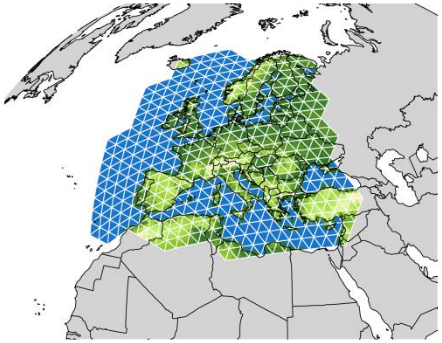
News on ICON-NWP

Recent model improvements and systematic resolution-dependence tests to determine needs for further development





Günther Zängl, on behalf of the ICON development team

COSMO-GM (online), autumn 2021

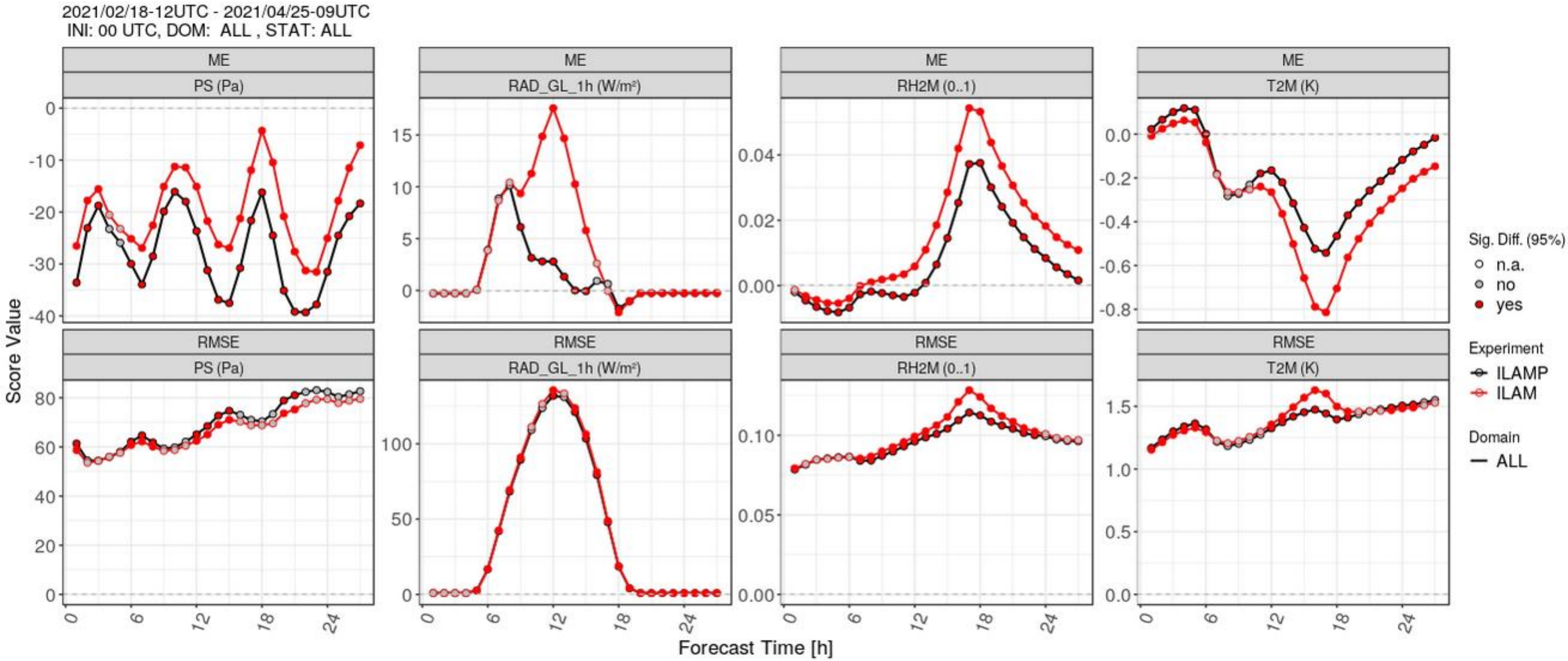
Overview

- Recent ICON-D2 model upgrades and their benefit for forecast quality
- Resolution-dependence of forecast quality: systematic investigations for global and regional scales
- Conclusions

Recent ICON-D2 model upgrades at DWD

- ICON-D2 became operational on February 10
- The major model physics upgrade announced last year (including ecRad and the „cp/cv“ energy conservation bugfix in the physics-dynamics coupling of the turbulence scheme) became operational on April 14 for all ICON configurations
- Extended coupling between model tuning parameters / boundary conditions and data assimilation was introduced on May 26 in order to further reduce T2M and pressure biases
- The forecast lead time was extended from 27 h to 48 h on June 23

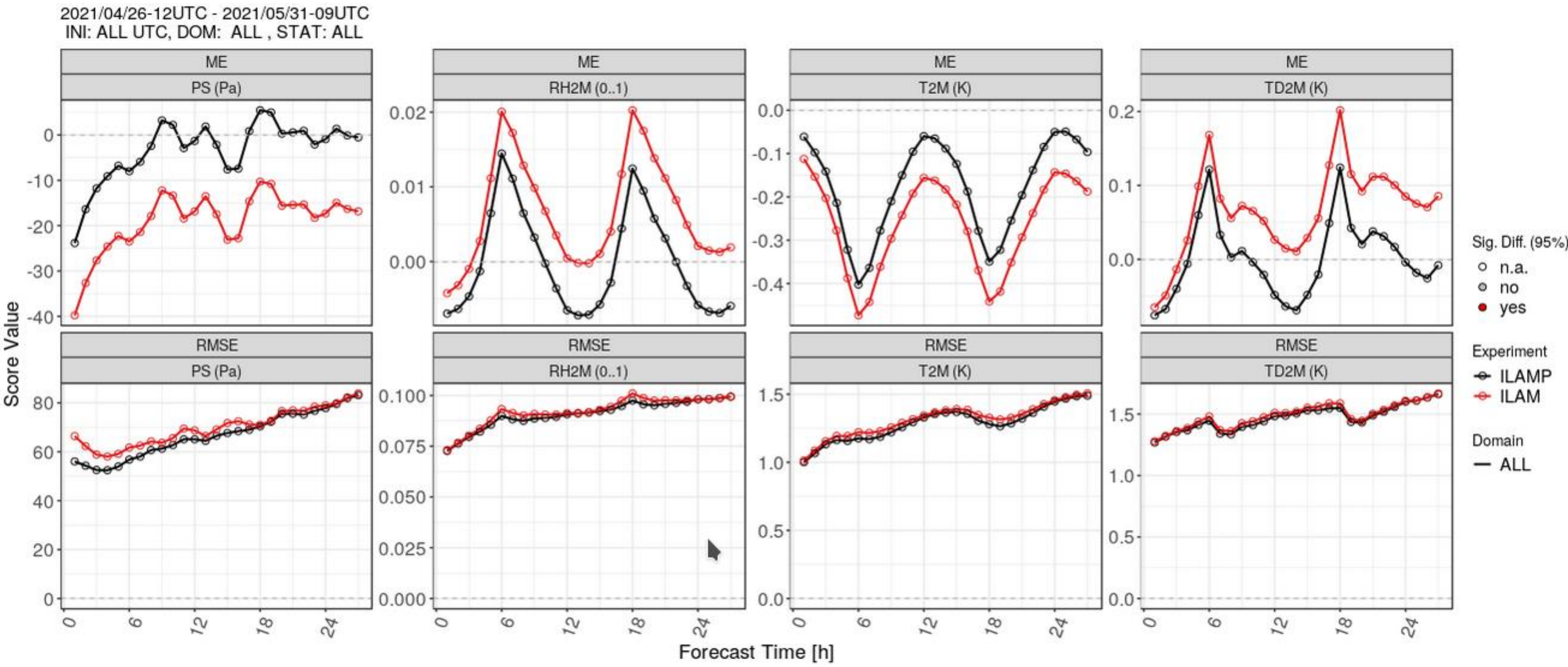
New vs. old physics configuration, Feb – Apr 2021



sfc pressure | global radiation | 2m humidity | 2m temperature

Extended model-DA coupling

- **Use time-filtered domain average of pressure increment at lowest model level to shift pressure imposed at lateral boundaries**
- **Use time-filtered temperature increment (in addition to RH increment) to dynamically adapt model parameters affecting surface evaporation (requires assimilation of T2M and RH2M!)**

