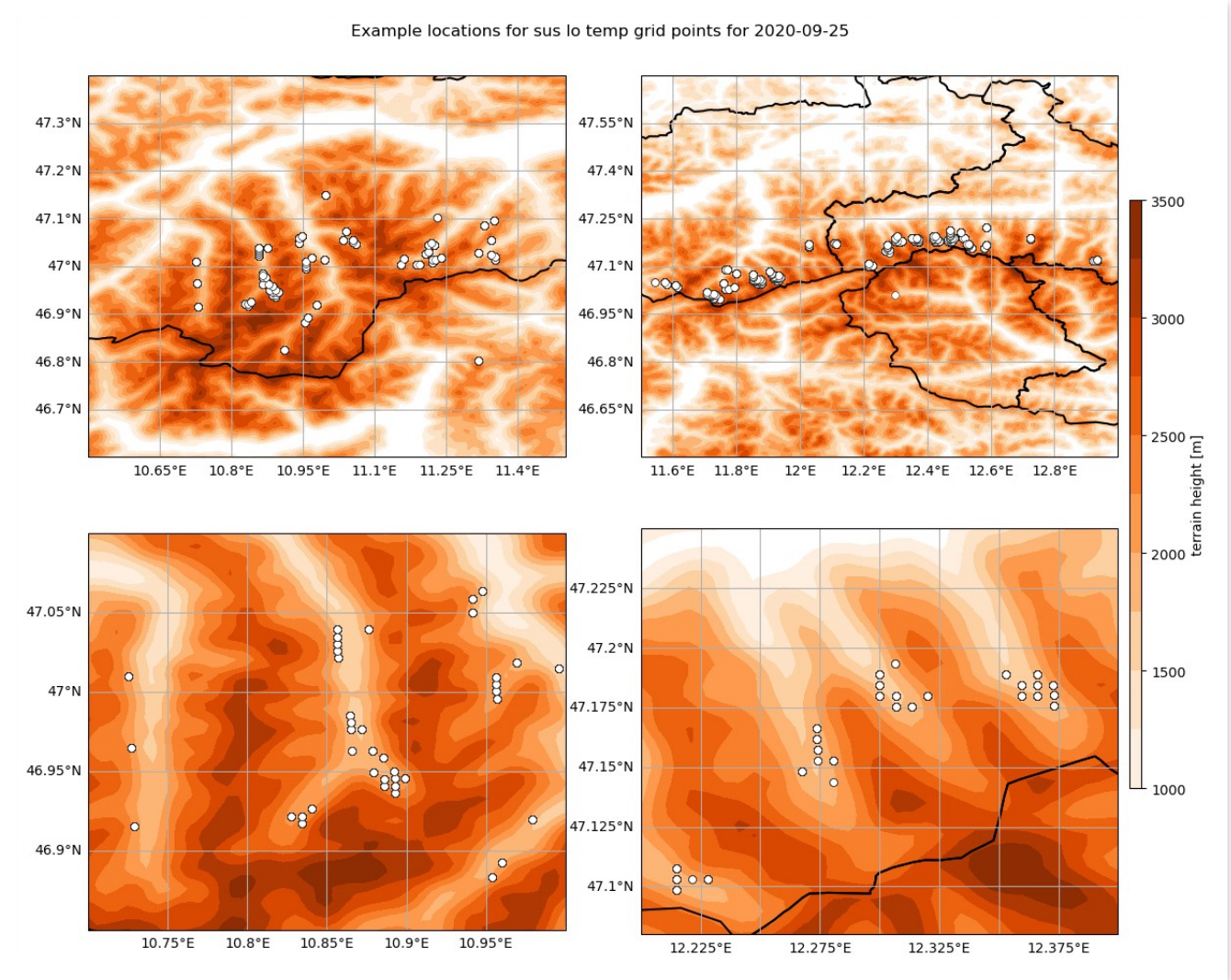
Warning: Comparing AROME-500 vs. AROME-Aut is not fair, but we want to know where we stand compared to operational systems

GeoSphere Austria

Notable Experiences

- No crashes* observed during 130 model runs
- But: SMILAG/ETADOT messages ocurred (more during wintertime)
- Indicates that references setup, in particular tstep=30 is close to the limit for stability
- But **no obvious problems** seen so far in the output fields (visual checks, fields look reasonable)
- Changing to quadratic/cubic truncation and/or reduction of tstep (15-20s) removes SMILAG/ETADOT

* problems with explicit snow scheme (ABORT with „suspicious low temp in layer …“ has to be commented).

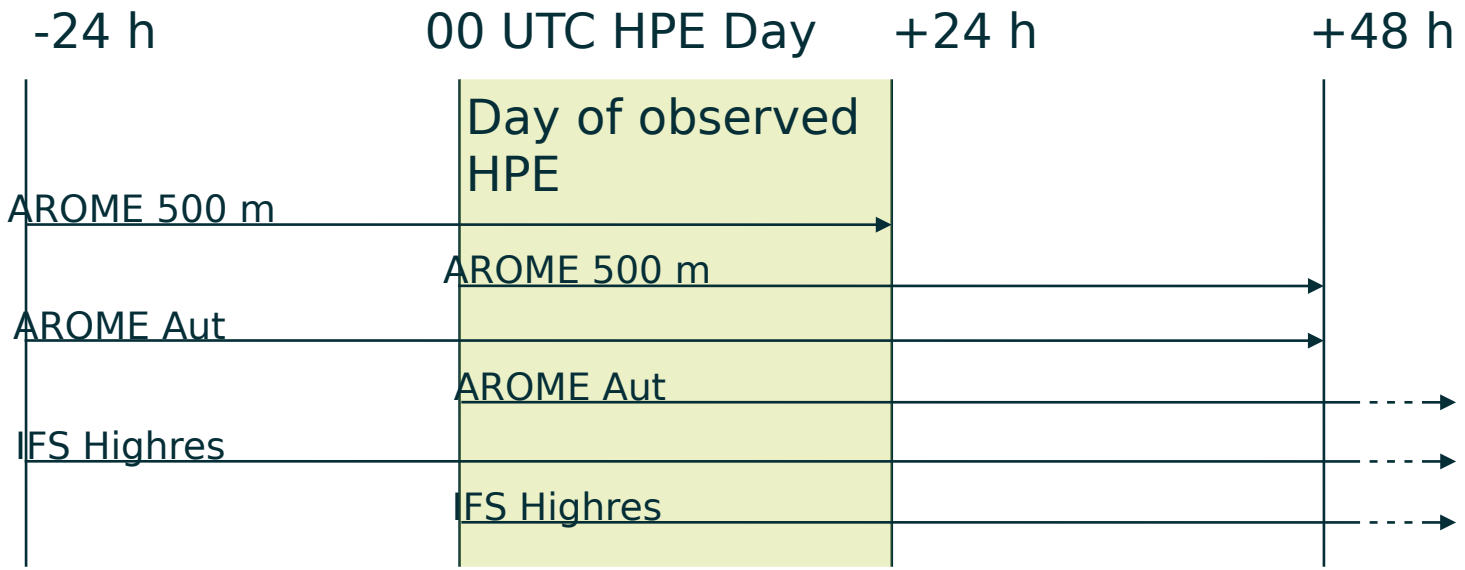


GeoSphere Austria

Verification Methodology

- Each selected day was covered with **two 500 m simulations**:
  1. init 00 UTC the same day
  2. init 00 UTC the day before

