Technical Memo

ECMWF
European Centre for Medium-Range Weather Forecasts

868

Towards optimal parameters for the prediction of near surface temperature and dewpoint

Anton Beljaars

June 2020

Series: ECMWF technical memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/publications/

Contact: library@ecmwf.int

Abstract

Screen level temperature and dewpoint forecasts from ECMWF's Integrated Forecasting System (IFS) show systematic errors with large scale geographical patterns, seasonal variation and pronounced diurnal cycles. The errors are hard to address because of the multitude of processes involved. To be able to optimize parameters it is important to have a fast and efficient testing environment. This report explores the potential benefit of so-called "relaxation" integrations, where the upper air model fields are gently relaxed towards the analyses. It has the advantage that a full annual cycle, following a realistic synoptic evolution, can be reproduced rather efficiently. Three aspects are discussed: (i) evaluation of the relaxation integrations, (ii) verification of daily maximum/minimum temperature and corresponding dewpoint at screen level, and (iii) sensitivity experiments in an attempt to find more optimum parameter settings in the atmosphere/land coupling.

It is concluded that relaxation experiments are useful, although it is necessary to account for differences with respect to operations. Evaluation indicates that about half of the root mean square errors in temperature and dewpoint are systematic errors (in the sense that these error are present in the local monthly means). The systematic errors have complex large scale geographical, seasonal, and diurnal patterns. Some of the errors can be related to orography, snow and bare soil, and some to land-atmosphere coupling. Parameter sensitivity experiments show that improvement is possible on some aspects, but it is very hard to find combinations of parameters that do not lead to deterioration of other aspects. The combination of a daytime dry and cold bias in summer is of particular concern, because soil moisture data assimilation relies heavily on the assumption that the model is unbiased.

1 Introduction

Direct model output of weather parameters is becoming increasingly important in applications of Numerical Weather Prediction (NWP). However, routine verification of 2m temperature shows non-negligible errors depending on area, time of the day, and meteorological circumstances. From experience with model changes and from sensitivity studies, it is well known that errors in 2m temperature forecasts are affected by uncertainties in a range of processes, e.g. clouds, turbulent diffusion, atmosphere to surface coupling, vegetation, orographic effects, presence of snow, and heat diffusion in soil/snow. A purely bottom up approach where model components are built from process knowledge and put together in a model is unlikely to be optimal, because of the high level of empiricism in the land surface coupling and the large uncertainty in many, often geographically dependent parameters. Earlier studies have shown that there are multiple options for optimization. An example is the study by Viterbo et al. (1999), who found that both the processes of soil moisture freezing and atmospheric turbulent diffusion could be used to mitigate a winter cold bias over Europe. Dutra et al. (2010) showed that more realistic snow density leads to more heat insulation and colder more realistic temperatures over Siberia. Repetition of old sensitivity experiments shows that this latest version of ECMWF's Integrated Forecasting System (IFS) is much more sensitive to changes in turbulent diffusion than the model version before the snow change (Beljaars 2012, Holtslag et al. 2013). More recently, a study by Betts and Beljaars (2017) shows that temperatures in the ERA-Interim re-analysis (ERA-I) are substantially biased over the Canadian Prairies, with a maximum bias in clear sky conditions. They suggest that the coupling of the atmosphere to the underlying soil is too strong. The soil heat fluxes in ERA-I are about a factor two too large at diurnal as well as seasonal time scales.

The purpose of this report is to work towards a procedure to optimize parameters in the IFS and to recommend priorities for improvement. To optimize parameters over a full annual cycle, it is crucial to have a fast and efficient experimental environment. Therefore "relaxation" of a long integration towards the analysis is explored. The advantage is that the synoptic variability of a particular year can be reproduced

Table 1: List of experiments

Short name	Description	Forecast type	Resolution	Exp ID
OPER	Operations	Daily short range	Tl1279	0001
CONTR	Control	Daily short range	Tco319	gmdt
OPENL	Control open loop surface	Daily short range	Tco319	gml8
RELAX	Relaxation	Single long run	Tco319	gmdv

in a single long run. The time scale of the relaxation is a compromise. It is selected to be sufficiently strong to ensure that the simulation follows the synoptic variability of the analysis, but weak enough to be able to see the effect of changes in model parameters.

Before considering parameter sensitivity, the relaxation procedure is evaluated. To be able to home in on a limited number of parameters, emphasis is further put on maximum and minimum temperature, because it can be considered at a global scale independent of time-zone, and on stratification of errors according to land cover and land characteristics.

2 Model configurations

The relaxation integrations are performed with model version CY43R1 at the resolution of Tco319L137 (equivalent to about 35 km in gridpoint space) for the full year of 2015. The relaxation uses ERA-Interim vorticity and temperature in spectral space truncated at T63. The vorticity relaxation is applied with a time scale of 12 hours on all vertical levels. Temperature relaxation is applied from level 33 (about 15 hPa) to level 82 (about 280 hPa) with a time scale of 120 hours. The purpose of the temperature relaxation is to prevent drift around the tropopause. Note that there is no relaxation of the divergent flow, moisture, clouds, and temperature in middle to lower troposphere. These variables are mostly controlled by the model itself and follow the synoptic variability imposed by the rotational part of the flow.

We compare the relaxation integrations with the operational high resolution deterministic forecasts (OPER) for 2015 (until 12 May IFS CY40R1; after that CY41R1) and with the operational analysis. The OPER forecasts and the operational analyses were at spectral resolution Tl1279 (approximately 16 km in grid point space). For verification, the 2m temperature and dewpoint fields of the operational analysis are used. This is based on the experience that the 2m analysis draws closely to the observations and is therefore a good reference for verification over land. Only the four synoptic times 0, 6, 12, and 18 UTC are available from the analysis system.

Daily short range forecasts with CY43R1 at Tco319L137 resolution are also performed to have a more clean assessment of the relaxation. Unfortunately, there is no technical framework yet to apply relaxation to soil moisture and temperature, and therefore also short range ”open loop” experiments are included. In these experiments the 24-hour forecast fields of the land surface variables of the previous day are used to initialize the forecast. This is equivalent to a scenario where the surface analysis is bypassed and no increments are applied to soil moisture, snow temperature, and snow depth. A summary of the relevant experiments is listed in Table 1. The short names for the different configurations are: OPER, CONTR, OPENL and RELAX.

3 General verification

Figure 1 shows the monthly averages of root mean square (RMS) error of geopotential height (Z500) in Northern and Southern hemisphere at the day-1 and day-2 forecast range from the OPER system and the Tco319 short range forecasts. They are both initialized from the same operational analyses at 0 UTC. As expected, the Tco319 forecasts are slightly worse than the operational forecasts because of the lower resolution. Figure 2 compares with the relaxation integration and indicates that its accuracy is equivalent to forecasts at about day 1.5.

Similarly, wind errors in the Tropics are shown in figs. 3 and 4. Again they indicate that the day-1 and day-2 errors of the low resolution forecasts are slightly worse than the high resolution ones. Comparison with the relaxation integration indicates that at 850 hPa the errors are equivalent to a day-2 forecast and at 200 hPa to about 1.8 days. Vorticity relaxation is obviously less efficient in the Tropics than in the extra-Tropics because in the Tropics the divergent part of the flow is less controlled by the rotational part.

It is worth noting that the large scale error measures of Z500 and wind with the short range open loop experiment (gml8) are nearly indistinguishable from the control short range forecasts (gmdt) at the same resolution (not shown).

Atmospheric moisture, surface processes, divergent motion and physical processes are only indirectly constrained by the relaxation of vorticity. Precipitation is considered here because it is an important variable for land surface interaction. As an example, Fig. 5 shows the 0-24 hr forecast of precipitation on 20150701 from OPER (Tl1279), from the 0-24 hr forecast at Tco319 initialized from the operational analysis, and from the Tco319 relaxation integration starting at 20150101. The short range Tco319 reproduces the synoptic features of OPER very well, both in the extra-tropics and in the Tropics. Also the relaxation integrations show the same synoptic features, which indicates that the relaxation is very effective, but the details are reproduced less well than in the 24 hr short range forecasts. Real precipitation verification is hard, and assuming that the short range OPER is the best, the latter is used here as a reference. Mean precipitation over NH, SH, TR, and RMS differences with respect to OPER are shown in Fig. 6 for July 2015. They indicate that the discrepancy between relaxation precipitation and OPER precipitation is slightly worse than in day-2 of the Tco319 forecasts. Precipitation is a field with lots of small scale variability, so it is understandable that differences grow more quickly e.g. compared to Z500 which has much less variability at small horizontal scales.

4 Errors of 2m temperature and dewpoint

Errors of temperature and dewpoint at screen level are highly variable and show seasonal, diurnal and geographical patterns. An attempt is made here to reduce the number of degrees of freedom by focusing on minimum and maximum temperature. This has the advantage that global statistical measures can be used whereas verification at particular times in UTC can only be interpreted at a regional scale. Maximum and minimum temperatures are not analysed and therefore the 6-hourly analysis is used that is closest to the maximum or minimum temperature. The particular synoptic times are diagnosed from the monthly mean analyses of temperature, resulting in a global time-zone mask for the maximum and minimum temperature. The masks for maximum and minimum temperature are shown in Fig. 7. They are used to compute global statistics of 2m temperature and dewpoint errors, corresponding to the maximum or minimum analysed temperature of a particular month.

From operational verification it is well known that 2T and 2D errors are only weakly dependent on fore-

cast range. Systematic biases in particular, are already present at very short lead times. Regarding the daily short range forecasts, emphasis is therefore put on lead times of 6, 12, 18 and 24 hours. Examples of monthly mean temperature errors are shown in figs. 8 and 9 for the months of January and July respectively. The left column of the figures contains from top to bottom: mean bias of maximum temperature (Mx-2T) of OPER, mean bias of Mx-2T of the control forecasts at Tco319, mean bias of minimum temperature (Mn-2T) of OPER, and mean bias of Mn-2T of the control forecasts at Tco319. The right hand columns contain the corresponding RMS errors. It can be concluded that: (i) the high resolution forecasts have smaller errors than the low resolution forecasts reflected in mean as well as RMS errors, (ii) the error patterns are different for day, night, summer and winter, illustrating the complex nature of the errors, and (iii) a substantial component of the errors is systematic and orography related. It is important to note that orographic corrections have been made on the basis of differences in orography between model and analysis. Another interesting point is that the resolution effects appear to be more pronounced at night and in winter, which points in the direction of stable boundary layer diffusion. In Appendix B it is hypothesized that meso-scale variability is a key player in the resolution dependence over orography.

Similar figures are shown for the dewpoint errors Mx-2D and Mn-2D in figs. 10 and 11. Variables Mx-2D and Mn-2D are not the maximum and minimum dewpoint but the dewpoint at the synoptic time where 2T has its maximum or minimum. The dewpoint errors show some similarity with the temperature errors, but not always.

To summarize the results, Fig. 12 shows the biases (left column) and RMS errors (right column) as monthly time series of averages over NH land. The left column contains Mx-2T, Mn-2T, Mx-2D, Mn-2D. Looking at the OPER Mx-2T, Mn-2T biases (top left panel, solid red and blue curves), we see that the minimum temperature is hardly biased and that the maximum temperature is too low by about 0.5 K, which results in an underestimation of the diurnal cycle of 0.5 K. Spring is an exception with a cold bias in both Mx-2T, Mn-2T of about 0.4 K. The rows represent the 4 configurations OPER, CONTR, OPENL and RELAX. As expected, the biases as well as the RMS errors increase from top to bottom. Interestingly, in the Tco319, the global average of Mn-2T develops a cold bias of a few tenth of degrees, improving the amplitude of the diurnal cycle.

The right hand column of Fig. 12 has an additional set of thin curves where the RMS is computed of the monthly means rather than from the individual days. Time averaging removes the day to day variability of the errors. This is a way of quantifying the systematic errors as it captures the RMS of the geographical patterns. Typically the RMS of the monthly mean is about half of the total RMS, so the systematic component of the 2T and 2D errors is substantial. All four configurations show similar seasonal behaviour although the magnitude of the errors increases from OPER to CONTR, OPENL and RELAX. Most of the increase is systematic.