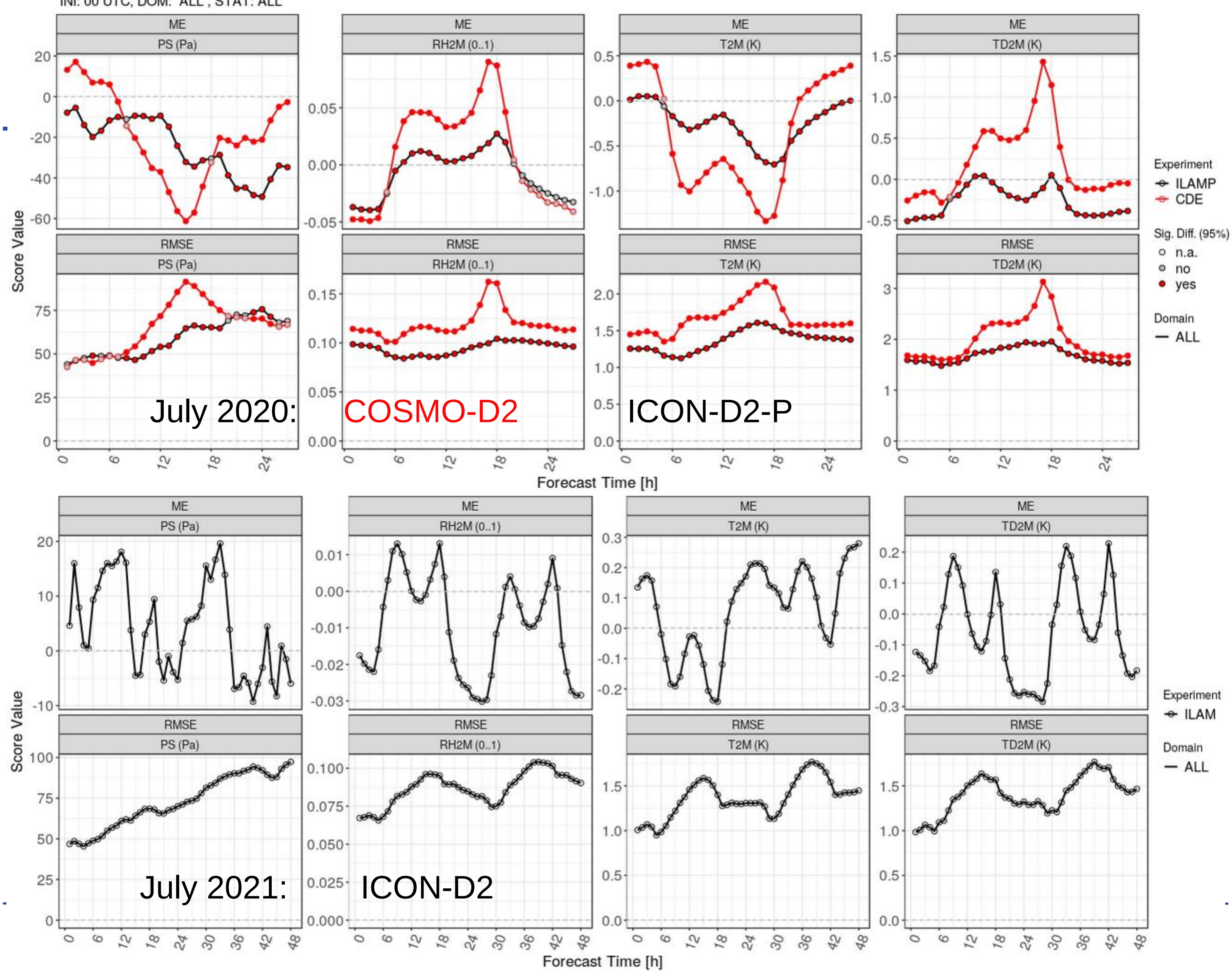


Estimation of total benefit achieved with the change from COSMO-D2 to ICON-D2

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- The following comparison considers COSMO-D2 for July 2020, ICON-D2-P for July 2020, and ICON-D2 routine for July 2021
- Note that the forecast lead time is 48 h in the latter case
- The most important differences between ICON-D2-P (2020) and ICON-D2 (2021) are the new physics configuration, the assimilation of T2M and RH2M, and the adaptive parameter tuning building upon that

- Remark: the pressure bias correction has no direct dynamical impact but improves the performance of the LETKF by acting against the anticorrelation between surface pressure bias and lower-tropospheric temperature bias

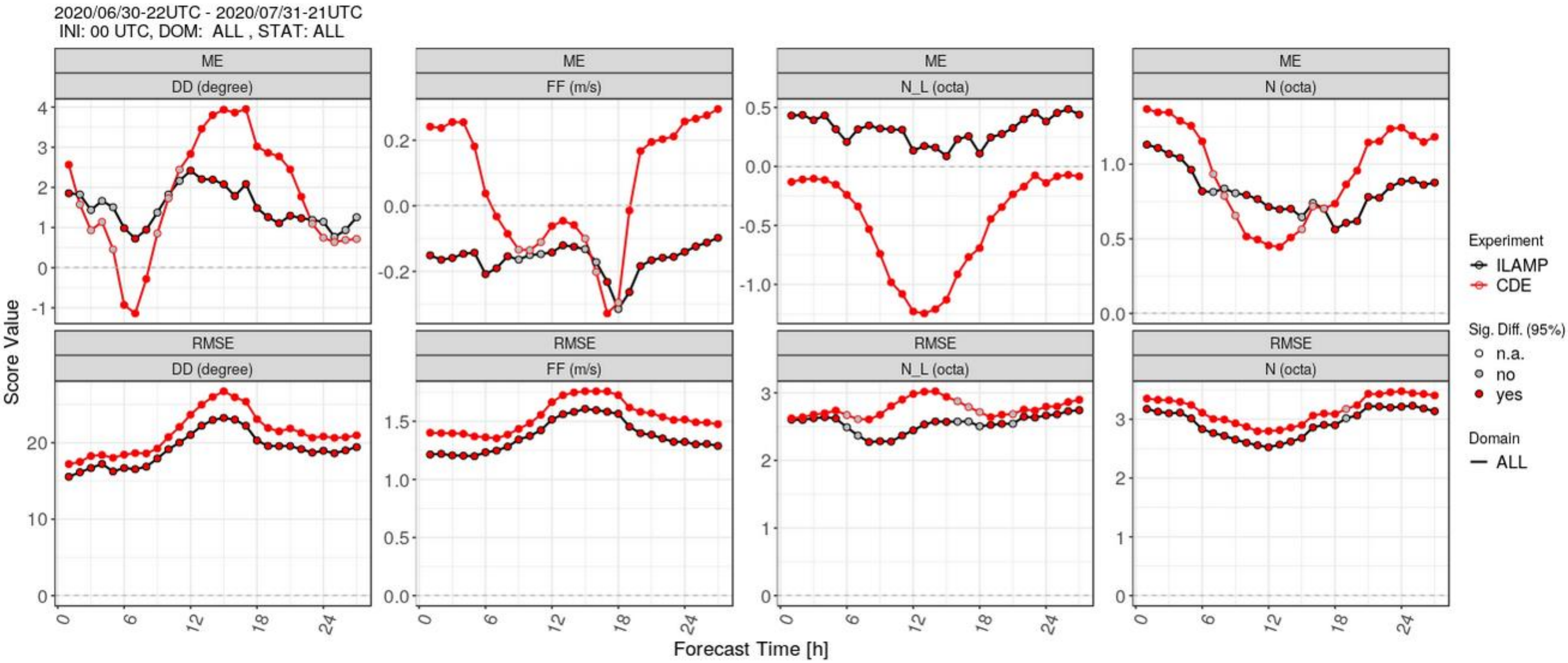

INI: 00 UTC, DOM: ALL , STAT: ALL
ME
PS (Pa)
RH2M (0..1)
T2M (K)
TD2M (K)
RMSE
Score Value
Experiment
ILAMP
CDE
Sig. Diff. (95%)
n.a.
no
yes
Domain
ALL
July 2020:
COSMO-D2
ICON-D2-P
Forecast Time [h]
July 2021:
ICON-D2
Experiment
ILAM
Domain
ALL
Forecast Time [h]

Estimation of total benefit achieved with the change from COSMO-D2 to ICON-D2

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Most pronounced improvements

- ➔ Much smaller amplitudes of the diurnal cycle bias
- ➔ Much smaller RMSE; reduction of T2M and RH2M errors by about a factor of 1.5 compared to COSMO-D2
- ➔ As shown previously, wind and cloud cover are improved as well



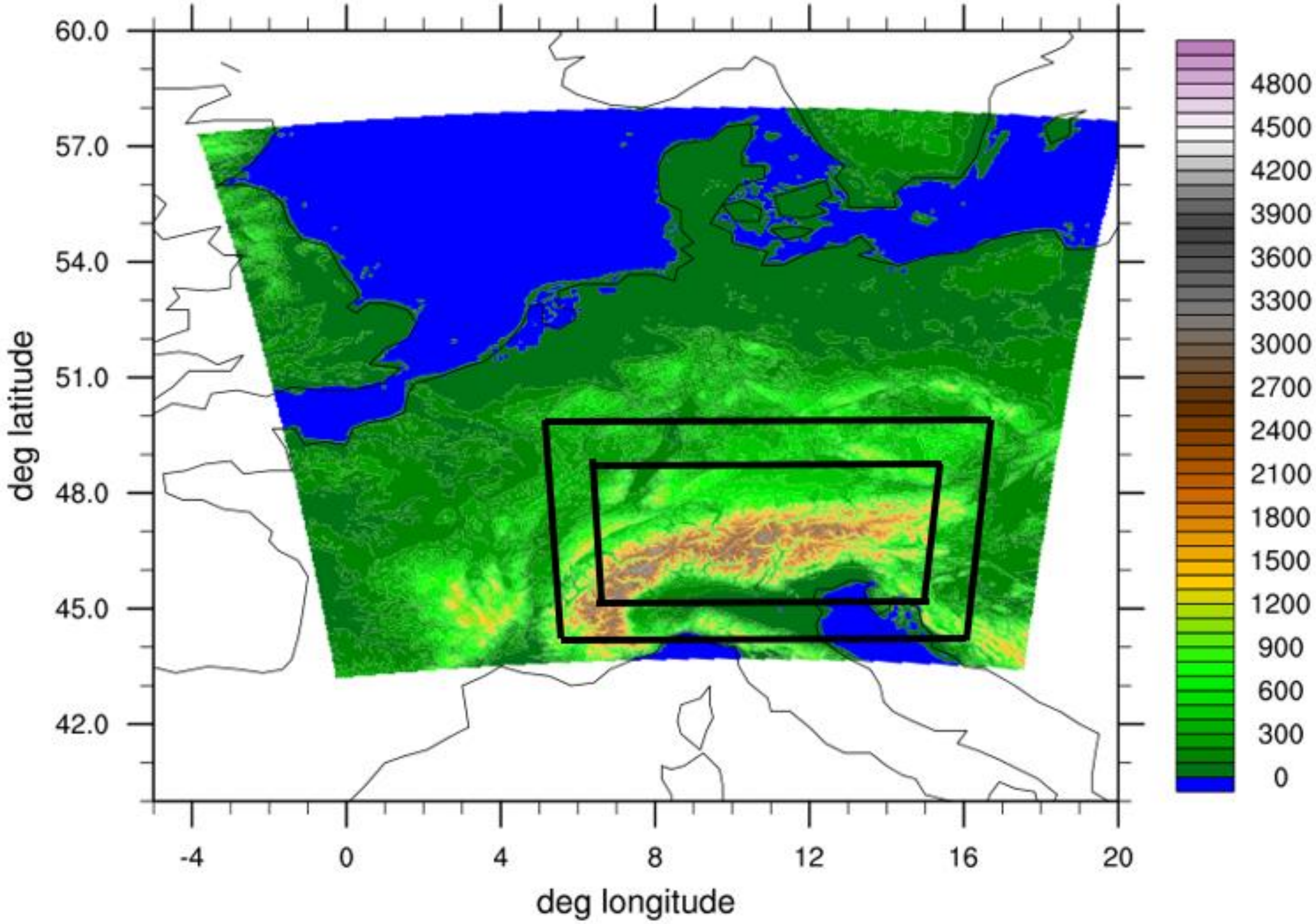
Resolution-dependence of forecast quality: Tests with higher resolution over the Alps