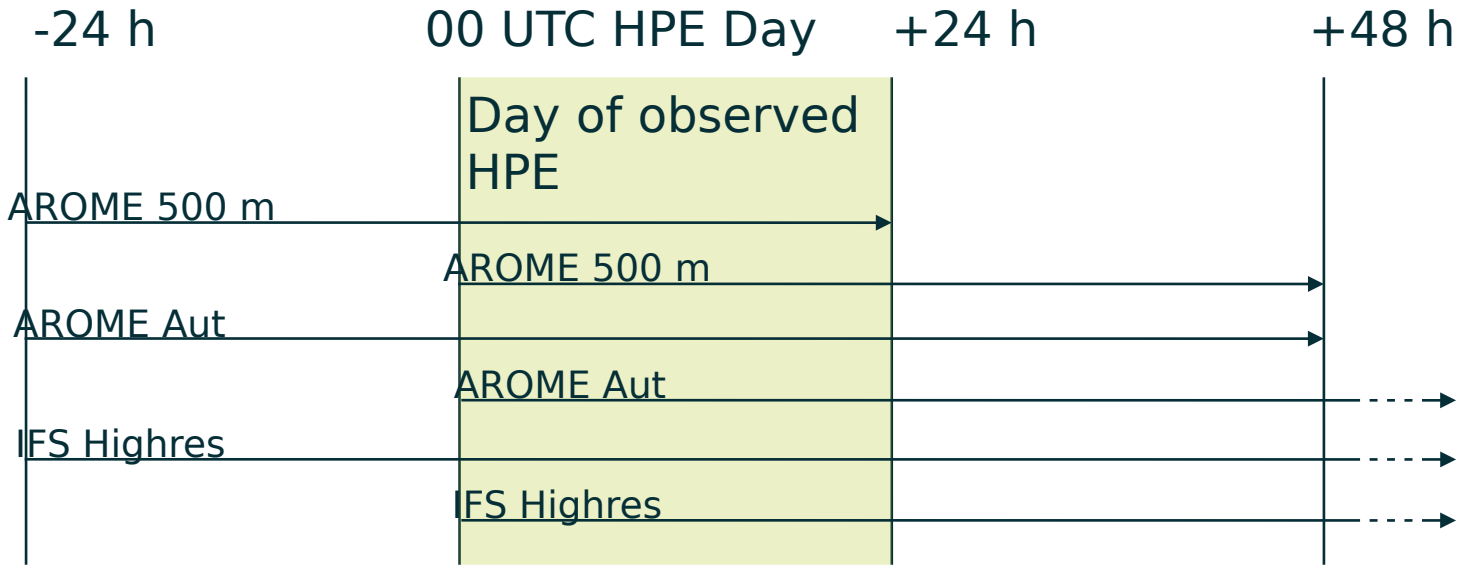
- Verified variable: **24 hour accumulated precipitation of the marked period**



GeoSphere Austria

Verification Methodology

- Each selected day was covered with **two 500 m simulations**:
  1. init 00 UTC the same day
  2. init 00 UTC the day before
- Verified variable: **24 hour accumulated precipitation of the marked period**

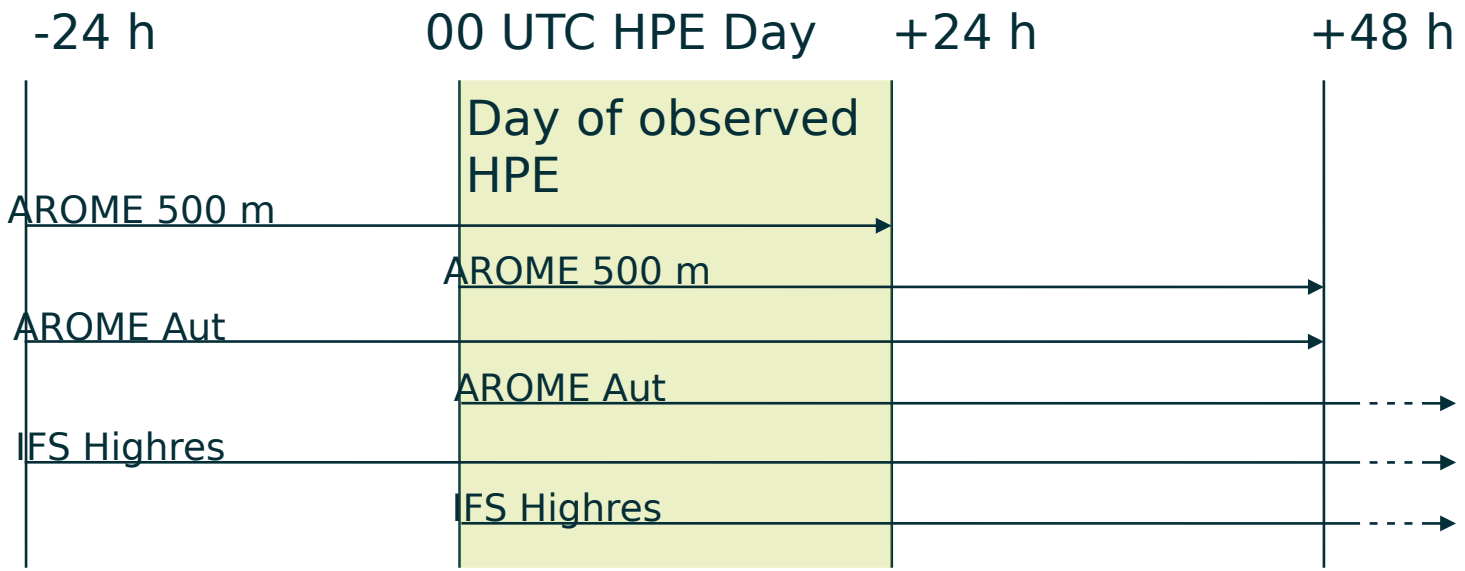


- Each run of Panelification compares these 6 runs, FSS is aggregated over all 65 days.
- **130 runs** of AROME 500, AROME-Aut, and IFS Highres **(390 model runs compared in total)**

GeoSphere Austria

Verification Methodology

- Each selected day was covered with **two 500 m simulations**:
  1. init 00 UTC the same day
  2. init 00 UTC the day before
- Verified variable: **24 hour accumulated precipitation of the marked period**
- Verification metric:
  Fraction Skill Score
  Aggregated FSS
  Relative Aggregated FSS
- Consecutive days of one longer HPE are verified as two separate days



- Each run of Panelification compares these 6 runs, FSS is aggregated over all 65 days.
- **130 runs** of AROME 500, AROME-Aut, and IFS Highres **(390 model runs compared in total)**

GeoSphere Austria

Using the FSS for more than one period

Fraction skill score for one threshold and window siz

$$FSS = \frac{2 f_o f_m}{f_o^2 + f_m^2}$$

f_o ... observed frequency summed over all grid points

f_m ... simulated frequency summed over all grid points

Lowest value: 0

Perfect value: 1

Useful/skillful threshold: **$0.5 * (1 + f_o)$**

Using the FSS for more than one period

Fraction skill score for one threshold and window siz

$$FSS = \frac{2 f_o f_m}{f_o^2 + f_m^2}$$

f_o ... observed frequency summed over all grid points

f_m ... simulated frequency summed over all grid points

Lowest value: 0

Perfect value: 1

Useful/skillful threshold: **0.5 * (1 + f_o)**

$$FSS_{Averaged} = \sum_{t=0}^{t_n} \frac{2 f_{o,t} f_{m,t}}{f_{o,t}^2 + f_{m,t}^2}$$

Averaging the FSS gives **equal weight to all events** (periods), which means that weaker precipitation periods can heavily distort results for higher precipitation thresholds.