To find out whether these errors are related to land use types, the following dominant types are defined for the Northern Hemisphere (20^o to 90^o N): NHHV (High Vegetation without snow), NHLV (Low Vegetation without snow), NHBS (Bare Soil without snow), NHSH (Snow underneath High vegetation), NHSL (Snow on Low vegetation or bare soil), and NHSO (the area where the standard deviation of Sub-grid Orography is larger than 100 m). The first five are exclusive and their combined area should be equal to the land area of the Northern Hemisphere (NHLA). The land types are plotted for January and July 2015 (based on monthly mean analysed snow depth) in the top two panels of Fig. 13. The third panel outlines the area where the standard deviation of the sub-grid orography is larger than 100 m. The bottom panel of Fig. 13 shows the evolution of land fractions relative to the total land area of the Northern Hemisphere (NHLA).

The RMS of the monthly mean Mx-M2T and Mn-M2T errors is displayed in Fig. 14 for the different land use types. The M2T errors characterise the magnitude of bias as geographical average. Comparison with the NHLA errors (i.e. all areas combined) provides candidates of land types for which parameters may not be optimal. The red curve represents NHLA, so we look for curves that are above the red one. For 2T, all results point to areas with exposed snow (NHSL) and large sub-grid orography (NHSO) as candidates for improvement. The results for 2D in Fig. 15 point to bare soil as a candidate for improvement during day time. At night, 2D behaves like temperature and the NHSO and NHSL types show the largest errors. This is to be expected, because at night the air is often saturated which makes dew point equal to temperature.

Since snow occurs more frequently at high elevation, it is possible that the large errors seen over snow (NHSL and NHSH) are related to orography. For that purpose sub-areas are defined where the surface elevation is below 1000 m, with snow under high vegetation (NHSHL) and and exposed snow (NHSLL). They are represented in Figs. 14 and 15 by thin lines. Results in summer are erratic due to the small snow area and are excluded from the discussion. For the winter maximum temperature, we see that there are no big differences between snow below and above 1000 m. However, the minimum temperature errors are reduced substantially by excluding the high elevation areas (e.g. Greenland).

Conclusions from the different types of experiments (OPER, Low resolution, Open loop or relaxation) about the role of processes and land cover, are similar although the magnitude of the errors is different. This suggests that relaxation integrations are sufficient to study the impact of processes and the role of land cover.

5 Water cycle and the impact of data assimilation

The water cycle varies strongly from region to region, so a number of areas are selected, for which the surface water budget is considered together with the short range temperature and dewpoint errors. The areas are Northern Hemisphere land (NHEM), Western Europe (WEUR), and Iberia (IBER). Figures 16 to 18 describe the seasonal evolution of the water fluxes (panel a), soil plus snow water content (b), mean and RMS of 2T errors (c,d) and mean and RMS of 2D (errors). The water cycle figures are for the Tco319 control experiment (CONTR) and the open loop experiment (OPENL); the 2T/2D errors are also shown for OPER and RELAX. In the RMS error plots, the additional thin curves indicate the systematic part of the error. Typically the systematic part of the error is about 30 to 50% of the total RMS error.

The flux curves (panels a) show the usual seasonal evolution of fluxes with precipitation throughout the year at all the selected locations, maximum evaporation in summer, a drying of the soil in spring/summer and a small residual as runoff. Because the soil moisture analysis system adds or deletes water/snow from the land total column water (soil moisture + snow water equivalent), the 24-hour forecasts of land total column water is not the same as in the verifying analysis. The increments have a systematic component of typically 0.5 to 1 mm/day, which is a non-negligible part of the water budget. The openloop experiment does not have increments by construction. The RMS of soil moisture increments was also considered and was typically of the order of precipitation. This is perhaps not surprising, because precipitation is often intermittent, noisy and not always located at the correct place. Ideally, soil moisture analysis responds to short range precipitation errors.

The error ranking of the four experiments is typically OPER, CONTR, OPENL and RELAX, where OPER has the smallest errors and RELAX the largest. This also applies to the systematic part of the errors. The only difference between CONTR and OPENL is that OPENL takes the land surface initial conditions from the 24-hour forecast of the previous day, so it can be interpreted as a an experiment

without land data assimilation and where the land surface evolves through the forcing of the short range forecasts. The result that the 2T and 2D-errors are systematically smaller in CONTR indicates that the land data assimilation has a positive effect. This is not unexpected, because the land data assimilation uses the 2T and 2D SYNOP observations as data source, and is designed to reduce the errors in these parameters by adding or removing soil moisture (even if they are not due to soil moisture errors). However, there is not always a clear relation between soil water increments (light blue curve in panels a) and the systematic errors on Mx-2T and Mx-2D (panels c and e). Apparently, the Jacobians that translate 2T and 2D errors into soil water and soil temperature increments are rather complex. The IBER region is an interesting example. Most of the time, soil moisture is added by the data assimilation because of the dry bias. However, the month of May is an exception. In May soil moisture is removed, resulting in an even higher dry bias in 2D. The temperature error is reduced, so it suggests that the soil moisture increments mainly act to correct 2T bias. However, this leads to amplification of the dry bias. More discussion on the interaction between model biases and land data assimilation is given in Appendix A.

6 Sensitivity experiments

To explore the sensitivity of 2T and 2D errors to model parameters, a number of sensitivity experiments were performed with the relaxation configuration. Parameters were selected that affect thermal coupling between atmosphere, surface and sub-surface. The experiments are the following (see also Table 2, located between Figs. 18 and 19):

Z0Hd10 The roughness length for heat is divided by a factor 10 for all LOW vegetation types. The effect is that the aerodynamic conductivity between the lowest model level and the surface is reduced. Because it is applied to the surface boundary condition, the reduced conductivity is below the 2m level.

LSTd2 The skin layer conductivity in STable conditions is divided by 2 for all land use types. For all vegetation types the conductivity is reduced from 10 to 5 $Wm^{-2}K^{-1}$. For bare soil the conductivity of half the top soil layer is reduced from 15 to 7.5 $Wm^{-2}K^{-1}$. In this case stable means that the skin is warmer than the underlying soil or snow surface, which is predominantly the case during the day. The skin layer conductivity controls the thermal coupling between the skin and the underlying surface.

LUSd2 Similar to LSTd2, the skin layer conductivity in UnStable conditions is divided by 2 for all land use types.

LSNd2 The conductivity of the top half of exposed snow layers is reduced by a factor 2, i.e. from 7 to 3.5 $Wm^{-2}K^{-1}$.

RSMm2 The minimum stomatal resistance of low vegetation is increased by a factor 2, i.e from 100 to 200 sm^{-1} for crops, irrigated crops, short grass, and tall grass, and from 80 to 160 sm^{-1} for tundra.

NOAz0 An earlier model change was reverted. The earlier model change (introduced in May 2015 CY41R1; RD memo 16-042, 2016) consisted of the aggregation of roughness length of momentum before defining the surface layer heat transfer for each tile. The purpose was to reduce the fast temperature drop at sunset over Scandinavia in spring.

LHVd2 The skin layer conductivity of the vegetation classes ”Evergreen needle leaf trees”, ”Deciduous needle leaf trees”, ”Evergreen broad leaf trees”, and ”Deciduous broad leaf trees”, is reduced from 10 to 5 $Wm^{-2}K^{-1}$.

gd065 A canopy resistance stress function is applied to low vegetation dependent on atmospheric moisture deficit. The same function is applied as for tall vegetation i.e. $f_3(D_a) = exp(g_D D_a)$, where D_a is atmospheric moisture deficit (hPa), but g_D is set to 0.065 hPa^{-1}. This value has been selected to mimic the moisture deficit function proposed by Beljaars and Bosveld (1997) as optimal value for the Cabauw location. The formulation they used was $f_3(dq) = \{1 + a_q(dq - dq_r)\}^{-1}$, with $dq_r = 3\,g\,kg^{-1}$ and $a_q = 0.16\,kg\,g^{-1}$. The change is applied to the vegetation classes "Crops, Mixed farming", "Sort grass", "tall grass", "Desert", "Tundra", "Irrigated crops", "Semideserts", "Evergreen shrubs", and "Deciduous shrubs". It should be noted that Beljaars and Bosveld (1997) used this stress function in combination with a radiation function that is different in the ECMWF model. The Cabauw site experiences only rarely soil moisture stress.

Fig. 19 contains summary statistics of all relaxation experiments for Northern Hemisphere land and Iberia covering the full year of 2015. Bias (red dots for Mx, blue dots for Mn), RMS of the monthly mean error (red stars for Mx, blue stars for Mn), and bias of the amplitude of the diurnal cycle (black dots) are shown for all experiments. IBERIA has been selected as a sub-area because the errors are more pronounced. Figs. 20 and 21 show the same for winter (DJF) and summer season (JJA) respectively. The experiments are represented by symbols from left to right. The control relaxation experiments is marked with a vertical line.

The 2T and 2D output of all relaxation experiments has been bias corrected by adding the monthly mean difference between OPER and RELAX. In this way the systematic errors of RELAX look as those of OPER. This follows the idea that changes in the biases of the relaxation configuration also apply to the operational configuration. In other words it is assumed that small model optimizations do not affect the differences between e.g. high resolution operational forecasts and low resolution relaxation experiments. We are looking now for experiments that reduce the biases with respect to the operational errors. In OPER, Mx-2T has a general NH cold bias of about 0.5 K, a very small cold bias of Mn-2T and an underestimation of the amplitude of the diurnal cycle by about 0.4 K. The main contribution to these 12-month mean biases appear to come from the non-winter months. Iberia shows similar signals.

The experiments that modify the biases in a non-negligible way are: Z0Hd10 which slightly decreases the amplitude of the diurnal cycle, LSTd2/LUSd2 which increase the amplitude of the diurnal cycle and RSMm2/gd065 which reduce the day time cold bias. The only change that improves the day time temperature without being negative on the night time is with RSMm2 and gd065. The impact comes predominantly from summer where the day time temperatures are sensitive to evaporation. Looking at dewpoint, it is clear that the day time improvement by RSMm2/gd065 is offset by a deterioration in dewpoint. This is understandable, because a cold bias combined with a dry bias can not be explained by errors in the Bowen ratio only.

The combined day time cold/dry bias is also problematic for soil moisture assimilation, because the conflicting atmospheric temperature and humidity biases can not be reduced at the same time with soil moisture increments. So addressing these moisture biases in the IFS is an important issue. For the moment it is hard to imagine a change in the land surface coupling that reduces both temperature and humidity errors. It is very well possible or perhaps even more likely that the problem is not in the land coupling but in interaction of the boundary layer with the troposphere above. Too strong boundary layer top entrainment, too active shallow convection or too little precipitation evaporation will all lead to a dry bias in the boundary later.

In contrast to the general dry bias, there is evidence that evaporation is too high in spring. This can be reduced by increasing the canopy resistance (as demonstrated by the sensitivity experiments). This will be an improvement for spring (in both maximum temperature and dewpoint), but will result in an even

larger dry bias in summer. A sensible way forward to narrow down the uncertainty in canopy resistance is to review the literature and to do a comparison study with flux tower stations.

7 Discussion and conclusions

The potential of so-called relaxation integrations has been explored. The purpose is to have a testing environment where a full annual cycle can be run in "deterministic" mode through relaxation to the 6-hourly analyses. This turns out to work very well. At Tco319, a full year runs in about 24 hours, so it allows for fast experimentation with many parameter settings. The NH Z500 RMS errors are similar to those of 1.5 day forecasts, so the relaxation runs reproduce the real synoptic variability rather closely. The 200 and 850 tropical wind errors are equivalent to day-2 forecasts. Compared to the operational 24-hour forecasts, precipitation appears to be slightly worse than the day-2 forecasts. Verification of screen level temperature and moisture shows similar errors patterns as the short range forecasts, but the biases are larger. To be useful as sensitivity experiments, the relaxation runs have to be bias corrected, a posteriori, with respect to the higher resolution operational forecasts.

A verification strategy has been developed using operational screen level temperature and humidity analyses as reference. Proxies of maximum and minimum temperature are used for evaluation by selecting the closest location-dependent 6-hourly time slot. This allows the evaluation of daily maxima and minima, independent of time-zone. Furthermore, errors are stratified according to land use. About half of the NH RMS errors of temperature and dewpoint are systematic, where the latter is defined as the RMS of the point-wise monthly average. The errors have large scale geographical patterns, and evolve seasonally with pronounced differences between day and night. Some of the errors are related to areas with bare soil, orography and snow. Recent work has demonstrated that substantial improvement can be achieved with a multilayer snow scheme (Arduini et al. 2019, Dutra et al. 2010), and with an updated vegetation cover data that improves the bare soil issues (Johannsen et al. 2019, Nogueira et al. 2020). The night time orography signals are dependent on model resolution, which is believed to be due to the increase of resolved shear with resolution, having an impact on shear driven turbulence and therefore on night time mixing (see Appendix B).