- **Motivation: Upcoming Alpine field campaign TEAMx**
- **Two-step nesting approach incorporated into the ICON-D2 domain, with a mesh size of ~500 m over the Alps**

ICON-D2:
2.08 km,
542000 gp

1st nest:
1.04 km,
543300 gp

2nd nest:
520 m,
946000 gp



Design of the model experiments

- Hindcast experiments running continuously over one month (January 2019 and June 2020), driven by data from the ICON-EU assimilation cycle
- Continuous one-month forecast highlights systematic model errors better than an assimilation cycle with short-range forecasts
- On the other hand, this approach still allows deterministic verification against observations for small model domains like that of ICON-D2

Scientific questions:

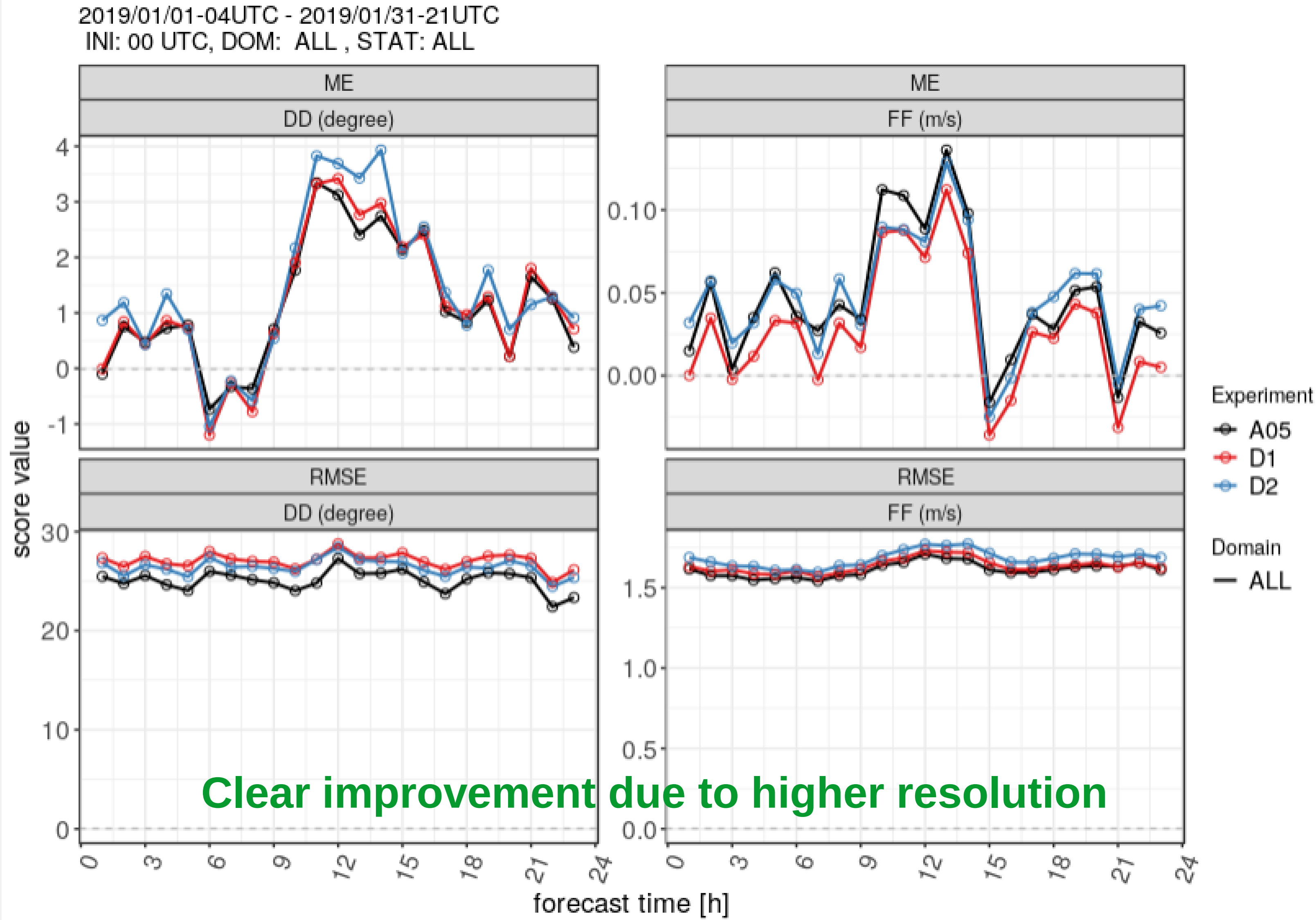
- Which forecast variables benefit directly from increased model (in particular, topography) resolution?
- Which set of parameterizations provides optimal forecast quality?
- At which components is further model development needed?

Tests discussed in the following

- **Systematic evaluation of resolution-dependence of forecast quality over the Alps in the range 2 km – 500 m**
- **Parameterization of orographic form drag: blocking part of SSO scheme (Lott and Miller, 1997) vs. TOFD scheme (Beljaars et al., 2006); parameterization needed at all at 500 m?**

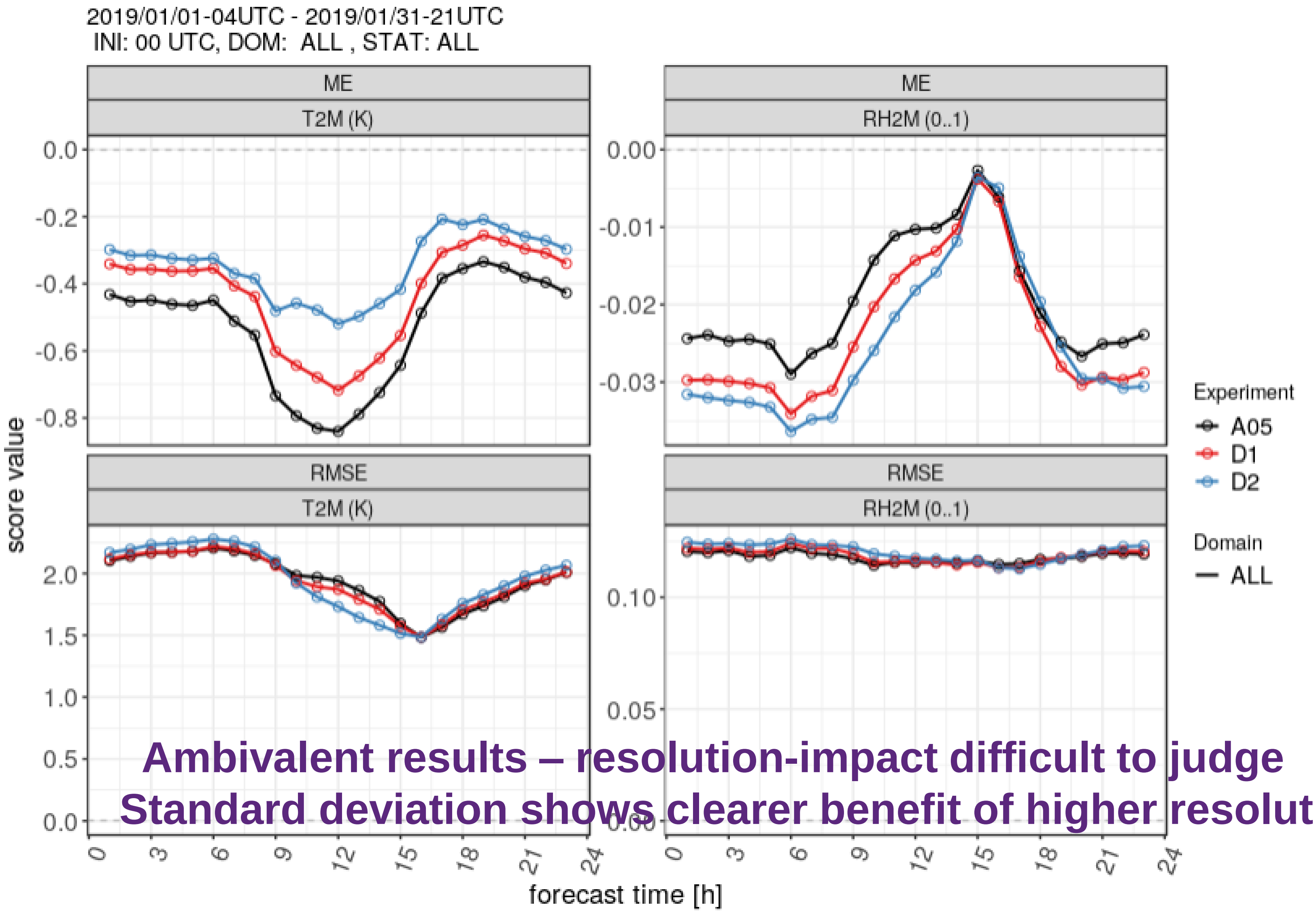
Resolution-dependence: 500 m, 1 km, 2 km wind direction/speed, January 2019