Using the FSS for more than one period

Fraction skill score for one threshold and window siz

$$FSS = \frac{2 f_o f_m}{f_o^2 + f_m^2}$$

f_o … observed frequency summed over all grid points

f_m … simulated frequency summed over all grid points

Lowest value: 0

Perfect value: 1

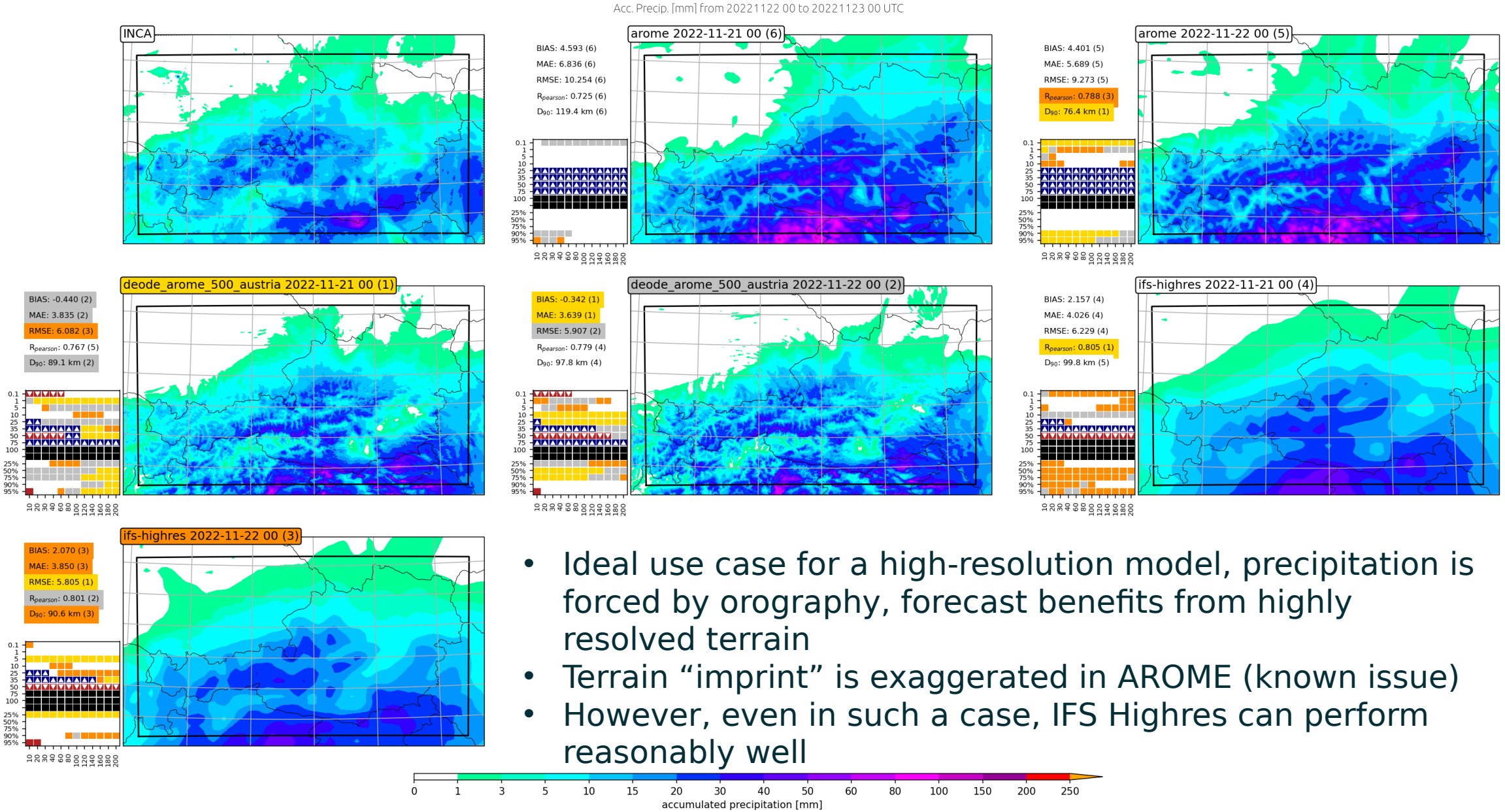
Useful/skillful threshold: **0.5 * (1 + f_o)**

$$FSS_{Averaged} = \sum_{t=0}^{t_n} \frac{2 f_{o,t} f_{m,t}}{f_{o,t}^2 + f_{m,t}^2}$$

Aggregating the FSS **weights per grid point, at which the threshold is exceeded**. However, this weights more widespread events more heavily.

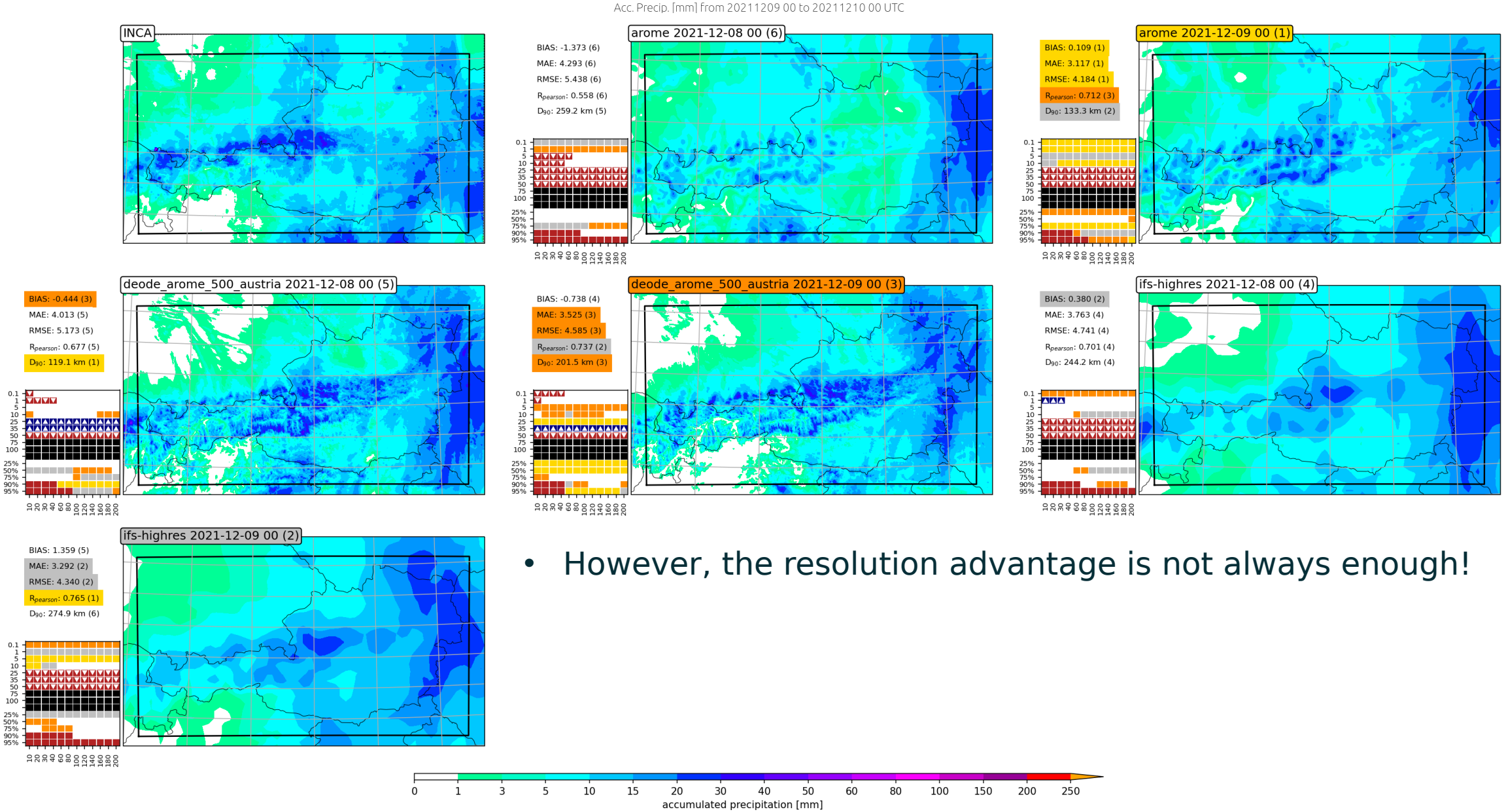
$$FSS_{Aggregated} = \frac{\sum_{t=0}^{t_n} 2 f_{o,t} f_{m,t}}{\sum_{t=0}^{t_n} f_{o,t}^2 + \sum_{t=0}^{t_n} f_{m,t}^2}$$