Results suggest that evaporation is too strong in spring and that the model has a tendency to get too dry in summer, particularly over bare soil. However, the day time dry bias in summer is in apparent conflict with a cold bias. Such an error characteristic can not be explained by soil moisture. Soil moisture data assimilation does reduce 2T and 2D errors but is obviously struggling because of the conflicting model biases in temperature and humidity (see Appendix A). In Appendix A it is also shown that May 2015 reduction of observation errors of 2T and 2D in the land data assimilation was beneficial to the forecasts of 2T and 2D. The improved diurnal cycle of convection slightly reduced the 2D errors (see Appendix A).

The combined cold/dry bias in the IFS is unlikely to be due to the land surface scheme only. Also the atmospheric model is probably a player. To analyse further, the vertical structure of temperature and humidity profiles using radio sondes has been studied (see Appendix C). It shows that at low winds in Southern European summer, moisture profiles are too well mixed in the IFS. This may be due to too strong turbulent diffusion, but also the effects of terrain heterogeneity in the observations could be an issue. SYNOP stations tend to be located in valleys where the soil is more moist than at higher elevations.

A series of sensitivity experiments have been performed that address the strength of atmosphere/surface coupling, evaporation and the interpolation between the lowest model level and the surface. Experiments show that improvement is possible on some aspects, but that often deterioration occurs on other aspects.

The amplitude of the diurnal cycle can be improved by changing the roughness length for heat. Increase of the canopy resistance results in lower evaporation and a day time temperature increase in spring and summer. However it also leads to an increase in the dry bias. To choose adequate canopy resistances it would be better to do a systematic study based on FLUXNET tower data.

As overall conclusion it is reasonable to say that relaxation is a powerful tool to do fast sensitivity experiments which give an impression of the impact of a model changes at all seasons. In practice, a combination of changes will be necessary, but it turns out to be very hard to find combinations of parameters that improve on some aspects and do not deteriorate on others.

Acknowledgement

The author would like to thank Emanuel Dutra for debugging the open loop forecast configuration and Tim Stockdale for providing the relaxation configuration. Gianpaolo Balsamo, Irina Sandu, and Nils Wedi made suggestions for improvement of the manuscript, which is highly appreciated. This study benefited from discussions with many colleagues in particular Gabriele Arduini, Souhail Boussetta, Johnatan Day, Thomas Haiden, Patricia de Rosnay, and Polly Schmederer.

References

Arduini, G., G. Balsamo, E. Dutra, J. J. Day, I. Sandu, S. Boussetta, and T. Haiden 2019. Impact of a multi-layer snow scheme on near-surface weather forecasts. *Journal of Advances in Modeling Earth Systems*.

Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts 2009. A revised hydrology for the ecmwf model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *Journal of hydrometeorology*, 10(3):623–643.

Beljaars, A. 1994. The impact of some aspects of the boundary layer scheme in the ecmwf model. In *ECMWF Seminar on Parametrization of Sub-grid Scale Physical Processes*, Pp. 125–161. https://www.ecmwf.int/sites/default/files/elibrary/1994/8035-impact-some-aspects-boundary-layer-scheme-ecmwf-model.pdf.

Beljaars, A. 2012. The stable boundary layer in the ecmwf model. In *ECMWF Workshop on diurnal cycles and the stable boundary layer*, Pp. 1–10. http://www.ecmwf.int/sites/default/files/elibrary/2012/8029-stable-boundary-layer-ecmwf-model.pdf.

Beljaars, A., A. R. Brown, and N. Wood 2004. A new parametrization of turbulent orographic form drag. *Quarterly Journal of the Royal Meteorological Society*, 130(599):1327–1347.

Beljaars, A. C. and F. C. Bosveld 1997. Cabauw data for the validation of land surface parameterization schemes. *Journal of climate*, 10(6):1172–1193.

Betts, A. K. and A. Beljaars 2017. Analysis of biases in era-interim over the canadian prairies. *Submitted to JAMES*.

Boussetta, S., G. Balsamo, A. Beljaars, A.-A. Panareda, J.-C. Calvet, C. Jacobs, B. van den Hurk, P. Viterbo, S. Lafont, E. Dutra, et al. 2013. Natural land carbon dioxide exchanges in the ecmwf integrated forecasting system: Implementation and offline validation. *Journal of Geophysical Research: Atmospheres*, 118(12):5923–5946.

De Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen 2013. A simplified extended kalman filter for the global operational soil moisture analysis at ecmwf. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1199–1213.

Drusch, M. and P. Viterbo 2007. Assimilation of screen-level variables in ecmwf's 2019 integrated forecast system: A study on the impact on the forecast quality and analyzed soil moisture. *Monthly Weather Review*, 135(2):300–314.

Dutra, E., G. Balsamo, P. Viterbo, P. M. Miranda, A. Beljaars, C. Schär, and K. Elder 2010. An improved snow scheme for the ecmwf land surface model: description and offline validation. *Journal of Hydrometeorology*, 11(4):899–916.

Fairbairn, D., P. de Rosnay, and P. A. Browne 2019. The new stand-alone surface analysis at ecmwf: Implications for land–atmosphere da coupling. *Journal of Hydrometeorology*, 20(10):2023–2042.

Holtslag, A., G. Svensson, P. Baas, S. Basu, B. Beare, A. Beljaars, F. Bosveld, J. Cuxart, J. Lindvall, G. Steeneveld, et al. 2013. Stable atmospheric boundary layers and diurnal cycles: challenges for weather and climate models. *Bulletin of the American Meteorological Society*, 94(11):1691–1706.

Johannsen, F., S. Ermida, J. Martins, I. F. Trigo, M. Nogueira, and E. Dutra 2019. Cold bias of era5 summertime daily maximum land surface temperature over iberian peninsula. *Remote Sensing*, 11(21):2570.

Lott, F. and M. J. Miller 1997. A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, 123(537):101–127.

Martens, B., D. L. Schumacher, H. Wouters, J. Mũnoz Sabater, N. E. C. Verhoest, and D. G. Miralles 2020. Evaluating the surface energy partitioning in era5. *Geosci. Model Dev. Discuss.*, Pp. https://doi.org/10.5194/gmd–2019–315. in review.

Nogueira, M., C. Albergel, S. Boussetta, F. Johannsen, I. F. Trigo, S. L. Ermida, J. Martins, and E. Dutra 2020. Role of vegetation in representing land surface temperature in the chtessel (cy45r1) and surfexisba (v8.1) land surface models: a case study over iberia. *Geoscientific model development discussions*, Pp. https://doi.org/10.5194/gmd–2020–49.

Viterbo, P., A. Beljaars, J.-F. Mahfouf, and J. Teixeira 1999. The representation of soil moisture freezing and its impact on the stable boundary layer. *Quarterly Journal of the Royal Meteorological Society*, 125(559):2401–2426.

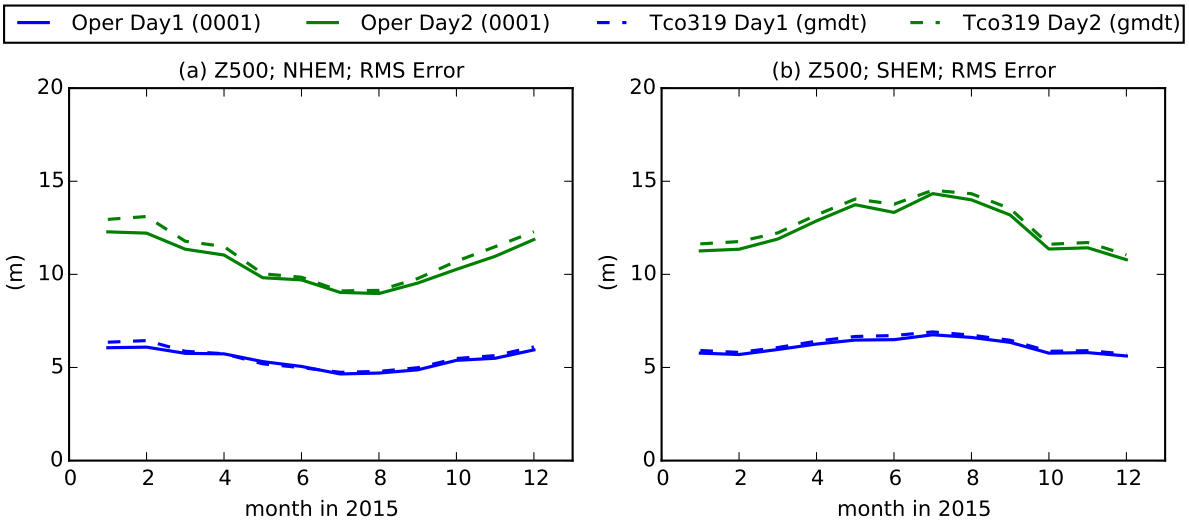

Figure 1: Monthly mean RMS errors of 500 hPa height over the Northern Hemisphere (left) and the Southern Hemisphere (right). The solid curves correspond to daily operational (OPER) day 1 (blue) and day 2 (green) forecasts initialized at 0 UTC. The dashed curves corresponds to the short range forecasts at Tco319 initialized from the operational analysis.

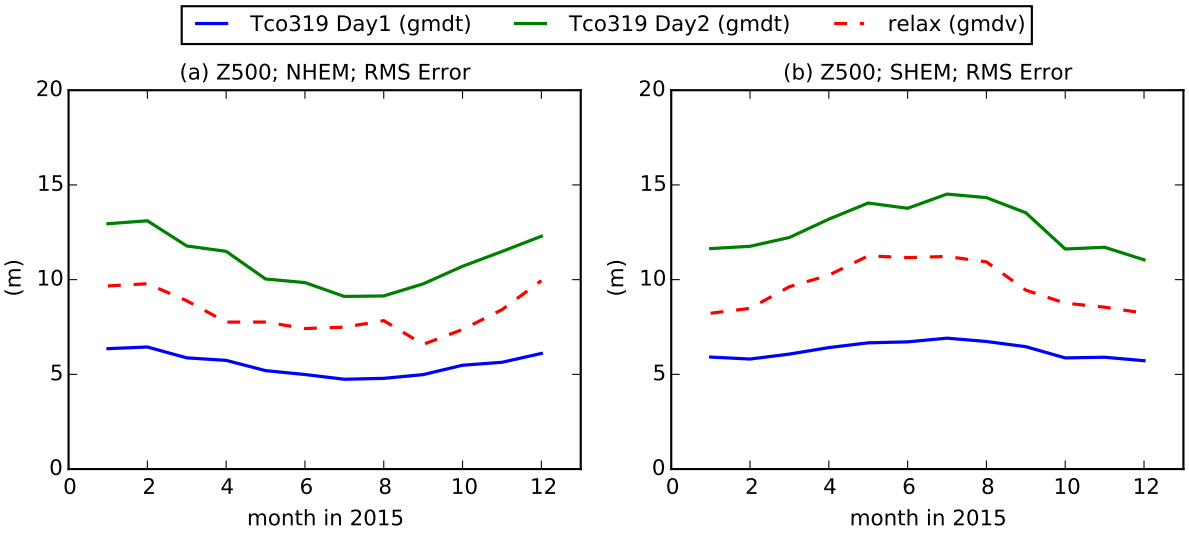

Figure 2: Monthly RMS errors of 500 hPa height over the Northern Hemisphere (left) and the Southern Hemisphere (right). The solid curves correspond to daily Tco319 forecasts from the operational analysis at 0 UTC at day 1 (blue) and day 2 (green). The red dashed curve corresponds to a single long integration at Tco319 which is relaxed to the ERA-I re-analysis.