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Clear improvement due to higher resolution

Resolution-dependence: 500 m, 1 km, 2 km
2 m temp./humidity, January 2019

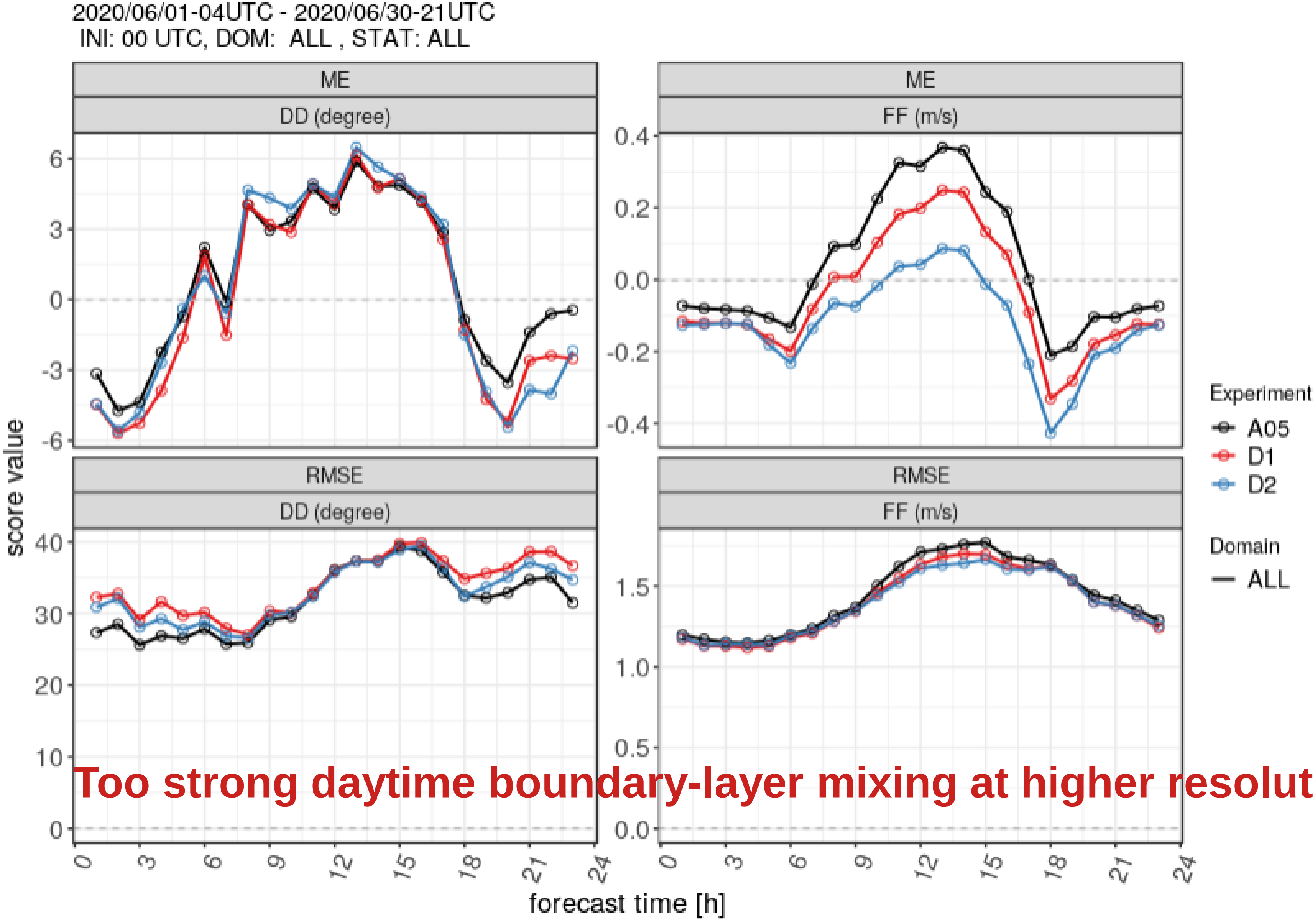


Ambivalent results – resolution-impact difficult to judge

Standard deviation shows clearer benefit of higher resolution

Resolution-dependence: 500 m, 1 km, 2 km
wind direction/speed, June 2020

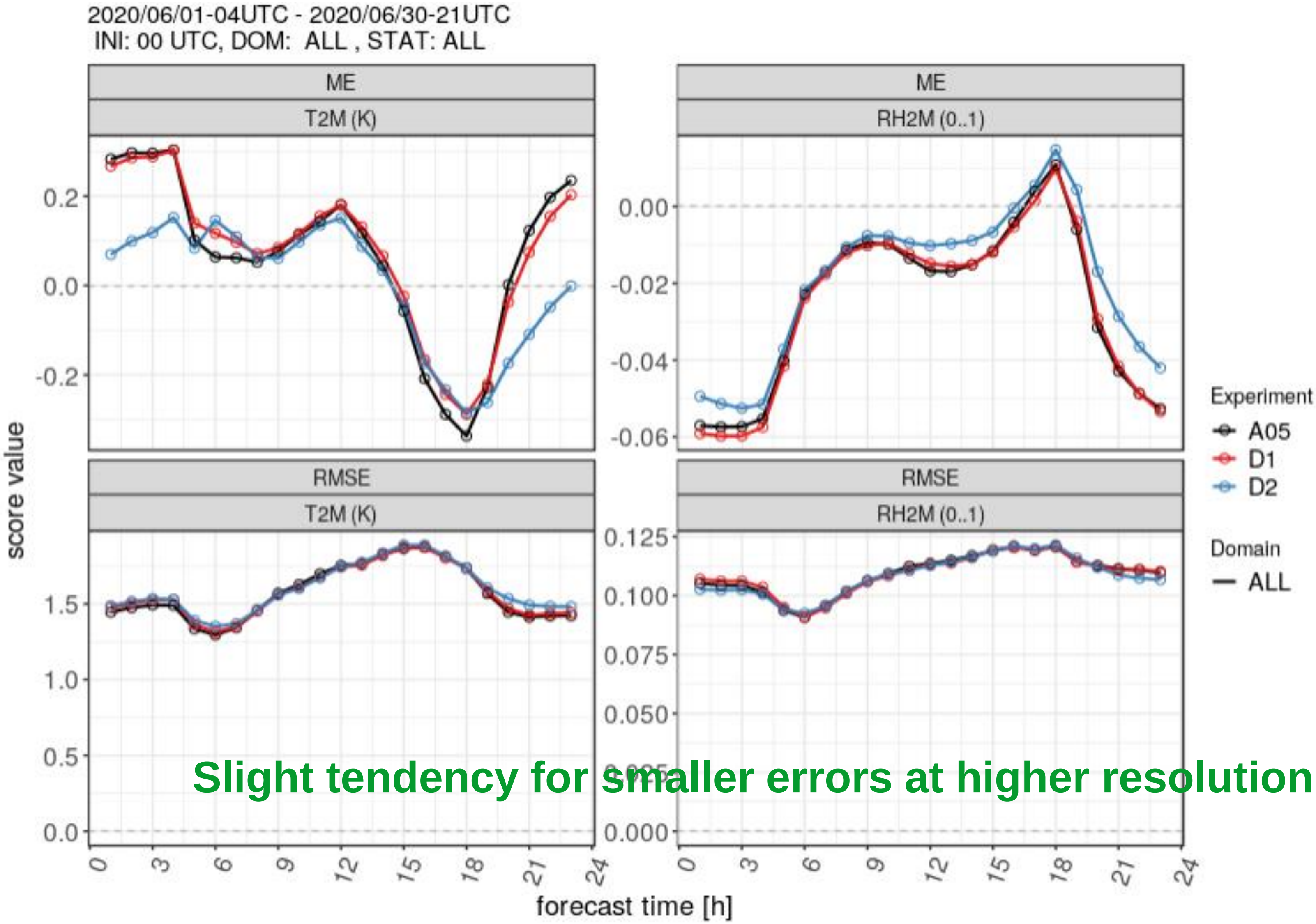
Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Too strong daytime boundary-layer mixing at higher resolutions