21 November 2022 – large scale, orographically forced



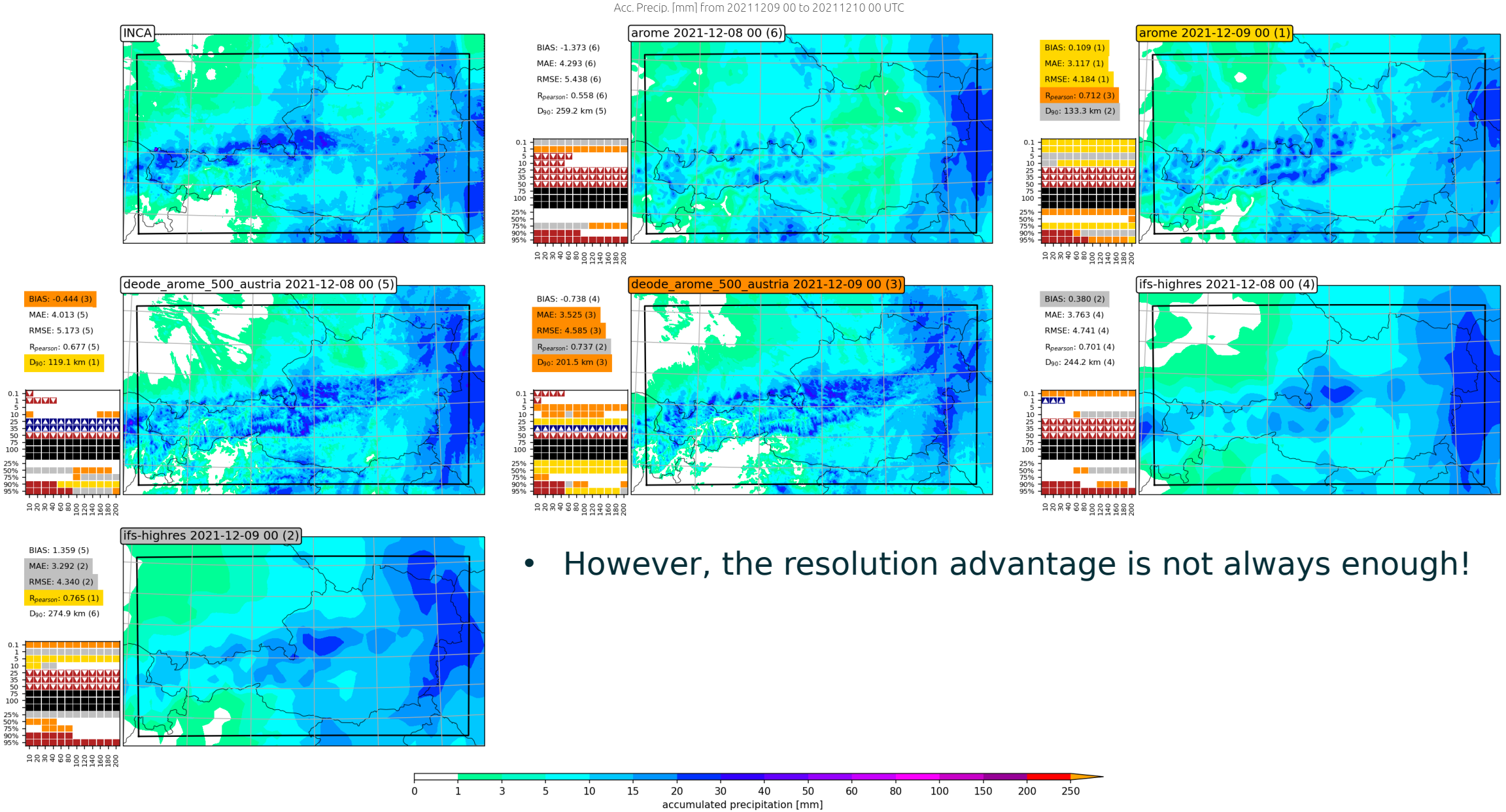
- Ideal use case for a high-resolution model, precipitation is forced by orography, forecast benefits from highly resolved terrain
- Terrain “imprint” is exaggerated in AROME (known issue)
- However, even in such a case, IFS Highres can perform reasonably well

9 December 2022 – large scale, orographically forced



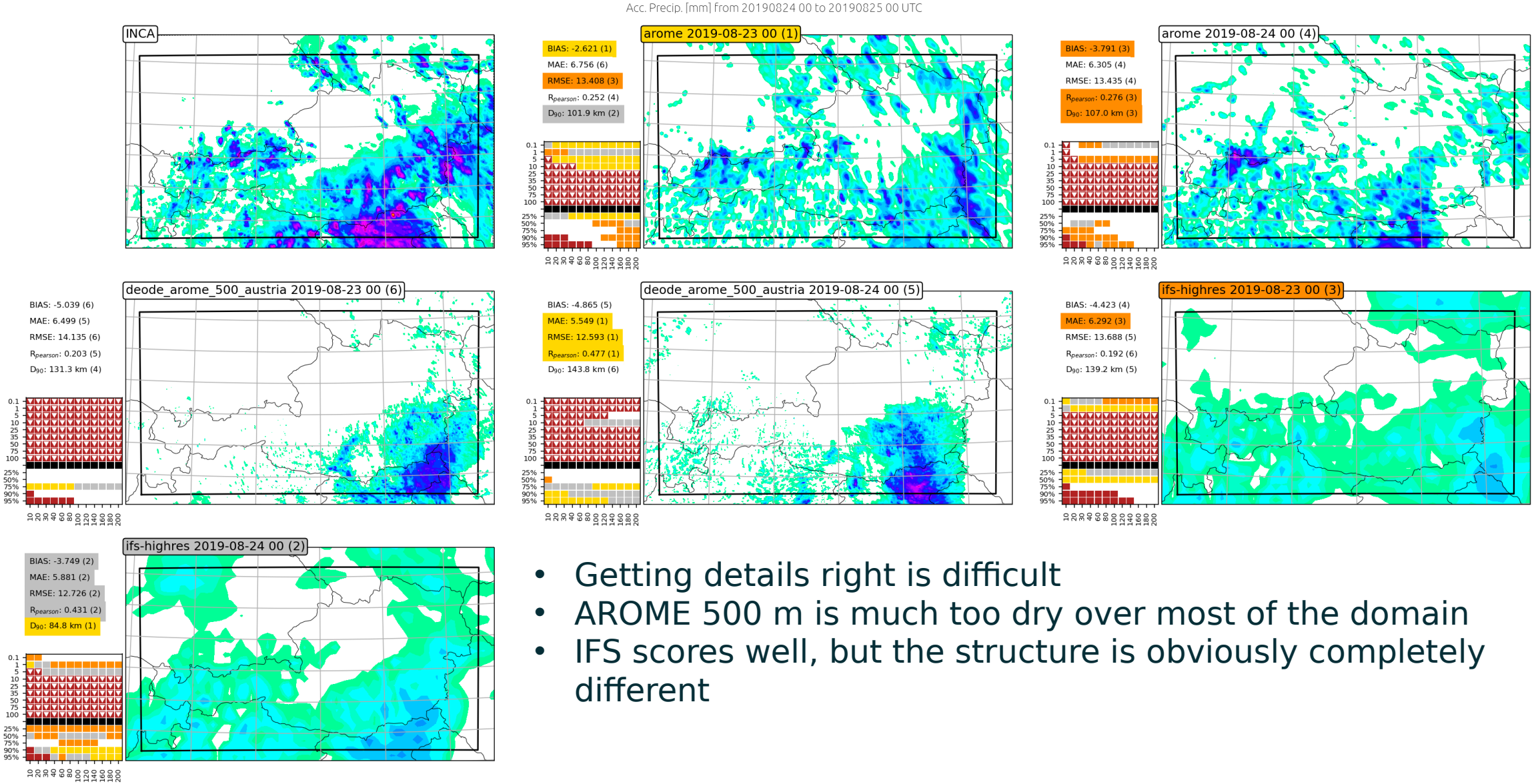
- However, the resolution advantage is not always enough!