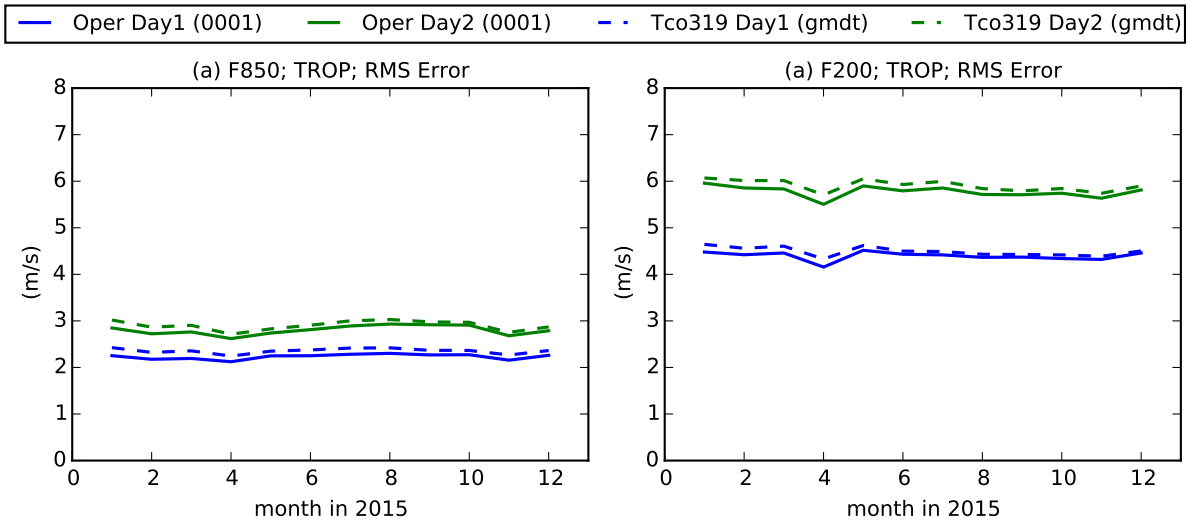

Figure 3: Monthly RMS errors of the 850 hPa wind (left) and the 200 hPa wind (right) in the Tropics. The solid curves correspond to daily operational (OPER) day 1 (blue) and day 2 (green) forecasts initialized at 0 UTC. The dashed curves corresponds to the short range forecasts at Tco319 initialized from the operational analysis.

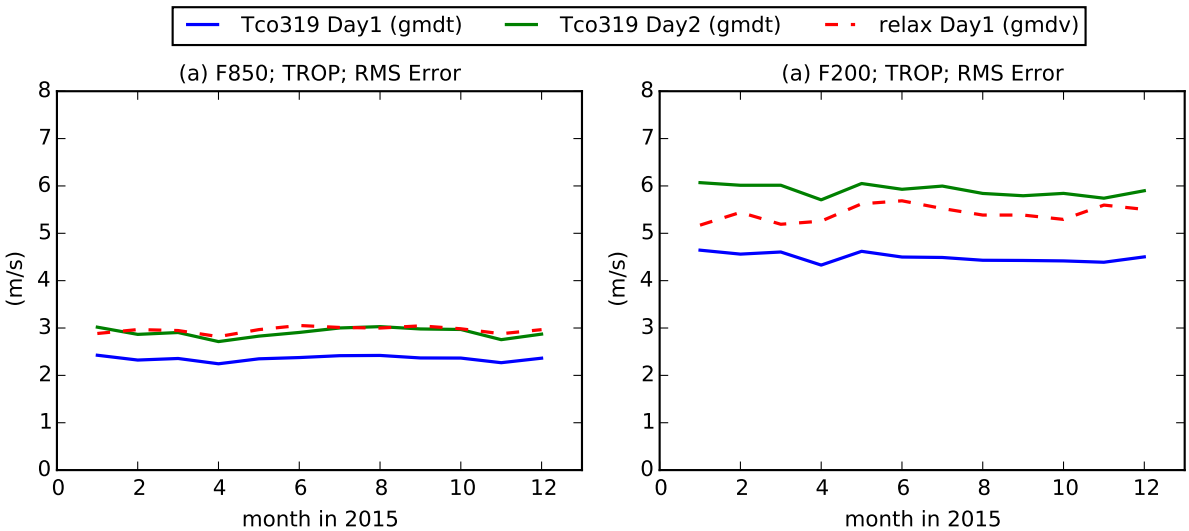

Figure 4: Monthly RMS errors of the 850 hPa wind (left) and the 200 hPa wind (right) in the Tropics. The solid curves correspond to daily Tco319 forecasts initialized from the operational analysis at 0 UTC at day 1 (blue) and Day 2 (green). The red dashed curve corresponds to a single long integration at Tco319 which is relaxed to the ERA-I re-analysis.

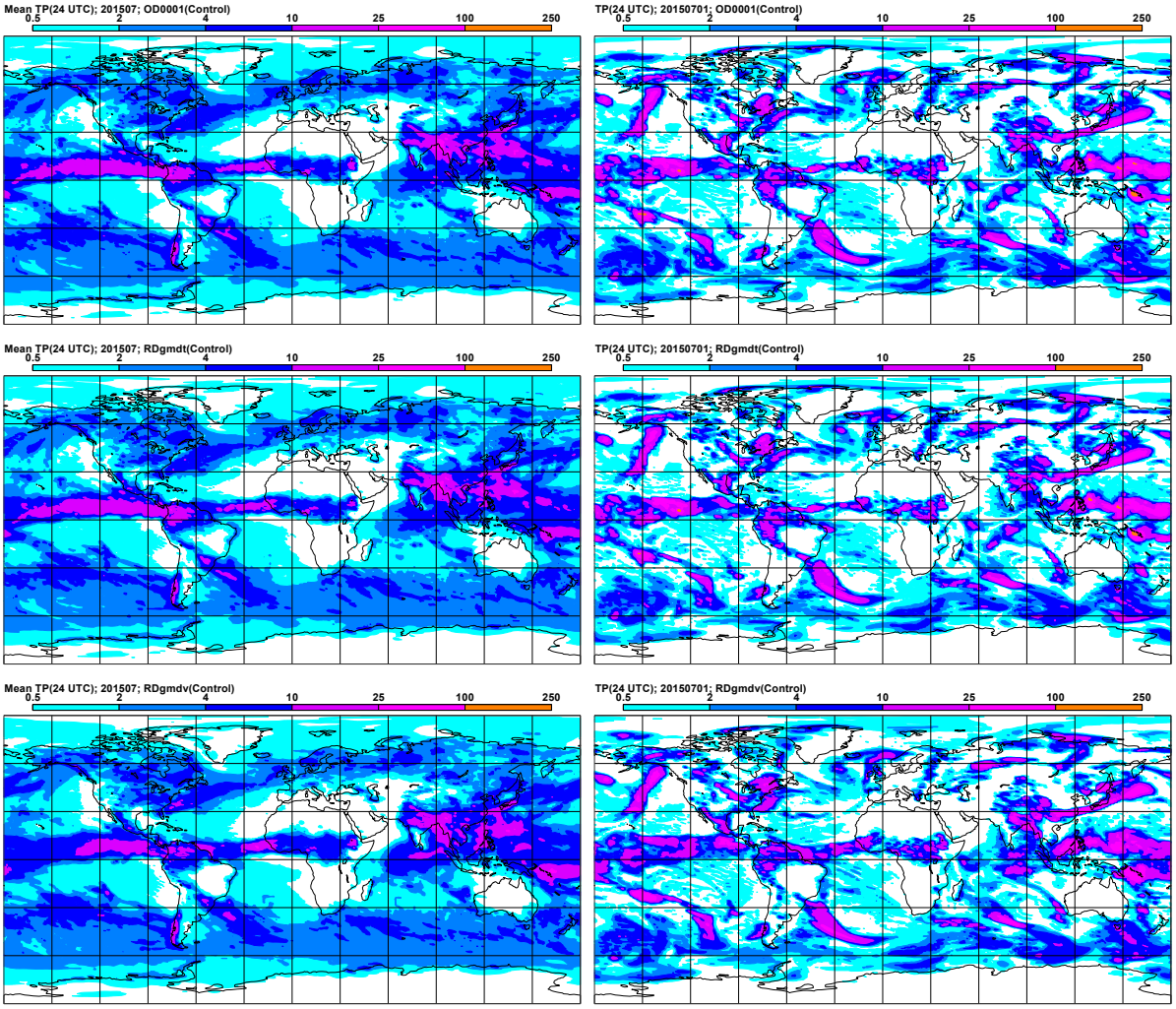

Figure 5: Mean precipitation (mm/day) for July 2015 (left) and the 24-hour accumulated value for the 1st of July (right). The top panels represent operational 24 hr forecasts, the middle panels the control short range forecasts at Tco319 and the bottom panels the long relaxation integration.

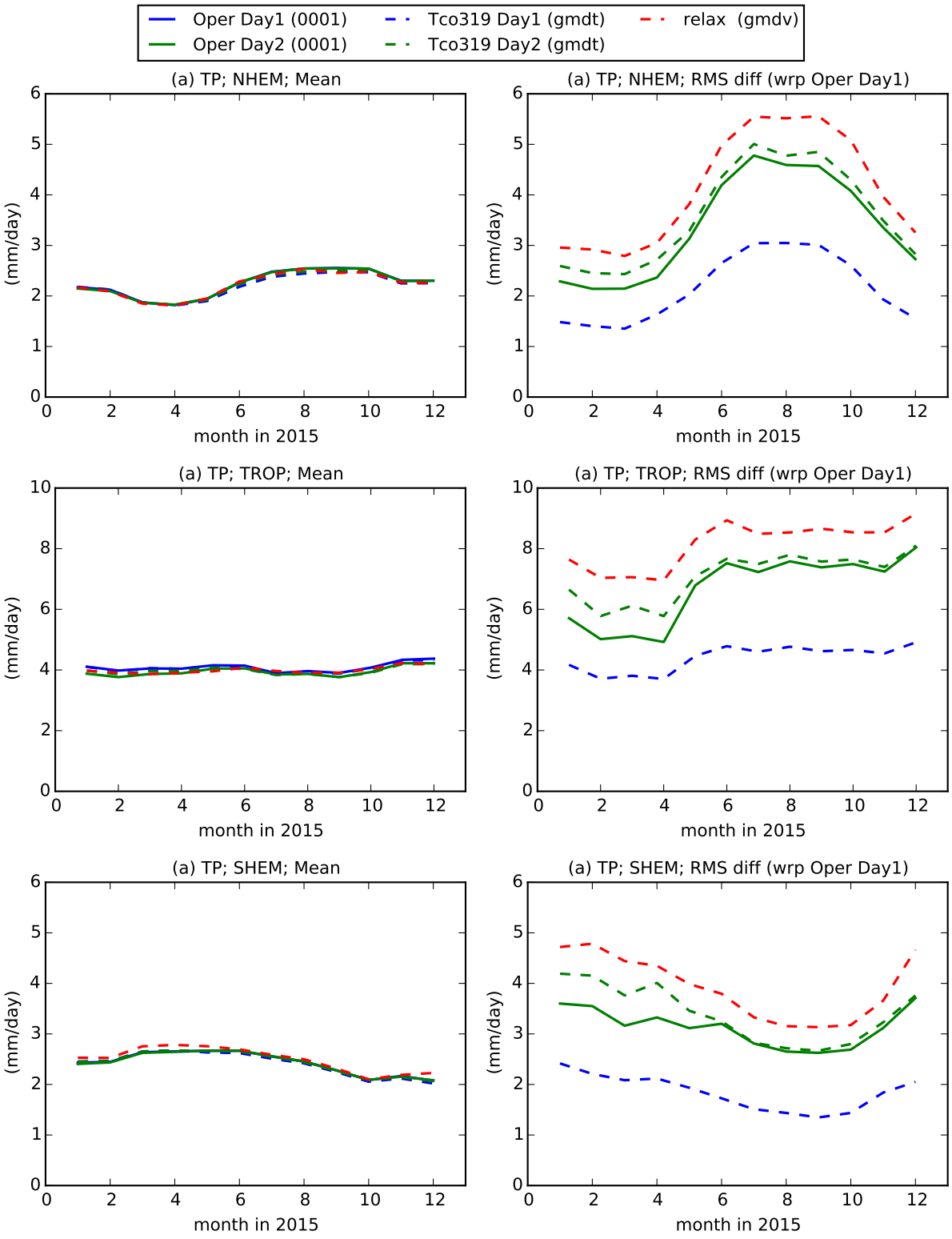

Figure 6: Monthly Means (left) and RMS differences (right) of total precipitation with respect to OPER 0-24 hour forecasts in the Northern Hemisphere (top row), Tropics (middle row) and Southern Hemisphere (bottom row).