Resolution-dependence: 500 m, 1 km, 2 km
2 m temp./humidity, June 2020

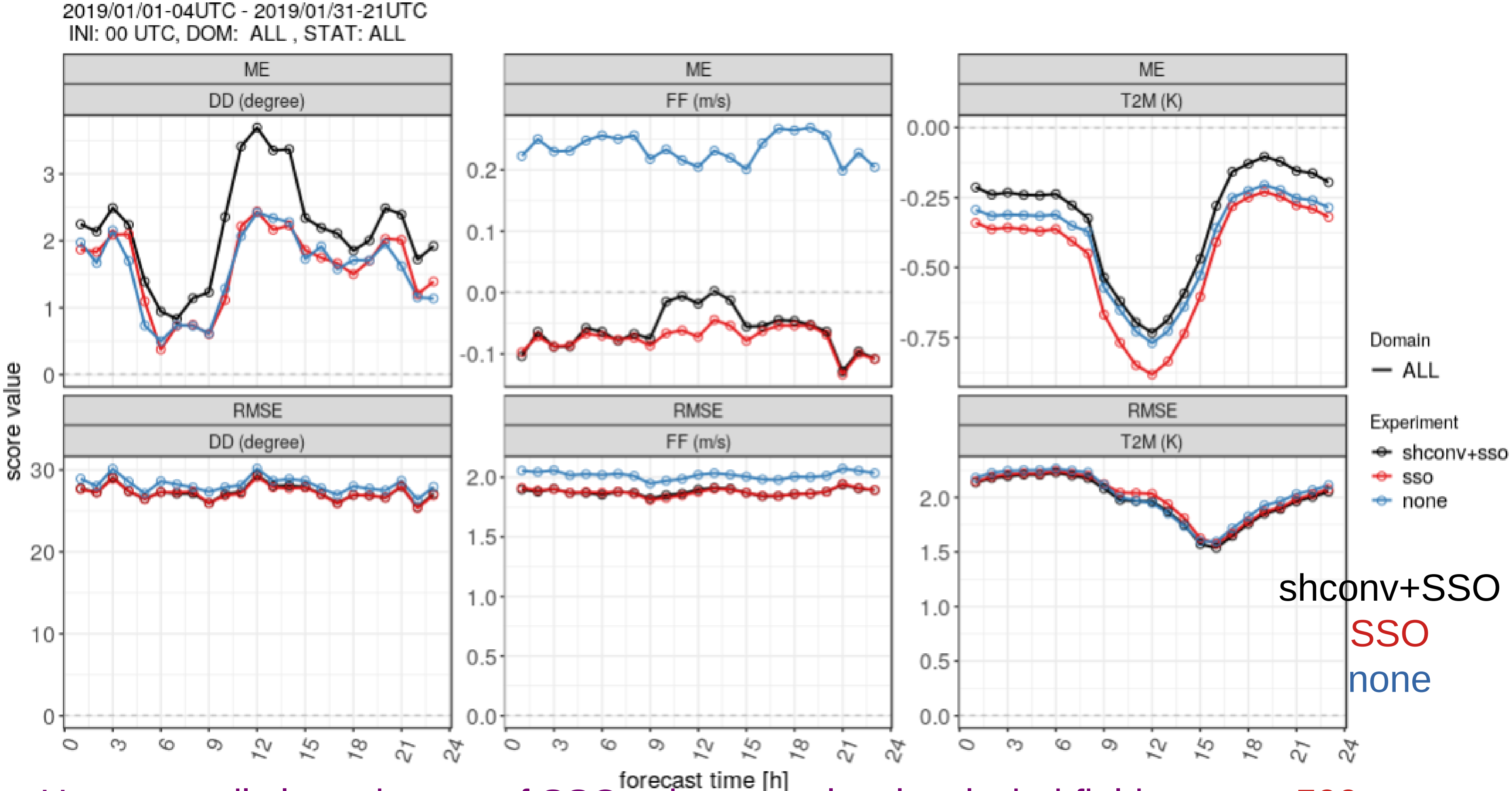
Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Slight tendency for smaller errors at higher resolution

Impact of SSO scheme and shallow convection; 500-m domain, January 2019

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

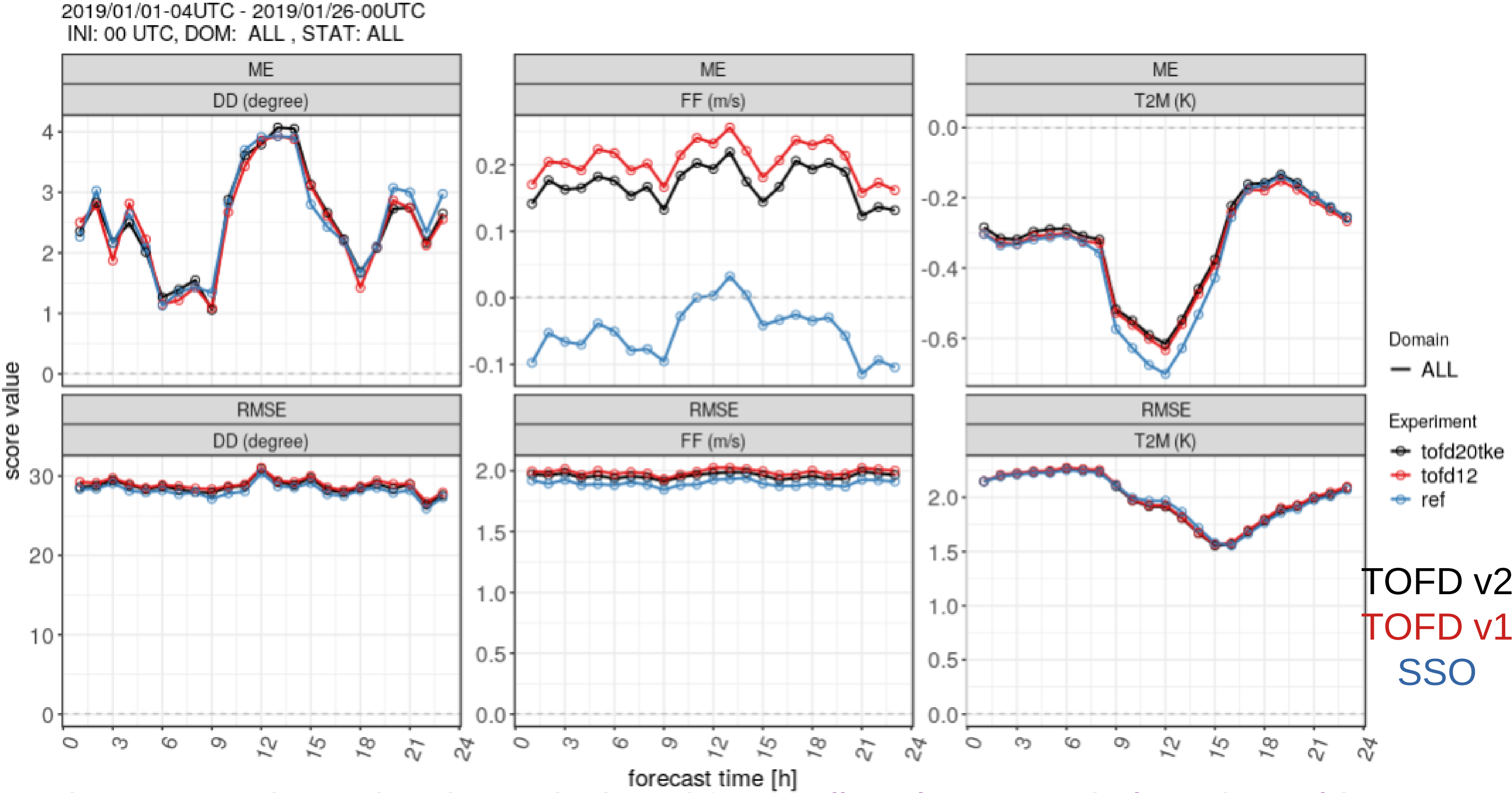


Unexpectedly large impact of SSO scheme on low-level wind field **even at 500 m**, minor positive impact of shallow convection on temperatures due to additional mixing

Note: the Lott-Miller SSO scheme has been developed for global models

TOFD vs. SSO scheme; 500-m domain, January 2019

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



The TOFD scheme has been designed for small scales; nevertheless, it would require significant further tuning to perform better than the SSO scheme!

Summary (LAM tests)

- **Under stable conditions, a direct benefit of enhanced resolution on the forecast quality is evident for 10-m wind speed/direction and to a lesser extent for 2-m temperature and humidity**
- **In convective boundary layers, total (resolved plus parameter-ized) mixing appears to increase with increasing resolution, which asks for further development work on the turbulence scheme; a resolution-dependent tuning of the shallow convection scheme has already been performed in order to counteract this effect**
- **Even at a mesh size of 500 m, parameterizing the drag contributed by subgrid-scale orography is necessary; from the schemes currently available in ICON, the SSO scheme delivers the best results**

Resolution-dependence of forecast quality: global experiments

- **Motivation; global convection-permitting modelling is pursued in various upcoming Exascale projects**
- **Their focus is mostly on technical aspects (GPU port, DSL, parallel I/O), tacitly assuming that existing models will deliver higher forecast quality just by operating them at convection-permitting resolution and turning off the deep-convection scheme**

- **The following results will demonstrate that major investments in parameterization development are needed as well**

Test strategy

- **ICON forecast runs for January 2021 (only the first 5 days for the time being) at R3B9 (3.25 km), with references at R3B8 (6.5 km) and R3B7N8 (operational configuration with 13 km globally and 6.5 km over Europe)**
- **120 vertical layers extending up to 75 km**
- **Initial conditions interpolated from IFS analyses for atmospheric fields, combined with interpolated surface fields from ICON analyses**
- **This is to avoid a possible advantage for the currently operational configuration, which otherwise would start from its ‘own’ analysis**
- **Results are shown for 6.5 km with deep convection scheme and for 3.25 km with and without deep convection scheme**

Evaluation metrics:

- **Standard verification against SYNOP and TEMP observations**

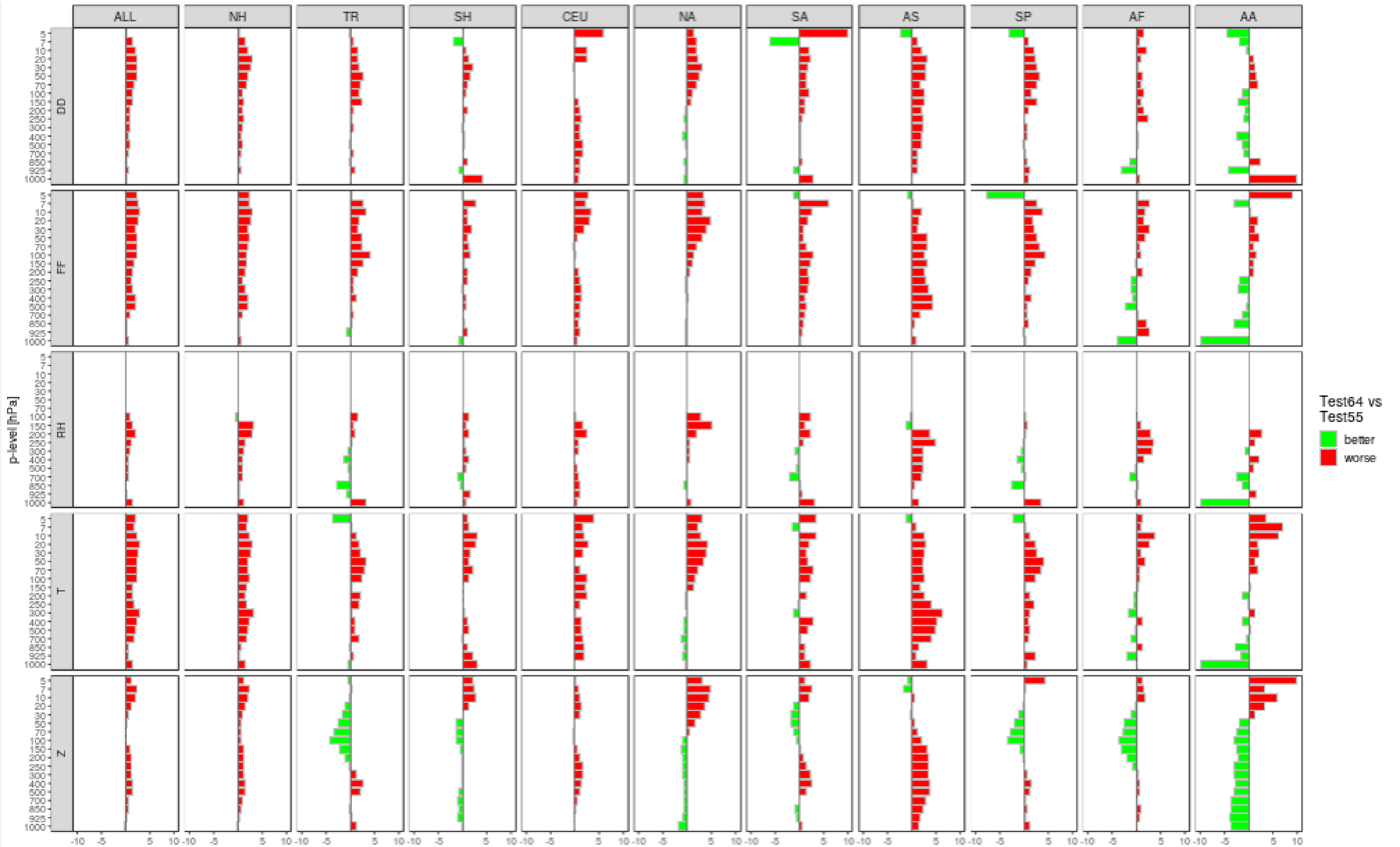
Some remarks

- **It is clear that 5 forecast runs are insufficient for statistical significance except for large changes; only the latter will be discussed in the following**
- **Orography data have a raw resolution of 30“ (~1 km), which is insufficient for a proper calculation of SSO parameters for a 3 km grid (and barely sufficient for 6.5 km)**

- **On DWD’s NEC SX Aurora, 45 nodes (with 64 cores each) are needed for R3B9L120 to fit into the memory; a 7.5-day forecast takes about 5 hours in this case**
- **A scaling test has not been conducted yet (would have to be attached to a maintenance downtime)**

Score card for verification against radiosondes, 13 km vs. 6.5 km (green: 6.5 km better)

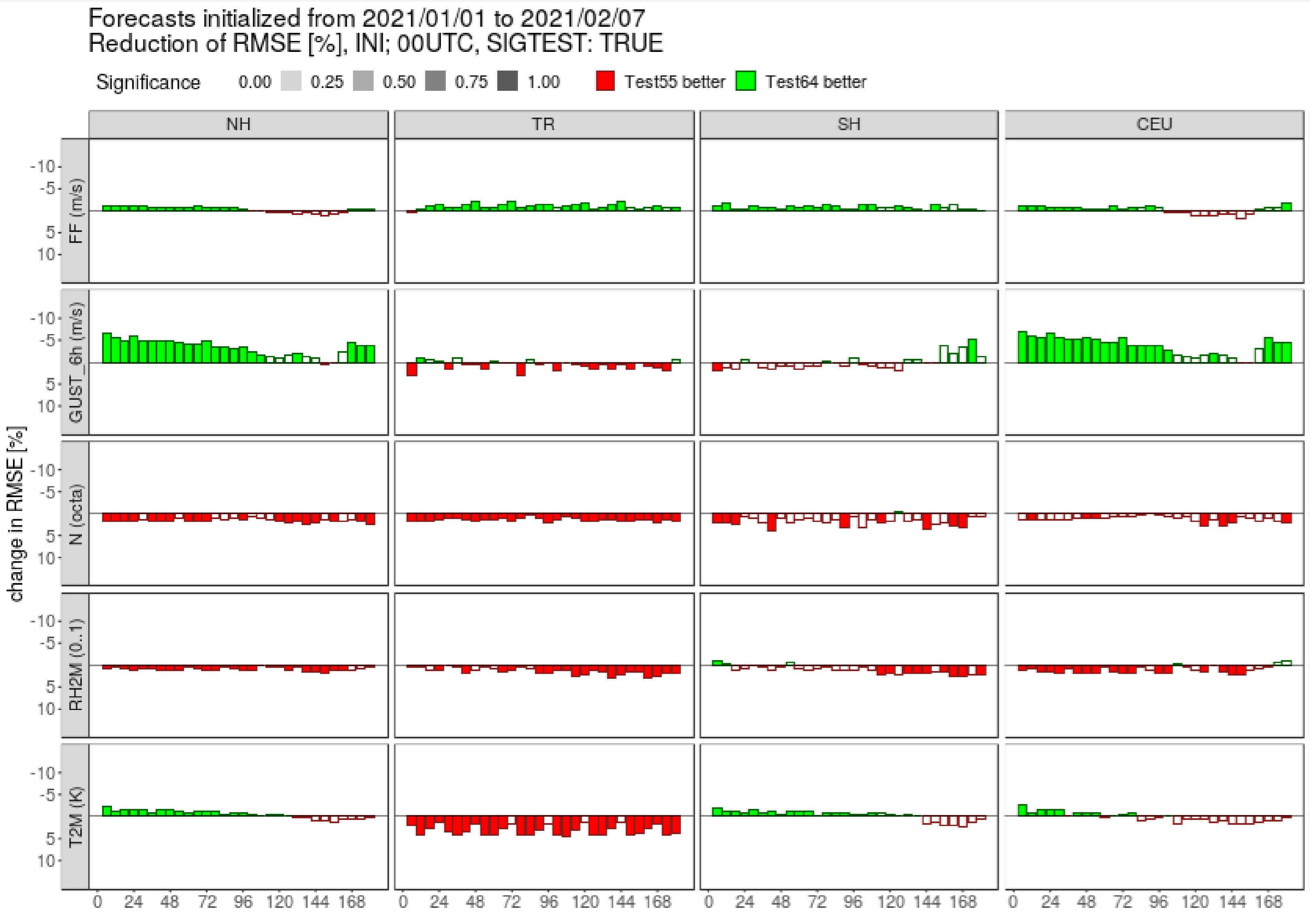
Deutscher Wetterdienst
Wetter und Klima aus einer Hand



quite disappointing results ...

Score card for verification against SYNOP data, 13 km vs. 6.5 km (green: 6.5 km better)