9 December 2022 – large scale, orographically forced



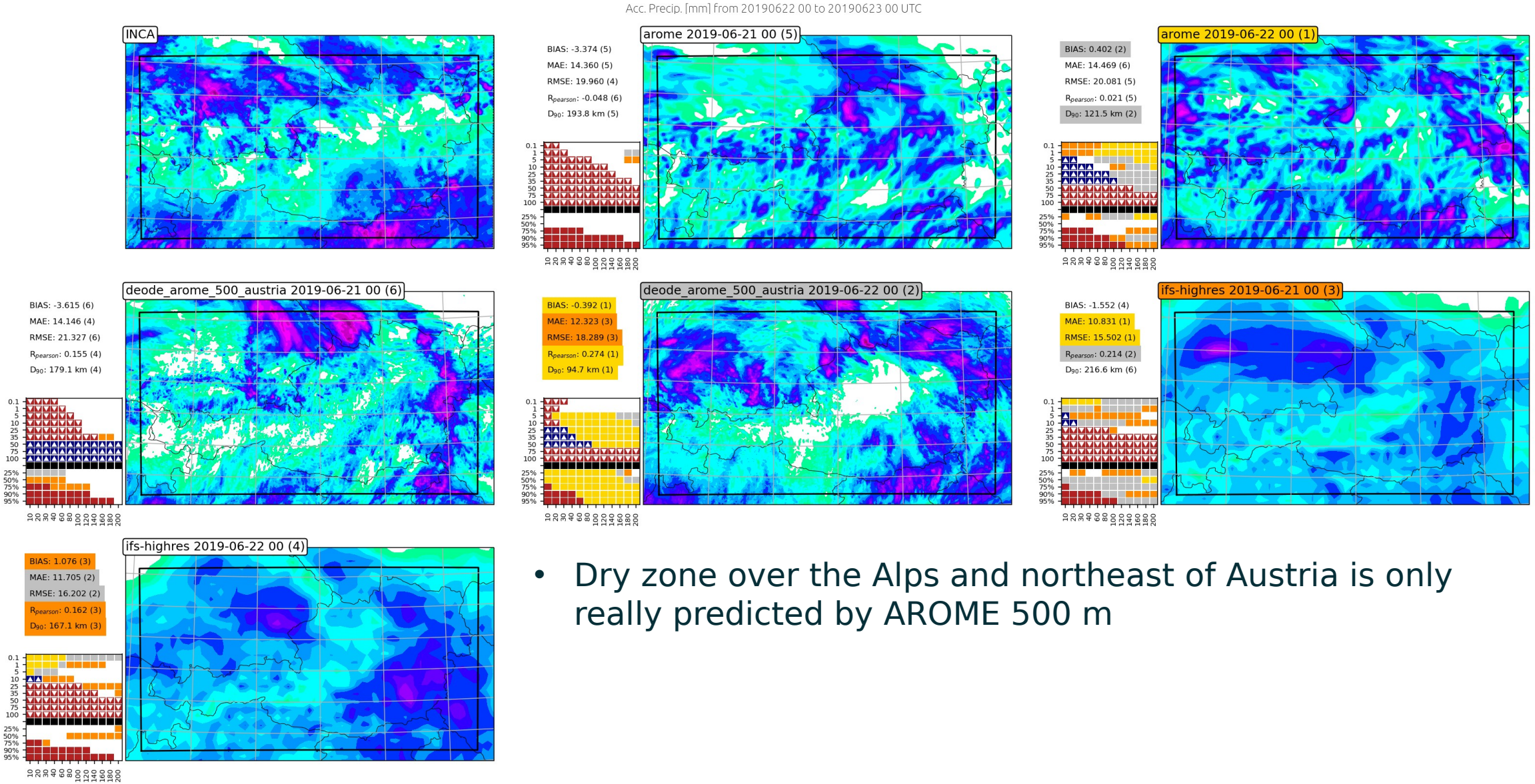
- However, the resolution advantage is not always enough!

24 August 2019 – scattered convection



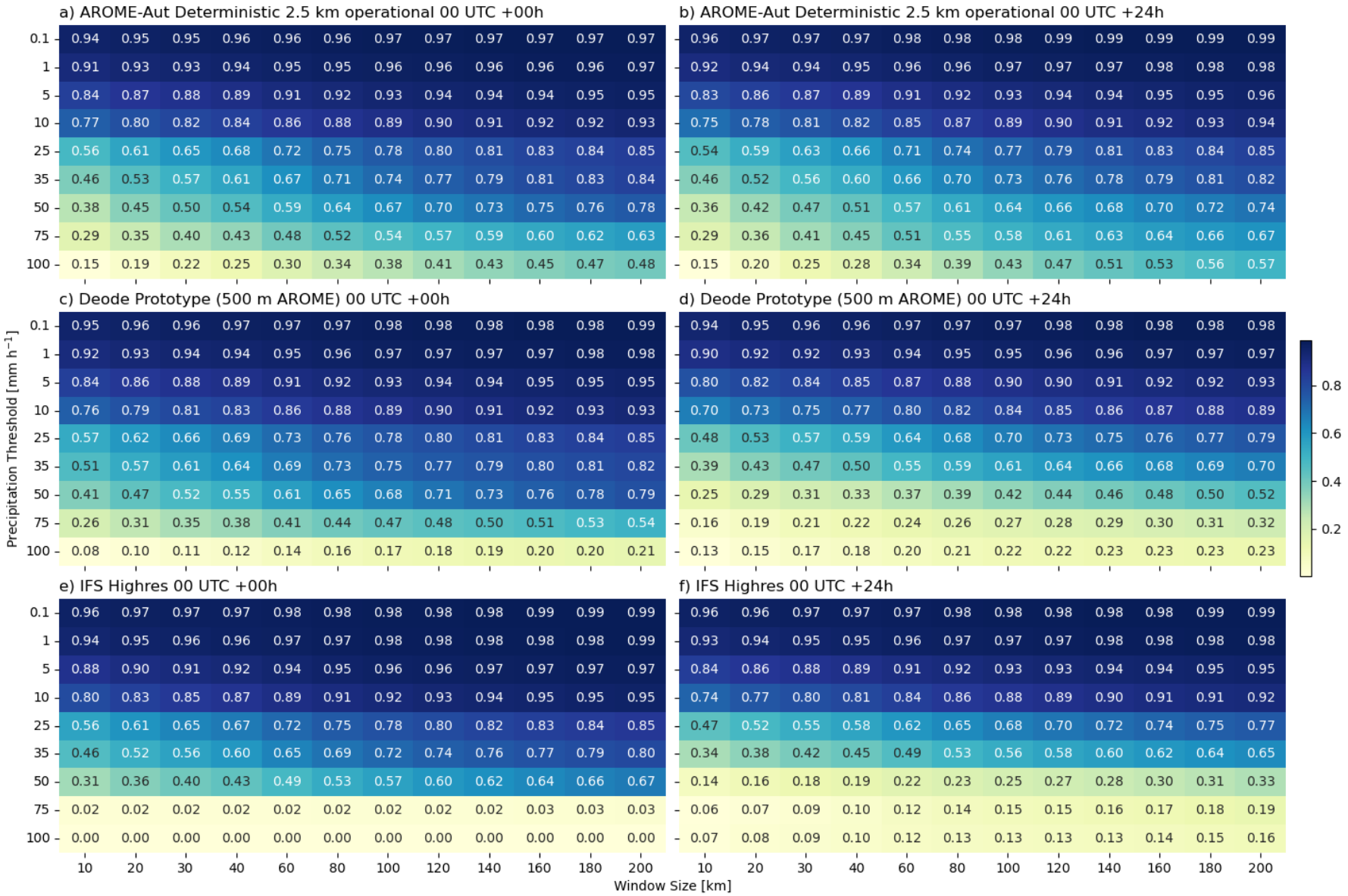
- Getting details right is difficult
- AROME 500 m is much too dry over most of the domain
- IFS scores well, but the structure is obviously completely different