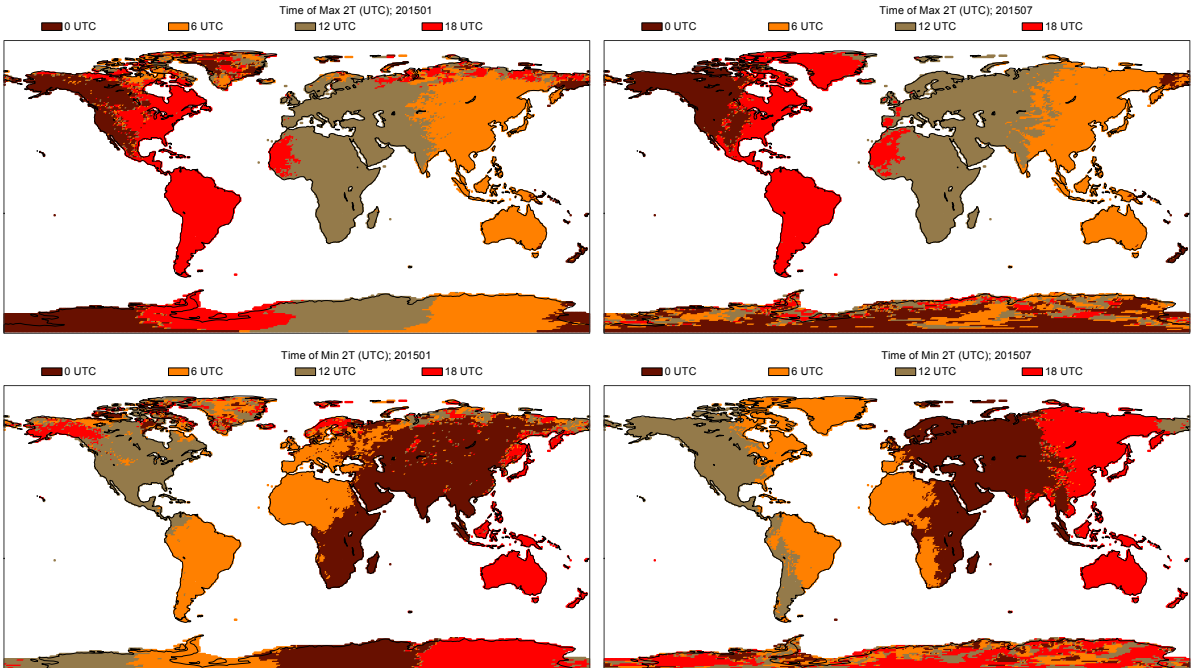

Figure 7: Selected Time in UTC for maximum (top row) and minimum temperature (bottom row) at 2m from monthly mean analyses in January (left) and July (right).

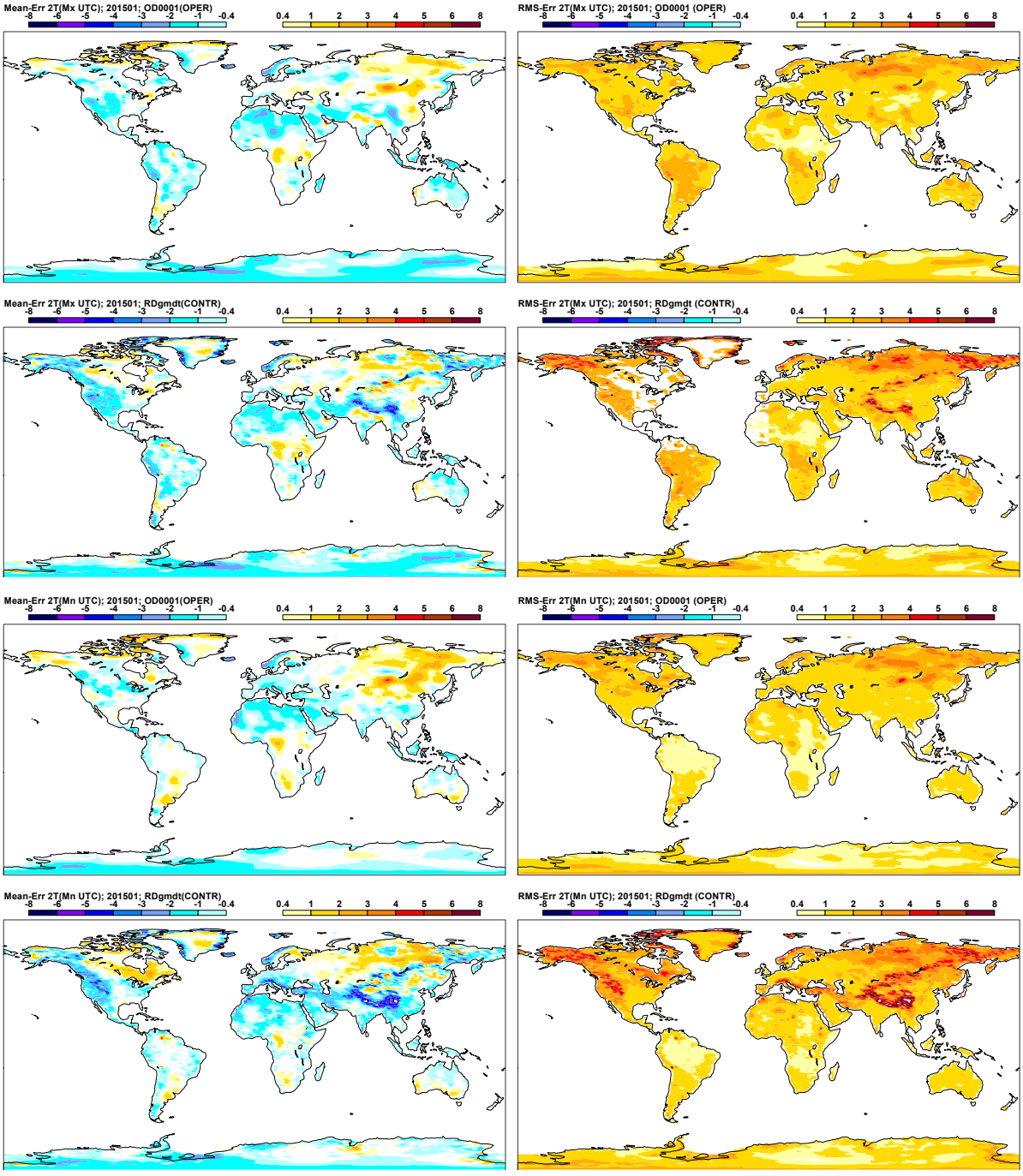

Figure 8: Mean (left) and RMS errors (right) of 2m temperature for January 2015 with respect to the operational analysis. The top panels are for the OPER maximum temperature; the second row is for the control Tco319 maximum temperature; the 3th row is for the operational minimum temperature; and the bottom panels are for the Tco319 minimum temperature. The lead times used to sample the diurnal cycle, are 6, 12, 18 and 24 hours.