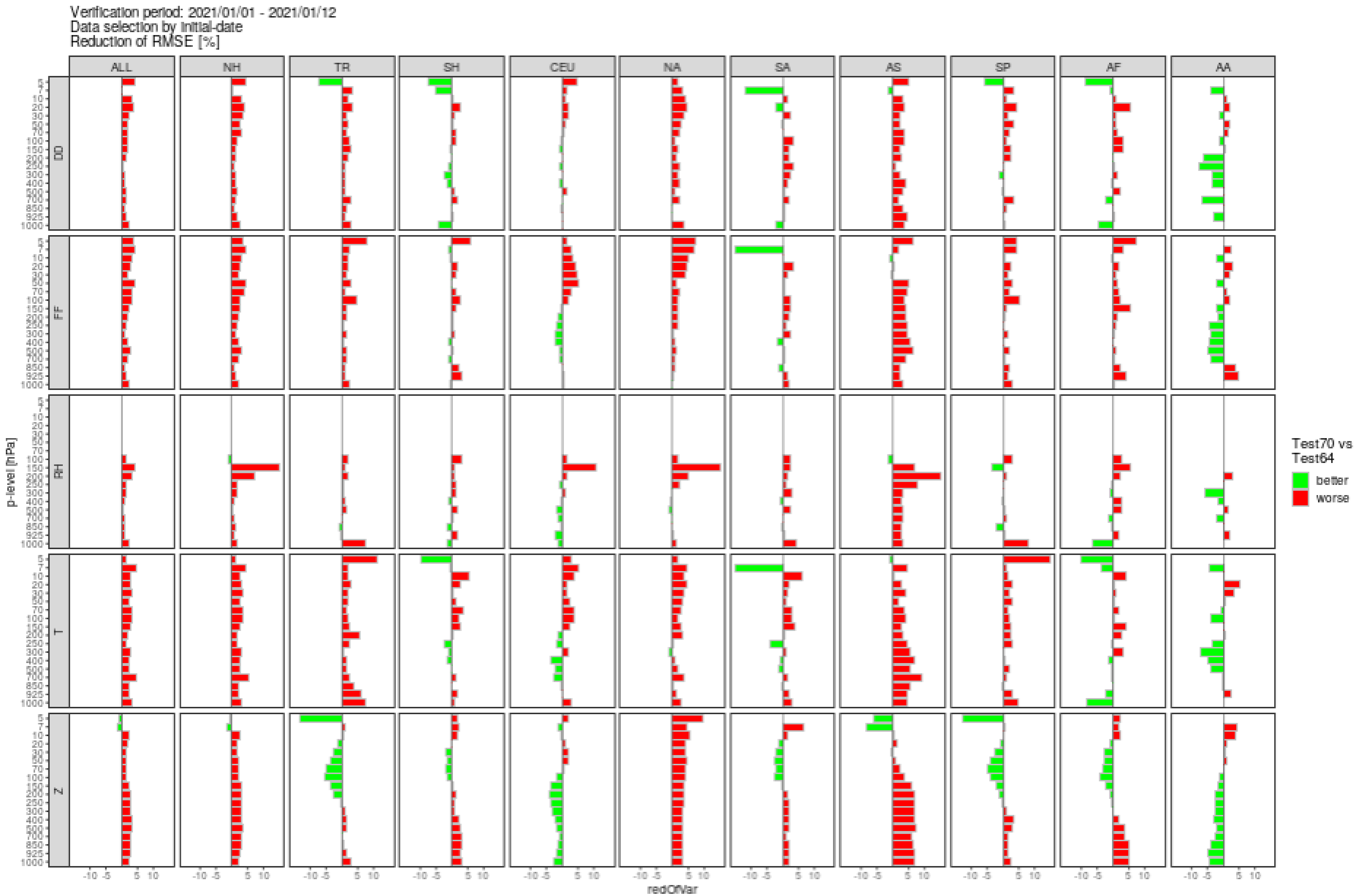
Deutscher Wetterdienst
Wetter und Klima aus einer Hand



… at least, some near-surface based quantities get better

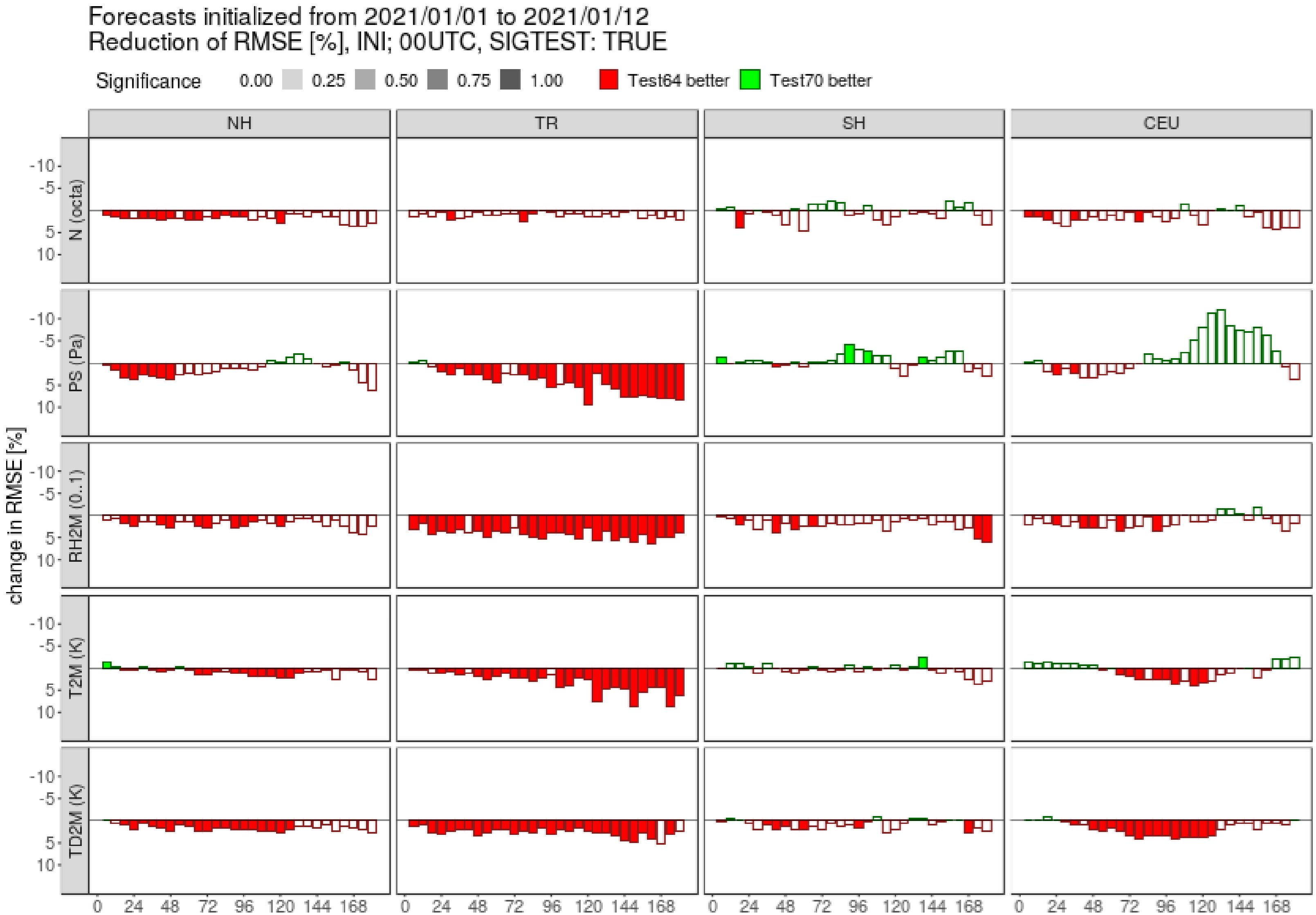
Score card for verification against radiosondes, 6.5 km vs. 3.25 km (green: 3.25 km better)

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Score card for verification against SYNOP data, 6.5 km vs. 3.25 km (green: 3.25 km better)

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

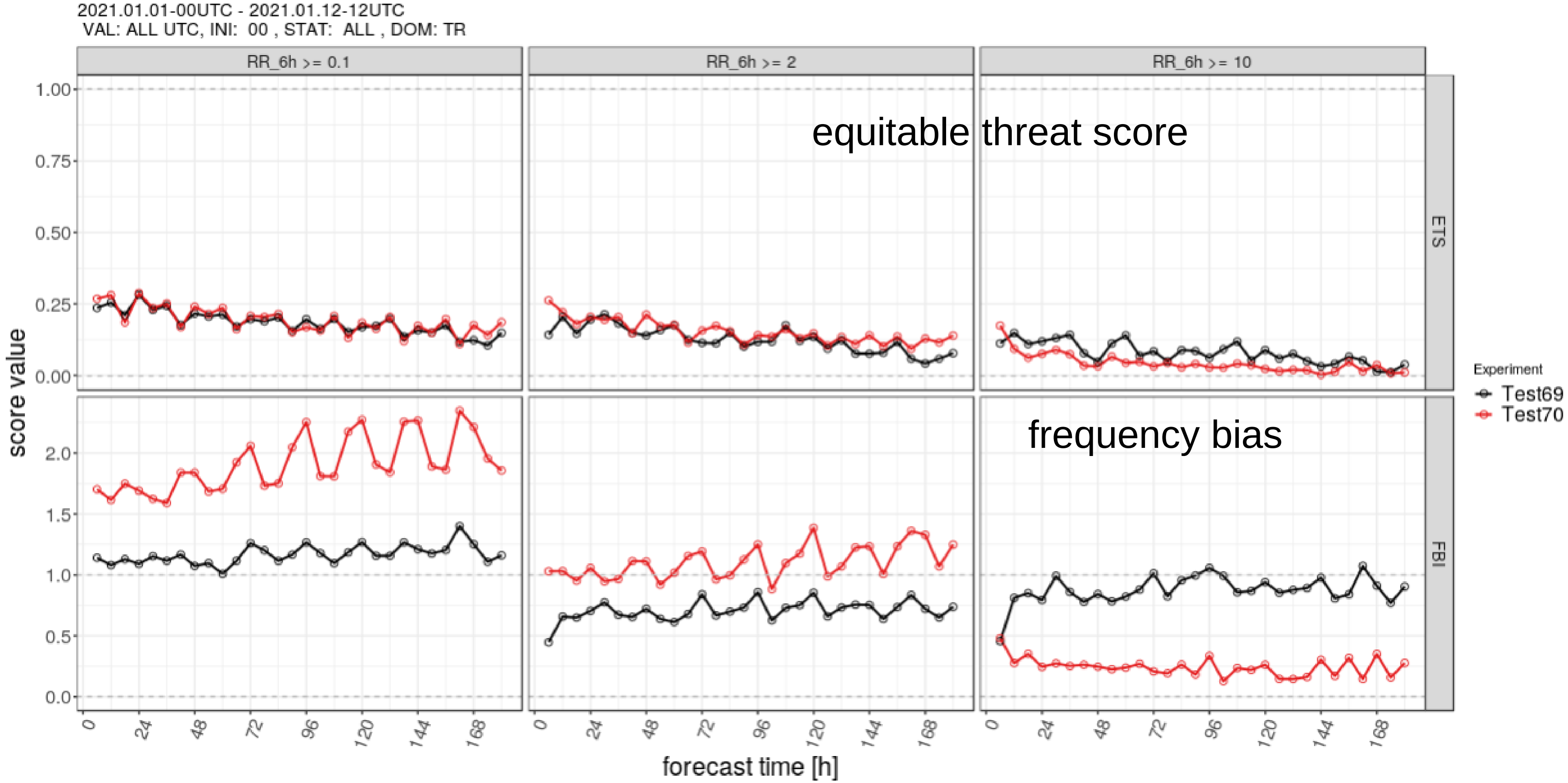


3.25 km performs even worse!