22 June 2019 – widespread convection



- Dry zone over the Alps and northeast of Austria is only really predicted by AROME 500 m

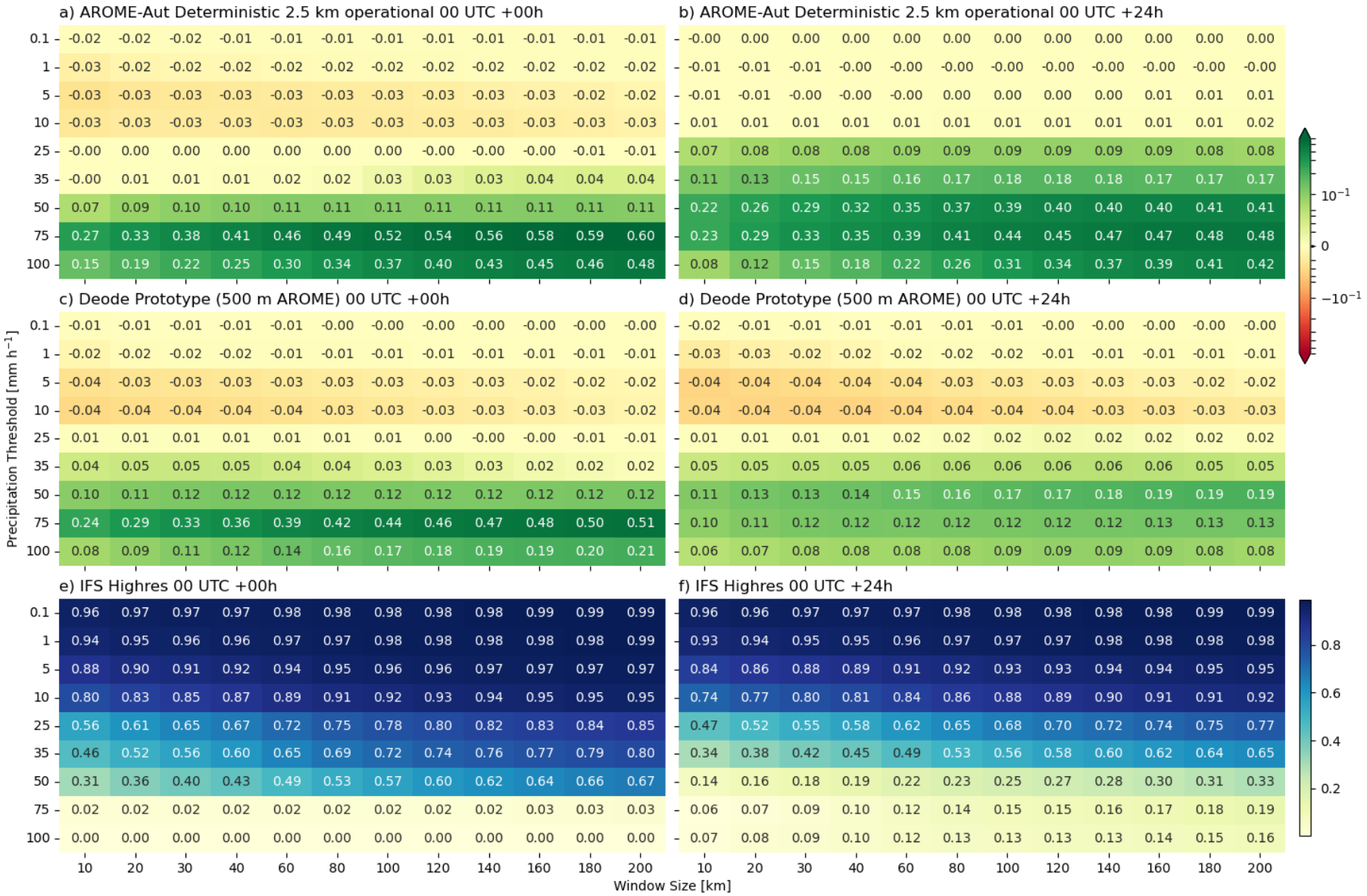
Overall Aggregated FSS



- Full tables are tedious to compare
- Useful/skillful threshold is not obvious for aggregated values, average over all cases would be an option
- For high thresholds, scores are often well below 0.5
- So what can we do?