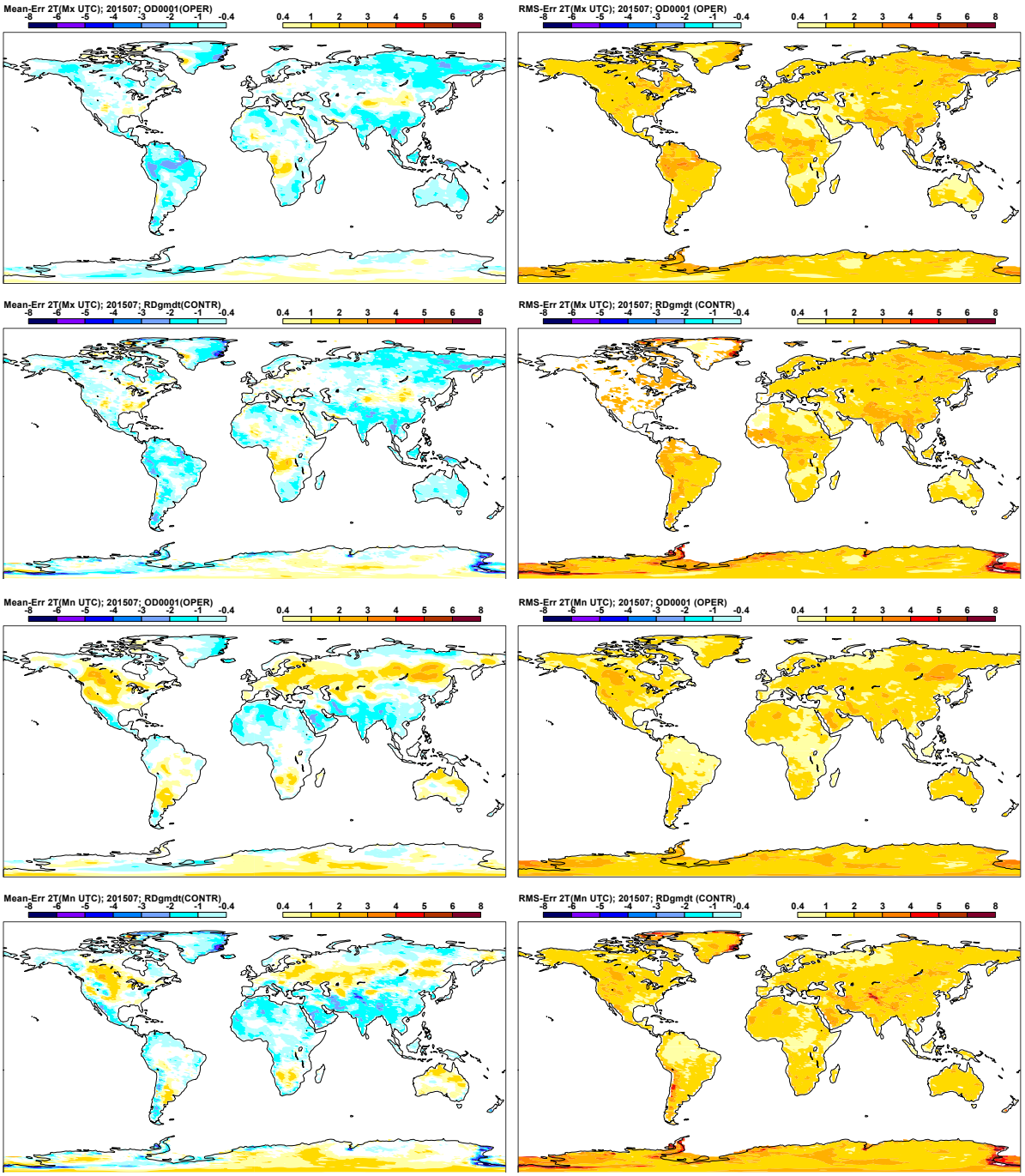

Figure 9: As figure 8, but for July 2015

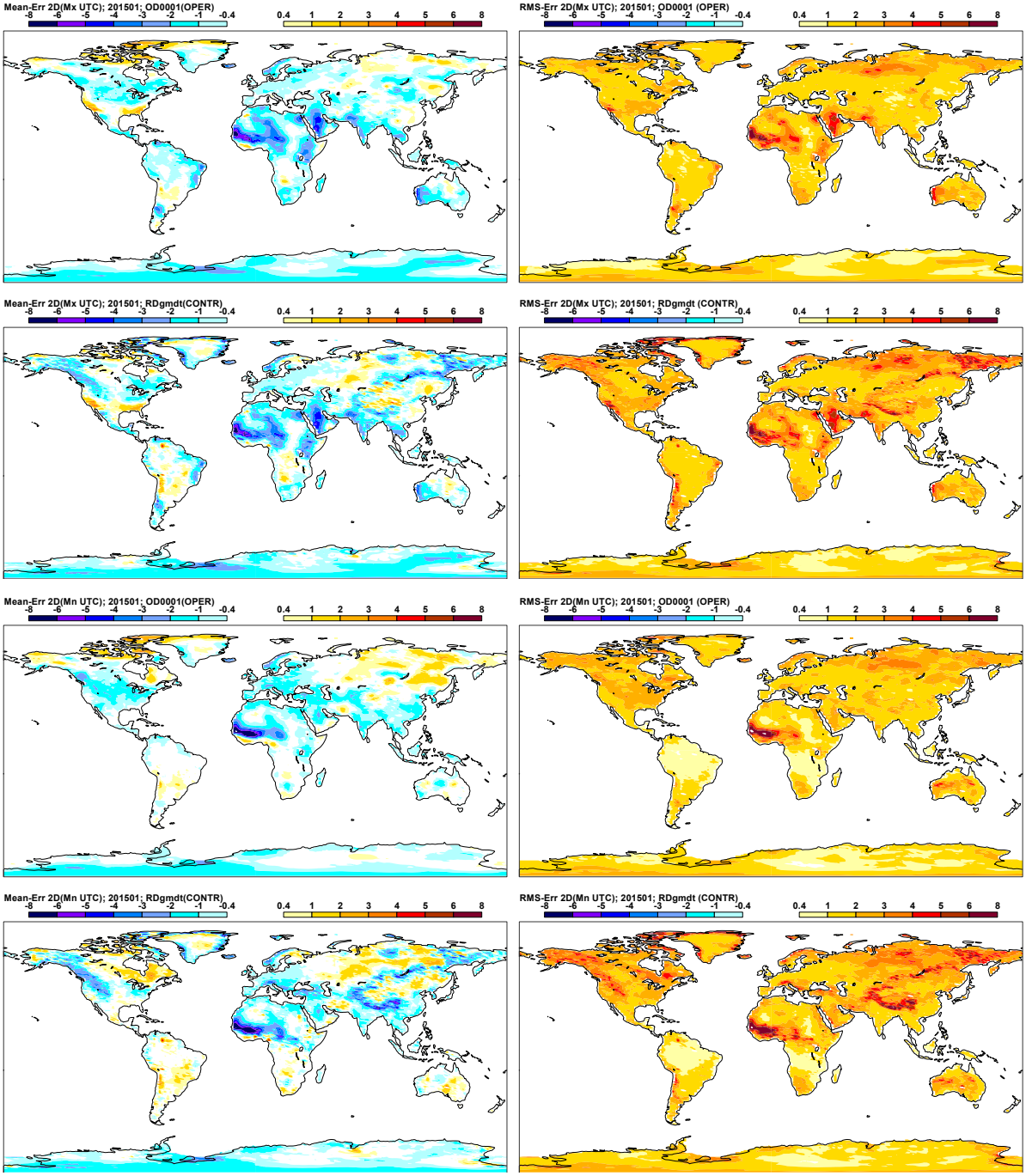

Figure 10: Mean (left) and RMS errors (right) of 2m dewpoint for January 2015. The top panels represent dew point errors of OPER at the time of maximum temperature, the 2nd row represents Tco319 forecasts also at the time of maximum temperature, the 3th row represents dew point errors of OPER at the time of minimum temperature, and the bottom panels represent Tco319 at the time of minimum temperature. The lead times used to sample the diurnal cycle, are 6, 12, 18 and 24 hours.

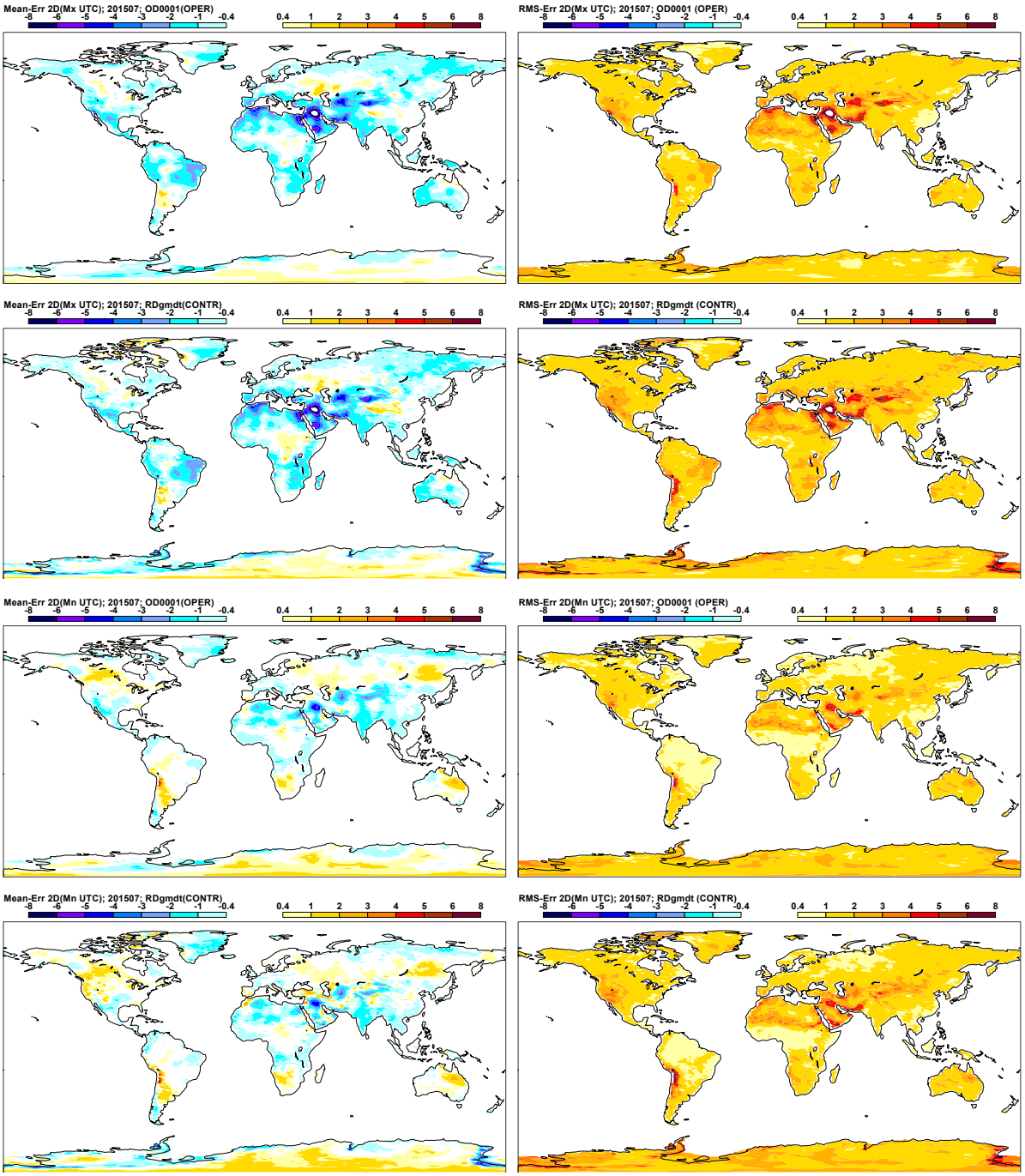

Figure 11: As figure 10, but for July 2015

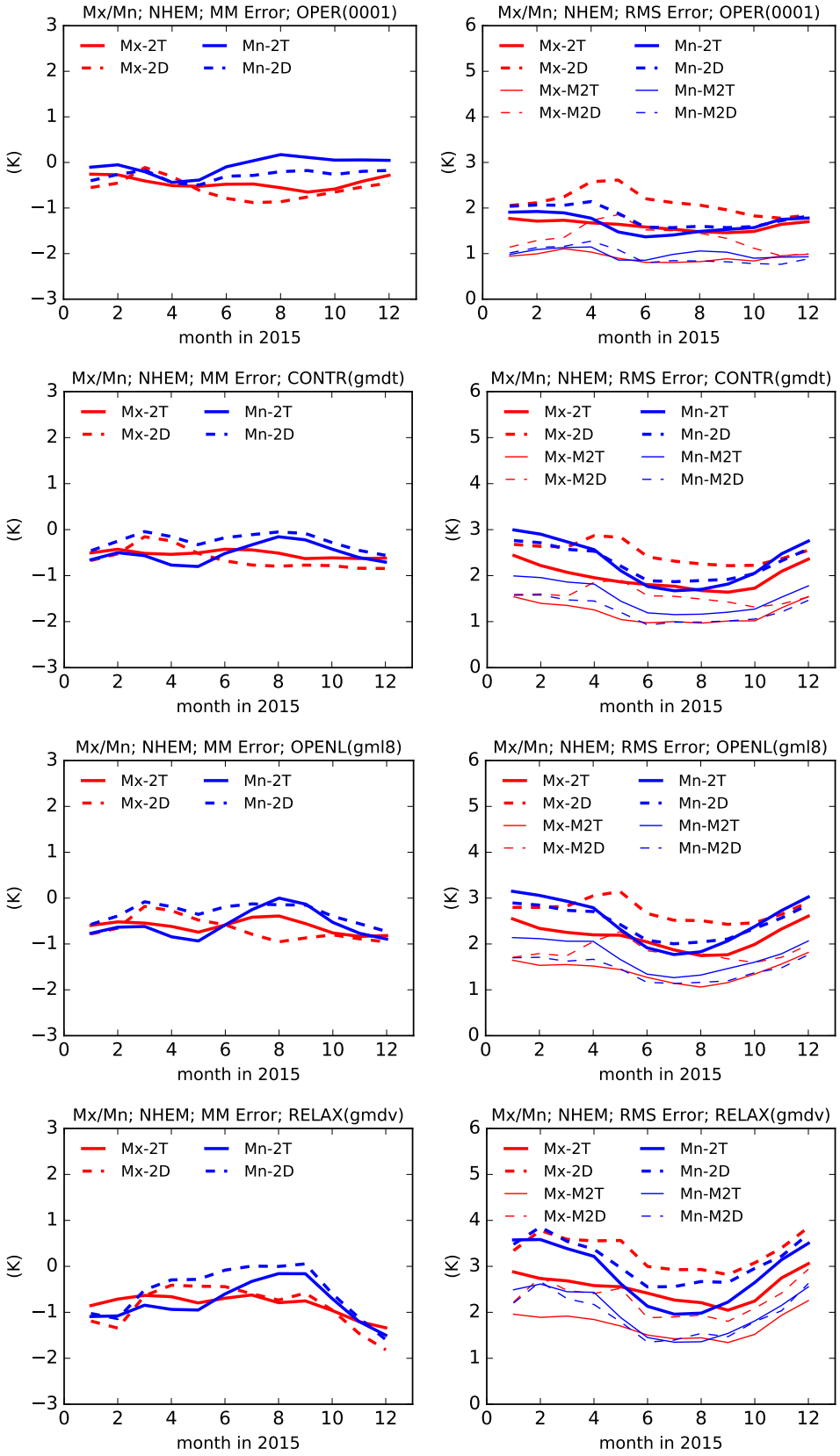

Figure 12: Northern Hemisphere monthly mean error (left column) and RMS error (right column) of Mx/Mn-2T and Mx/Mn-2D for the 4 experiments (rows). The RMS is computed in two different ways: The standard RMS (thick lines) and the RMS of the monthly averages of the error (thin lines).