Turn off parameterization for deep convection

Precipitation verification, tropics
without / with deep convection scheme

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

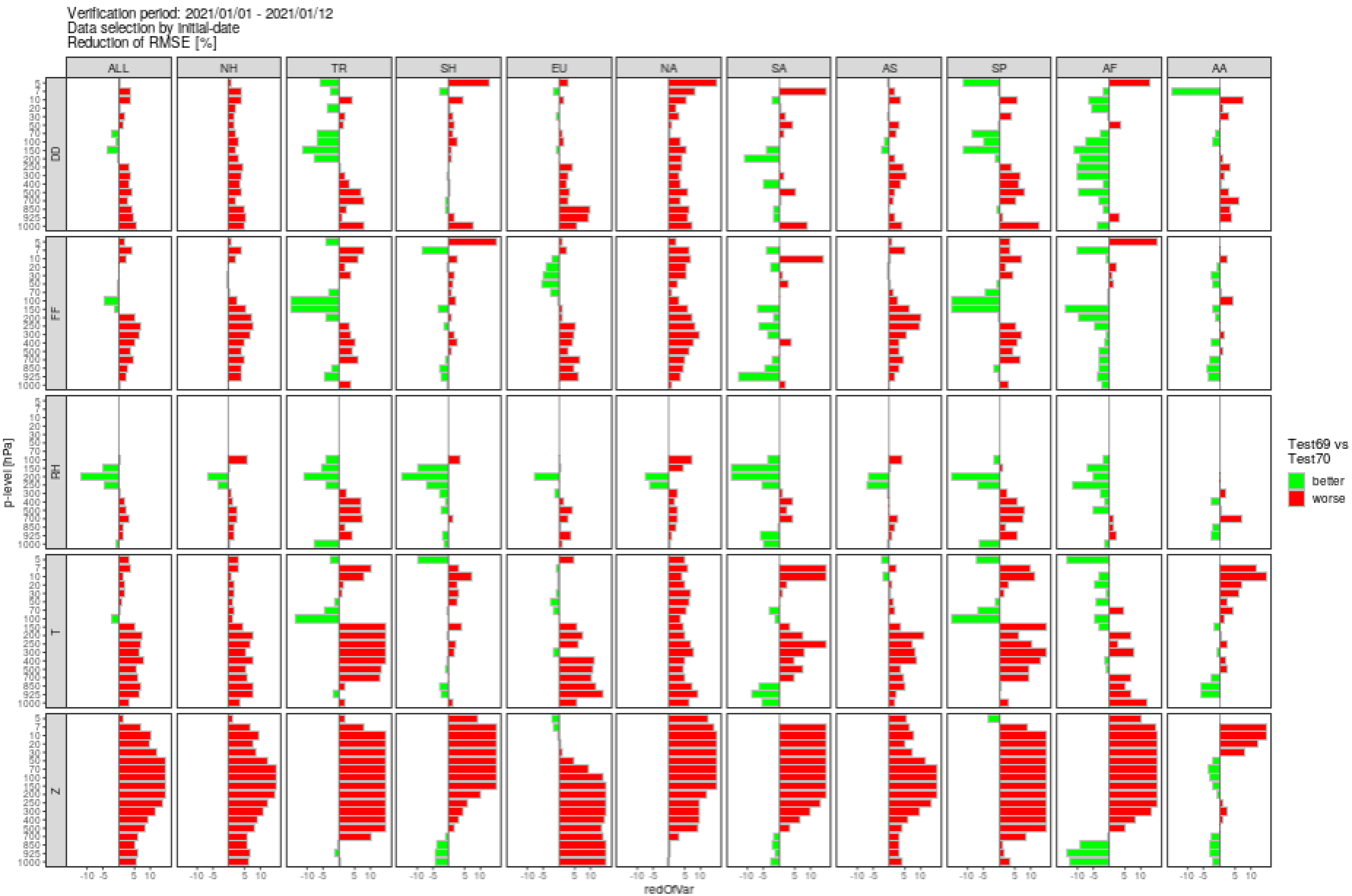


Much better representation of intensity spectrum

BUT...

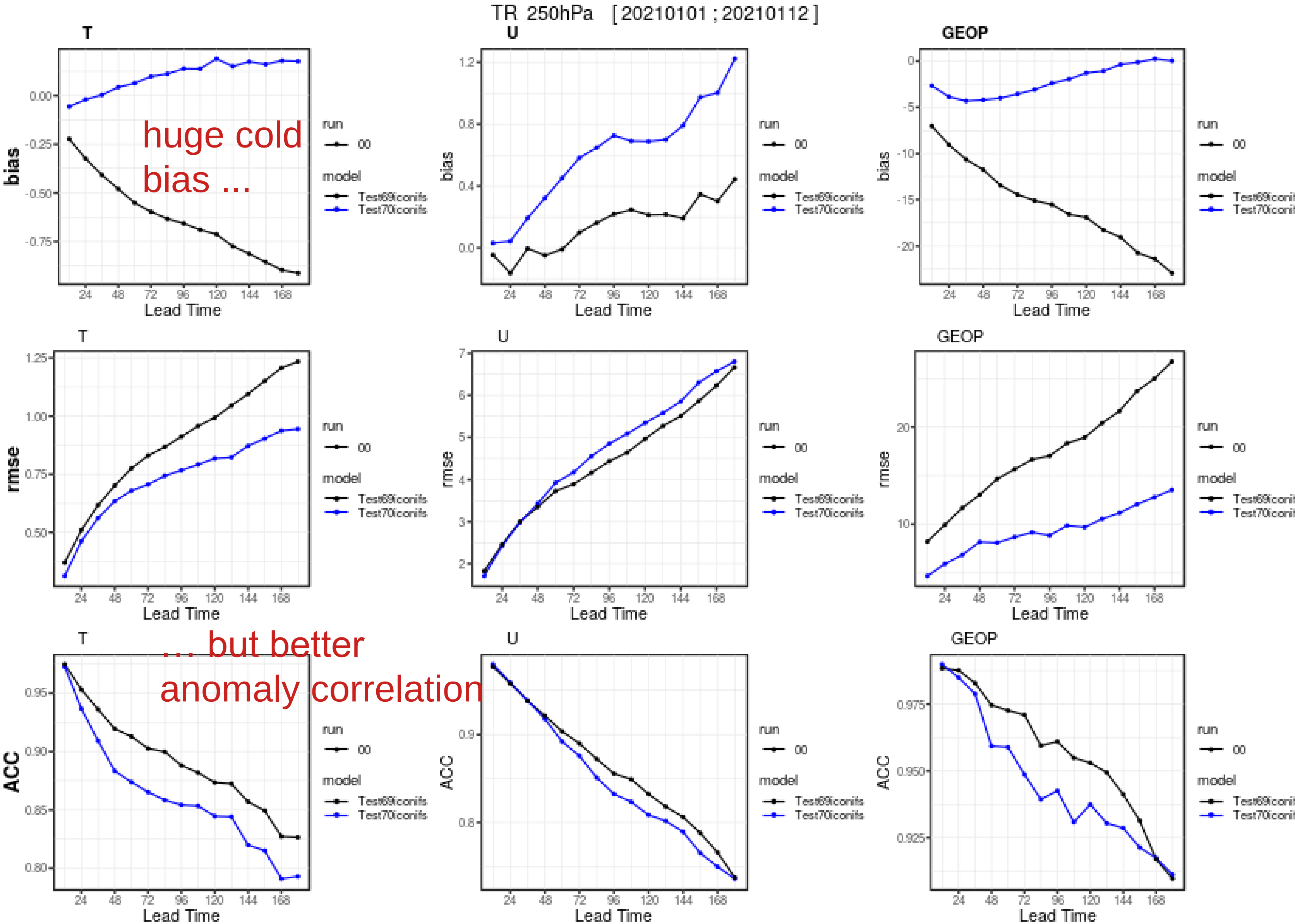
Score card for verification against radiosondes, 3.25 km (green: better without deep convection)

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Analysis verification, tropics, 250 hPa
without / with deep convection scheme

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Summary (global tests)

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Results are quite disappointing – the direct benefit from enhancing the horizontal resolution one gains at coarser scales appears to terminate around 10 km
- Forecast scores for the NH are probably degraded by the fact that the available orography data are too coarse for computing adequate SSO parameters, particularly for R3B9

- With explicitly simulated convection, some aspects that are well known to be notoriously misrepresented by parameterizations improve: intensity spectrum and diurnal cycle of convection, organization and propagation of mesoscale convective systems
- However, a huge cold bias arises that has its maximum in the upper tropical troposphere. It entails a planetary-scale redistribution of mass that compromises the pressure forecast over the whole globe
- **Convection-permitting is not convection-resolving!**

Conclusions

- **Enhancing the model resolution used to provide a direct benefit for forecast quality**
- **At least in the global ICON, this rule appears to be no longer valid in the convective gray zone (we don't know for LAM configurations – testing this in a meaningful way is not easy)**
- **Convection-permitting is not convection-resolving - and global model configurations are even more sensitive against bias issues than limited-area models**
- **Besides further development of our physics parameterizations for gray zone resolutions, more accurate and higher-resolved external parameter data are needed (work in progress, but their importance cannot be stressed too much)**
- **Leaving the convective gray zone means entering the gray zone of turbulence: the list of open issues will remain long!**