GeoSphere Austria

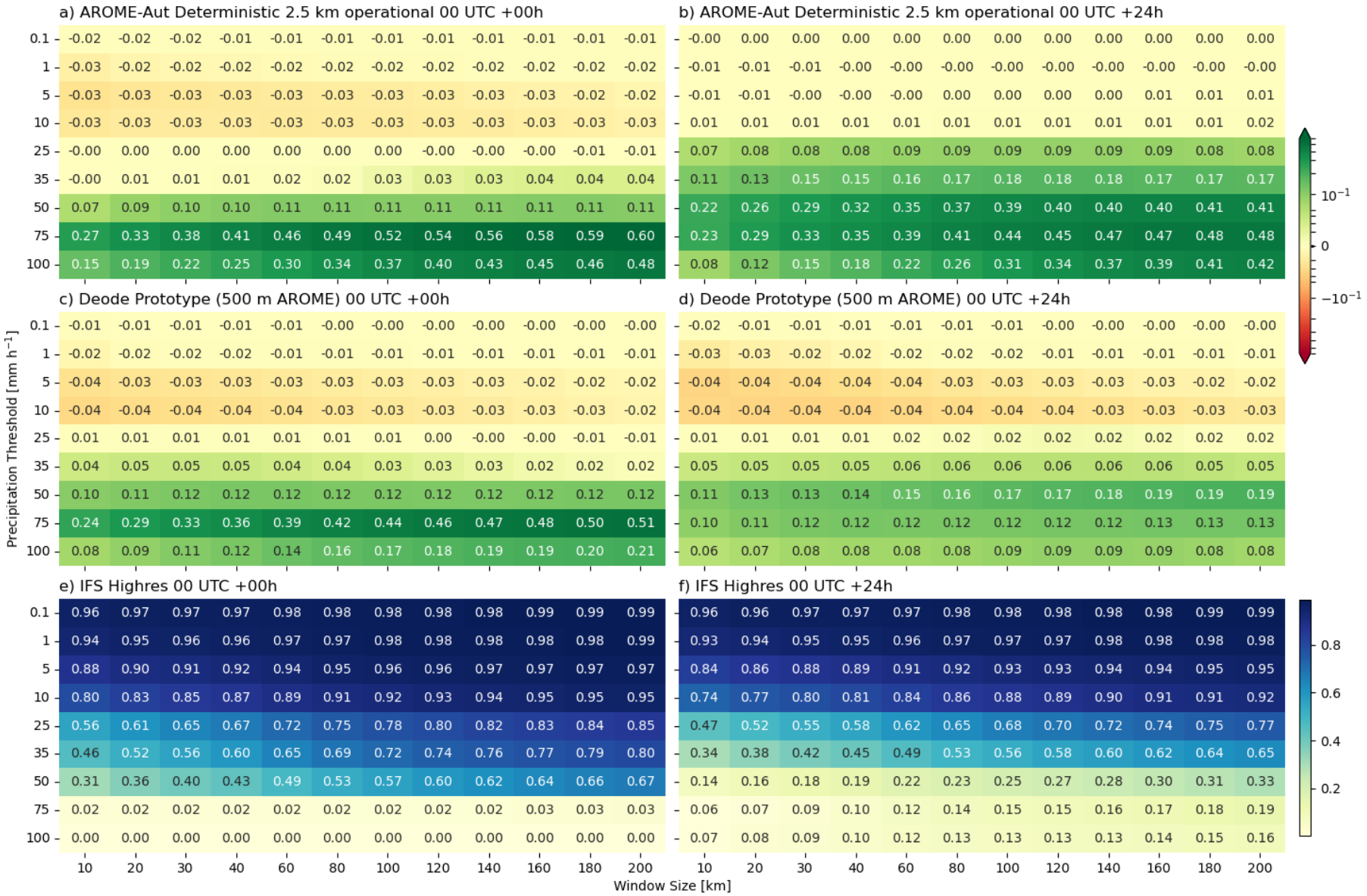
Overall Aggregated FSS



- Look at the differences between the models!
- With the right coloring, improvement and deterioration are easily visible

GeoSphere Austria

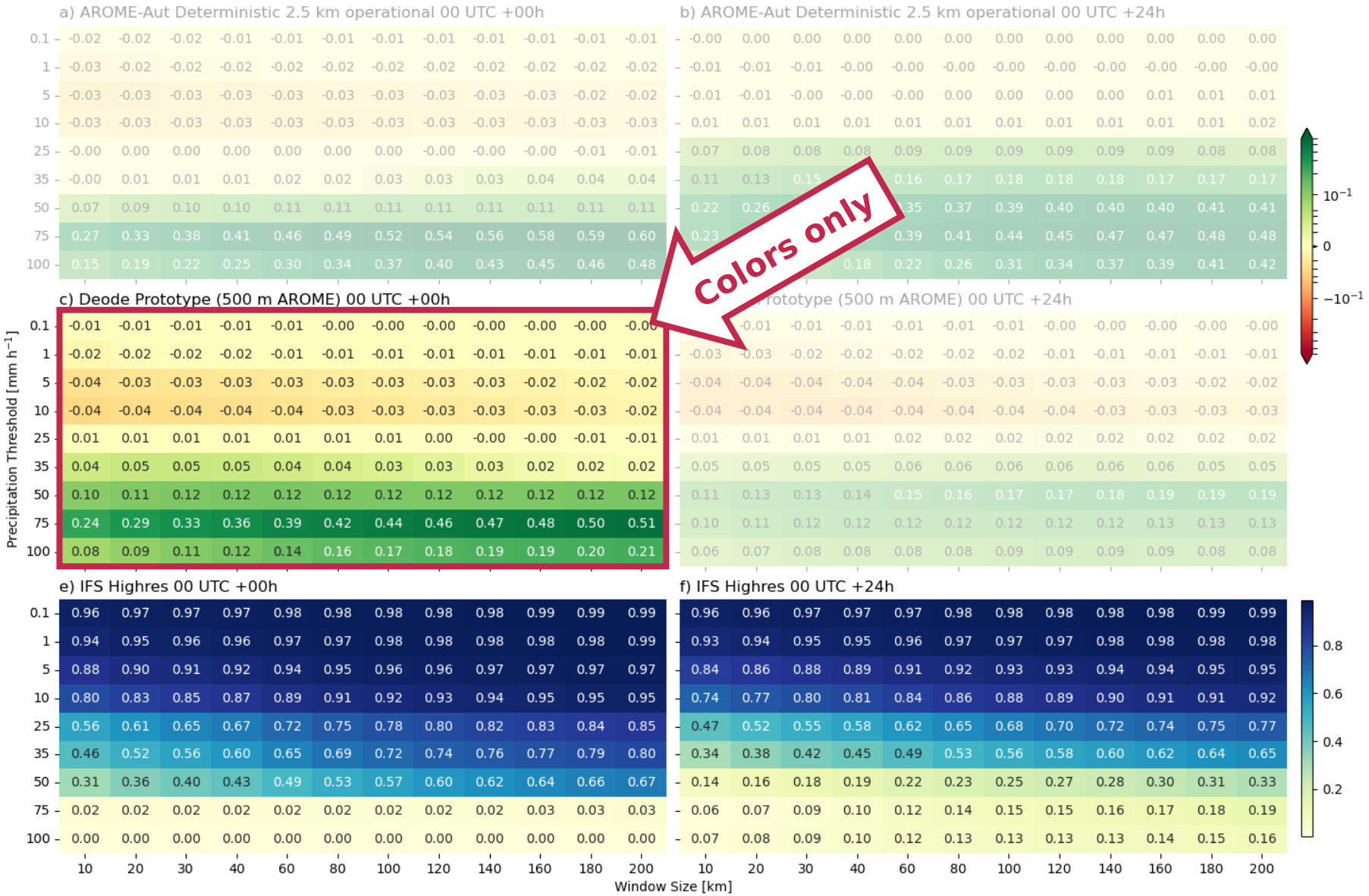
Overall Aggregated FSS



- Look at the differences between the models!
- With the right coloring, improvement and deterioration are easily visible
- AROME 500 m performs much better than IFS Highres for higher precipitation thresholds
- In comparison, only a small deterioration is seen for lower values.