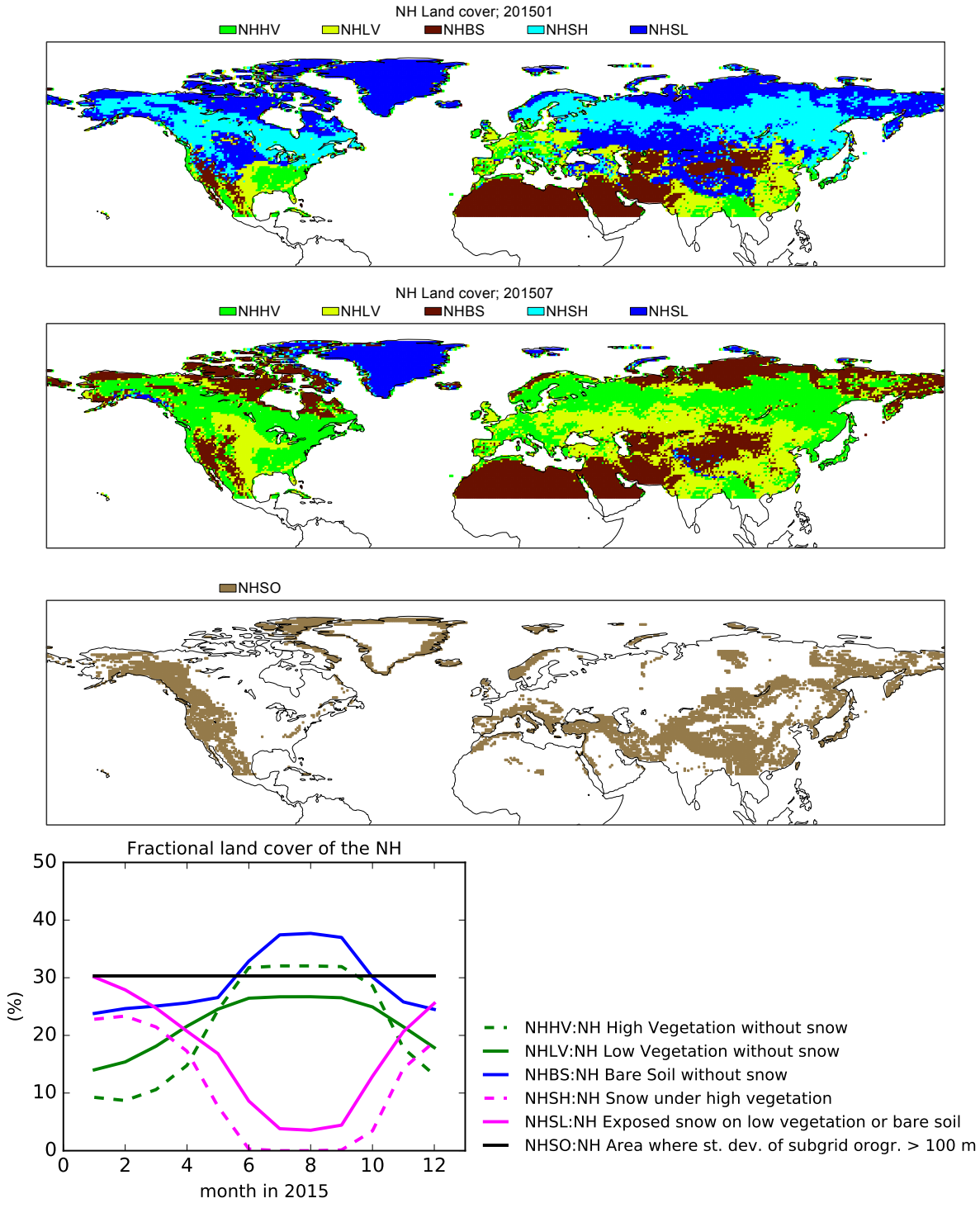

Figure 13: Surface fractions of dominant land use categories in January (upper panel), and July 2015 (second panel). The third panel outlines the area where the standard deviation of subgrid orography is larger than 100 m. The bottom panel shows the seasonal evolution of land use fraction relative to the considered land area (all land North of 20^o N). The underlying data for these plots is from climate fields as used by OPER and snow depth as analysed by the operational system in 2015.

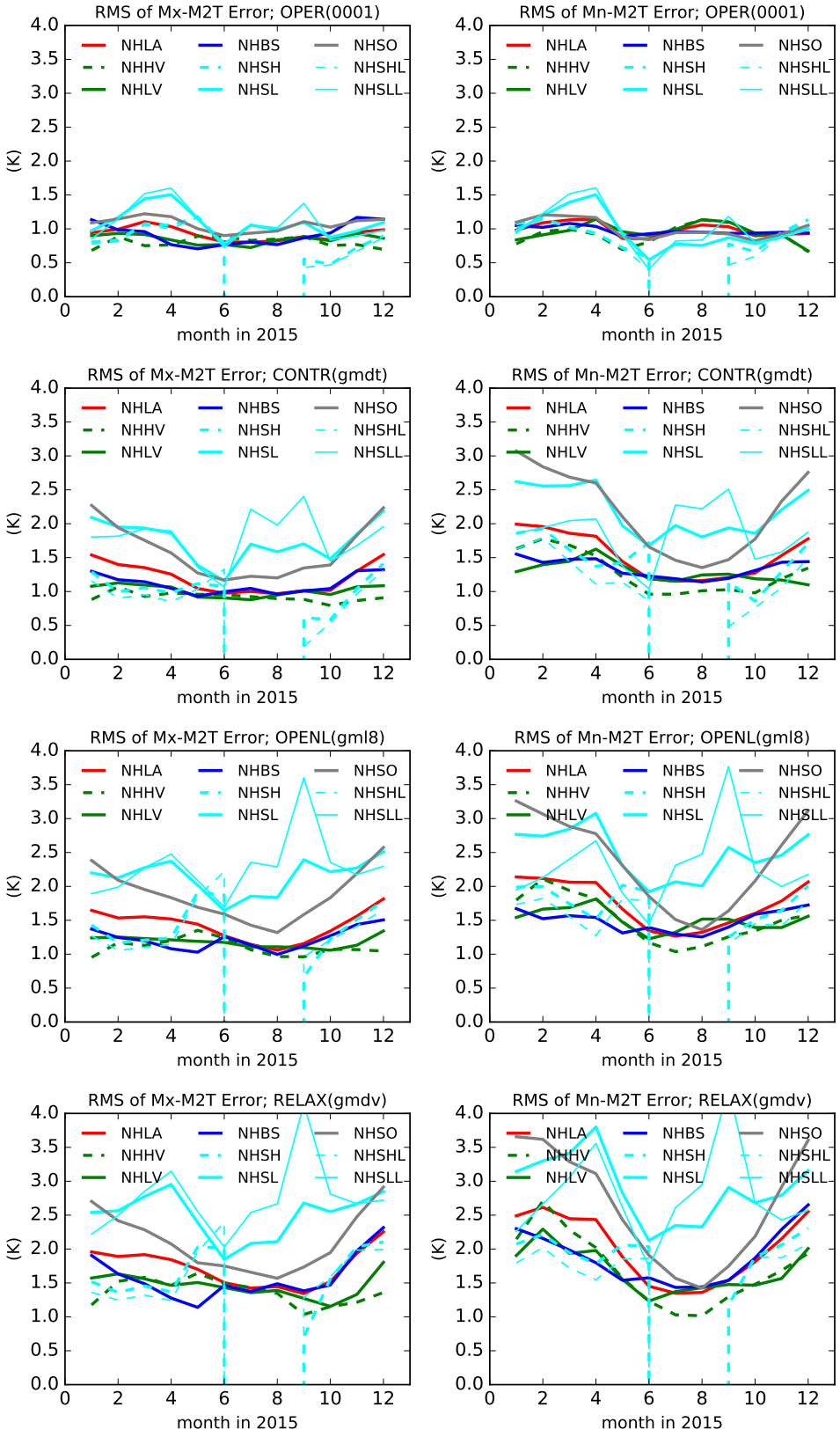

Figure 14: Northern Hemisphere RMS of the monthly mean error of Mx-2T (left) and Mn-2T (right) for the different surface types (lines). The 4 rows show the experiments. The thin lines (NHSHL, NHSLL) are also for the areas with snow, but exclude orography above 1000 m.

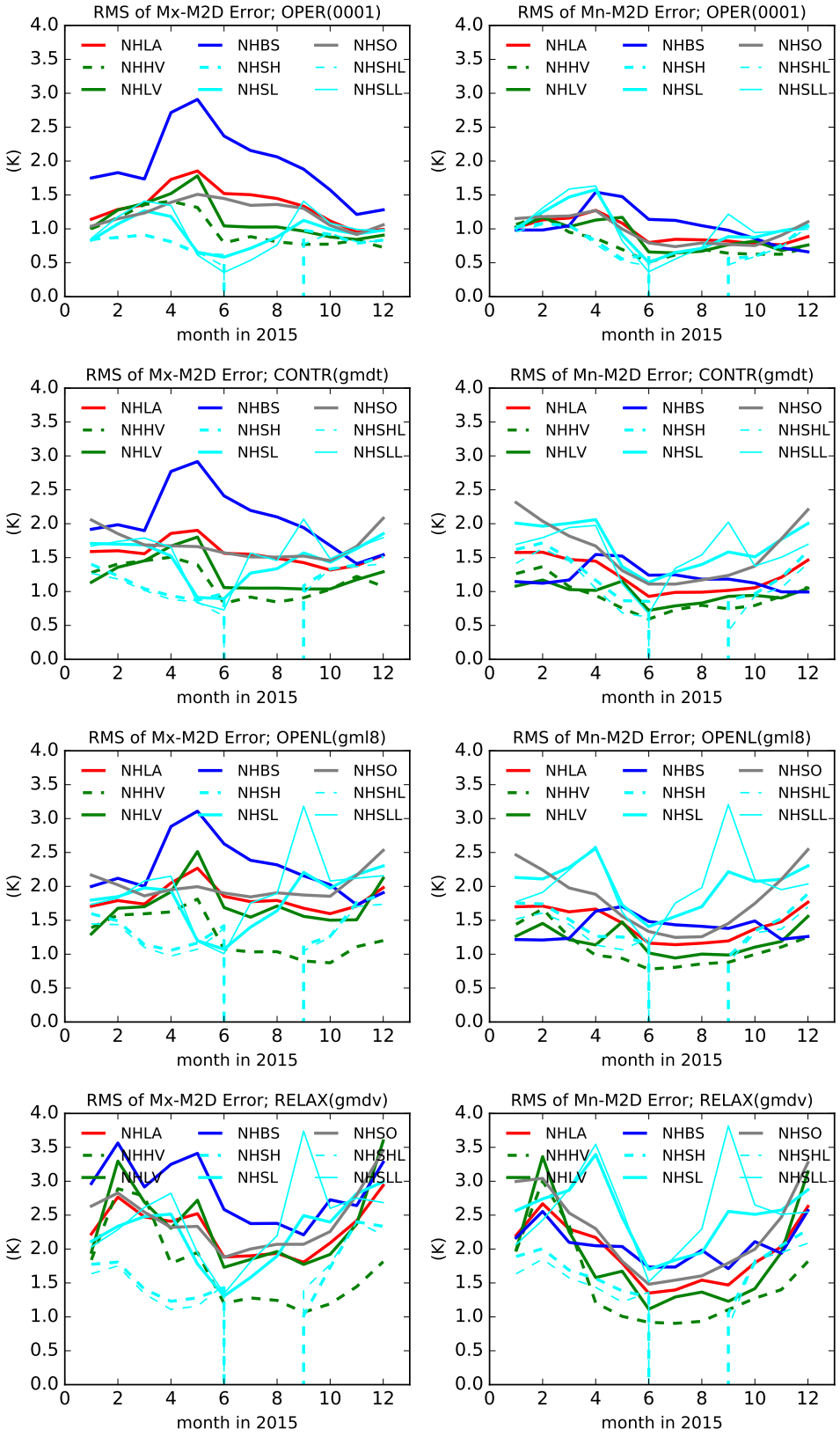

Figure 15: Northern Hemisphere RMS of the monthly mean error of Mx-2D (left) and Mn-2D (right) for the different surface types (lines). The 4 rows show the experiments. The thin lines (NHSHL, NHSLL) are also for the areas with snow, but exclude orography above 1000 m.