GeoSphere Austria

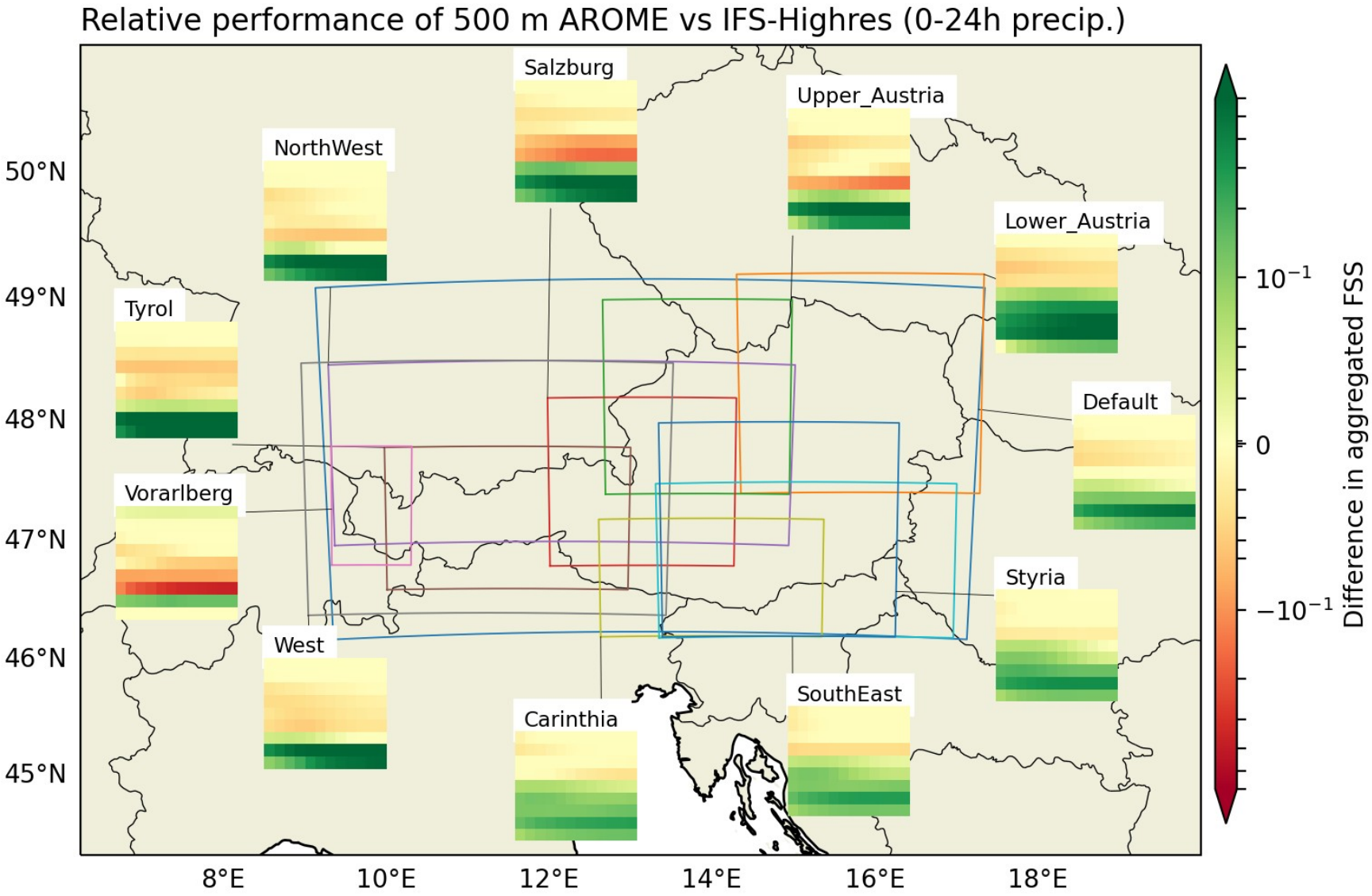
Overall Aggregated FSS



GeoSphere Austria

Results for smaller regions: 0 – 24 h forecasts

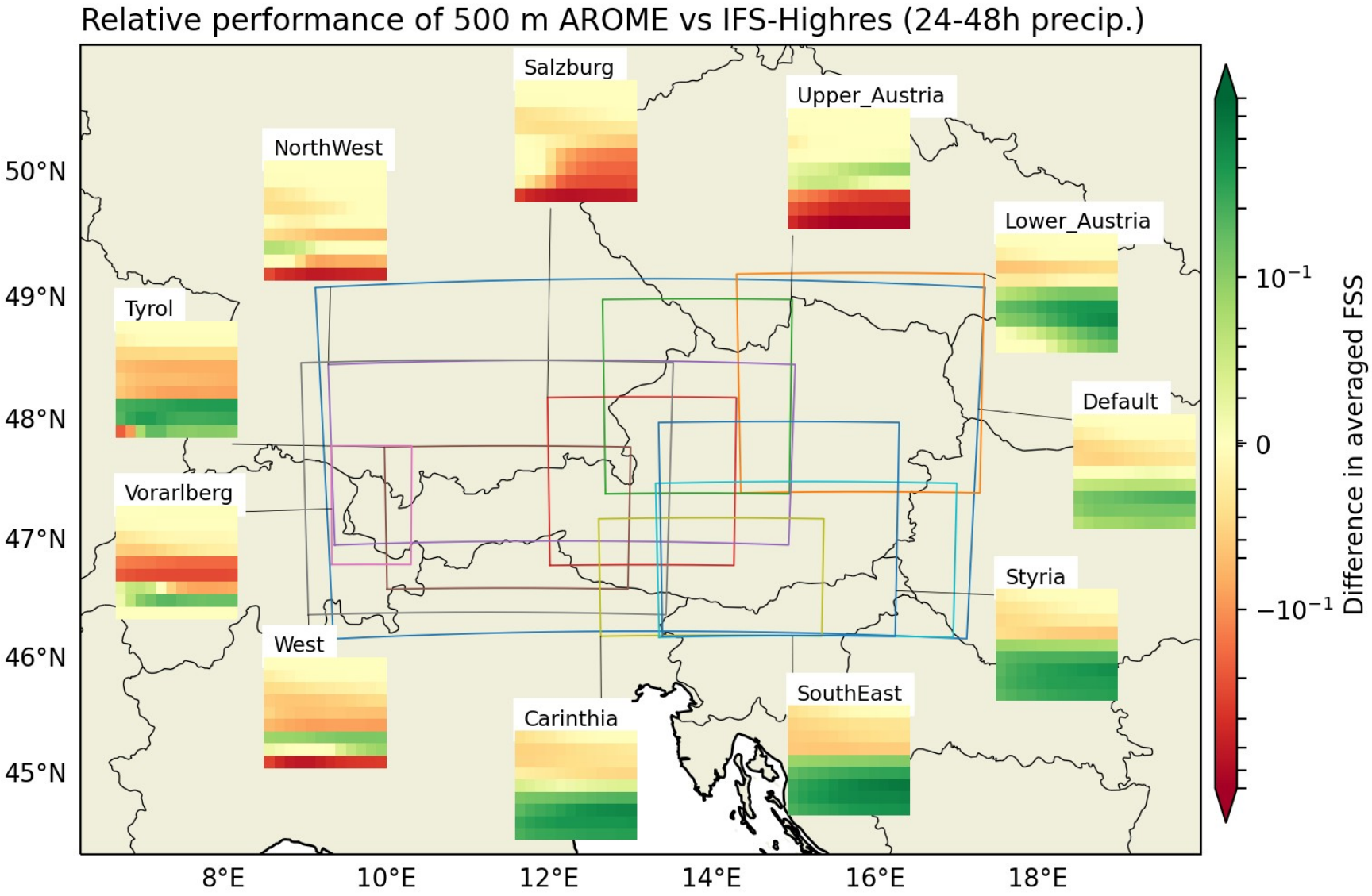
- Verification was repeated for multiple subdomains to examine the added value
- Smaller sample (of grid points) shows higher variability
- South and southeast is captured better than West and North.



GeoSphere Austria

Results for smaller regions: 24 – 48 h forecasts

- East-west contrast is more pronounced for 24 – 48 h predictions
- Some regions clearly do not benefit from 500 m resolution at these leadtimes
- Larger sample will be needed for definitive results



GeoSphere Austria

Summary

The **AROME 500 m simulations have a qualitative and quantitative added value** compared to IFS Highres for heavy precipitation forecasts.

Since the sample is still low at 65, the numbers for smaller regions are less clear.

Added value for the evaluated sample is **lower in the west and north but higher in the south and southeast** of Austria.

The **results of the operational AROME-Aut 2.5 km deterministic model are still better**. While this comparison is not entirely fair, it should still be used as ultimate benchmark (+ IFS HRES, GDT).

Funded by the European Union Destination Earth implemented by ECMWF esa EUMETSAT