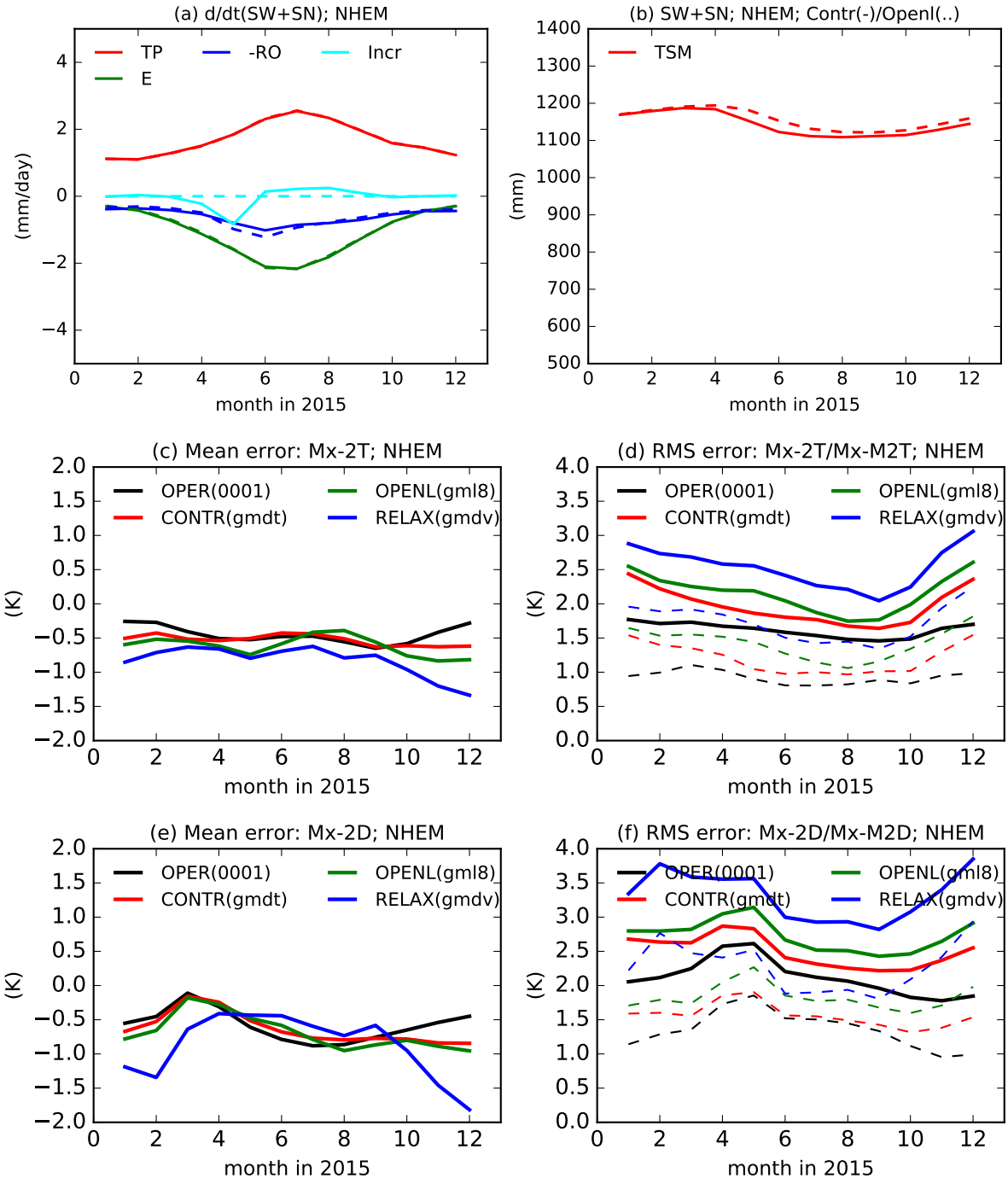

Figure 16: Northern Hemisphere (NHEM; $20^o - 90^oN/180^oW - 180^oE$) monthly mean soil water budget for CONTR (solid) and OPENL (dashed) (panels a,b), combined with Mean/RMS errors of 2T (panels c,d) and Mean/RMS errors of 2D (panels (e,f). The terms in panel a are total precipitation (TP), minus runoff (-RO), soil moisture + snow water increments (Incr), and evaporation (E). The residual of these fluxes drives the evolution of soil moisture + snow water as indicated in panel b. The mean errors in 2T,2D are plotted in panels c,e for OPER, OPENL, CONTR and RELAX. Panels d,f display the RMS errors. The standard RMS of Mx-2T/Mx-2D error is obtained by averaging the squared errors over all points and all days (thick solid line). The RMS of Mx-M2T/Mx-M2D is obtained by first taking the monthly mean errors and then computing the RMS by averaging over all points (thin dashed lines).

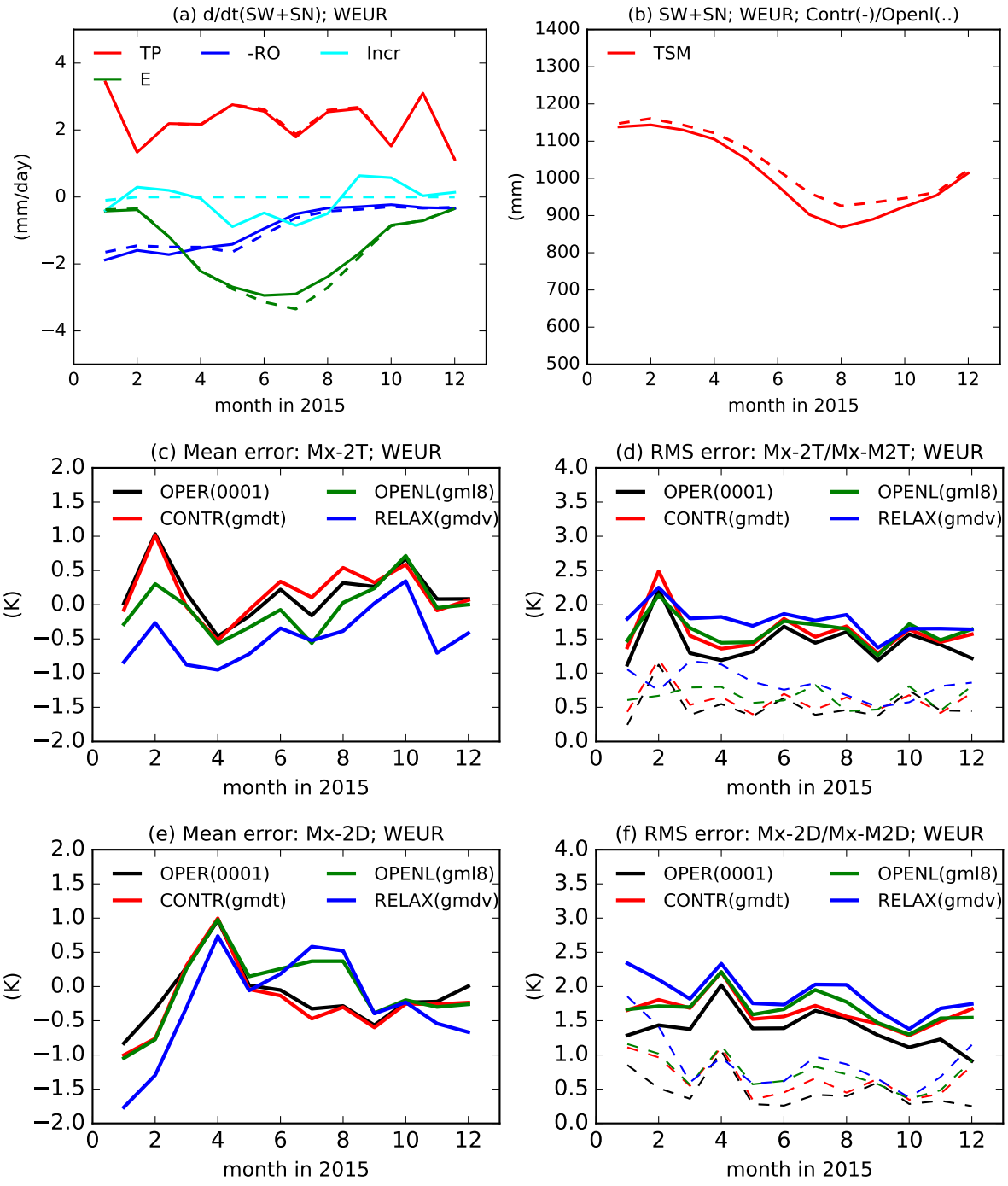

Figure 17: As Fig. 16 for Western Europe (WEUR; $47^o - 52^oN/6^o - 12^oE$).

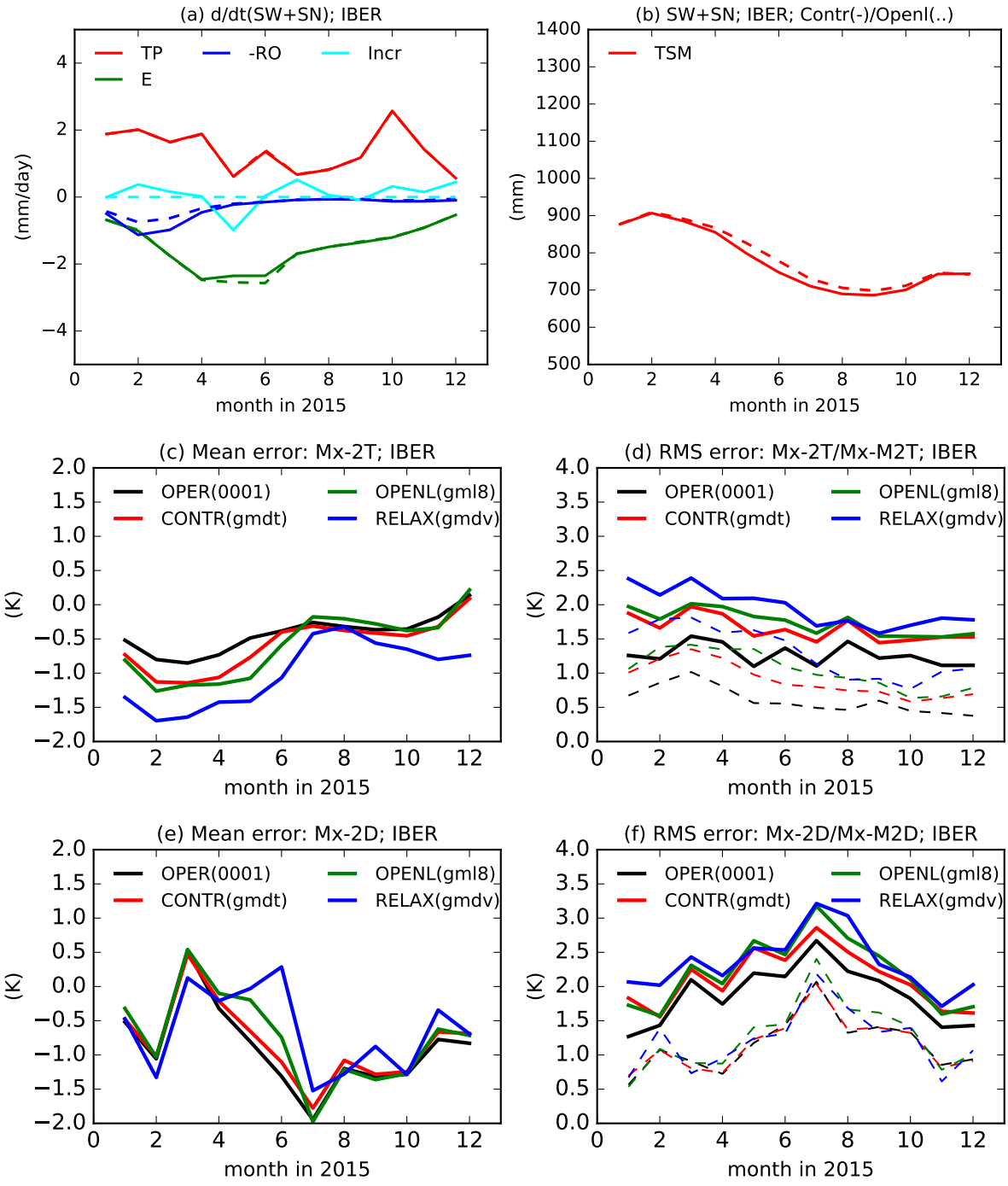

Figure 18: As Fig. 16 for the Iberian area (IBER; $37.5^o - 43.5^oN/1^o - 8^oW$).

Table 2: List of Tco319 sensitivity experiments

Short name	Description	Forecast type	Exp ID
CONTR	Control	Short range	gmdt
RELAX	Control	Relaxation	gmdv
Z0Hd10	Z_{oh} divided by 10	Relaxation	gqc2
LSTd2	$\Lambda_{sk,st}$ stable divided by 2	Relaxation	gqcj
LUSd2	$\Lambda_{sk,unst}$ unstable divided by 2	Relaxation	gqck
LSNd2	$\Lambda_{sk,snow}$ divided by 2	Relaxation	gqd9
RSMm2	$r_{s,min}$ multiplied by 2 (low veg)	Relaxation	gqda
NOAz0	No aggregation of z_{0m}	Relaxation	gqve
LHVd2	$\Lambda_{sk,hveg}$ divided by 2	Relaxation	gqw2
gd065	$g_D = 0.065$ in q-stress function(low veg)	Relaxation	gsrp

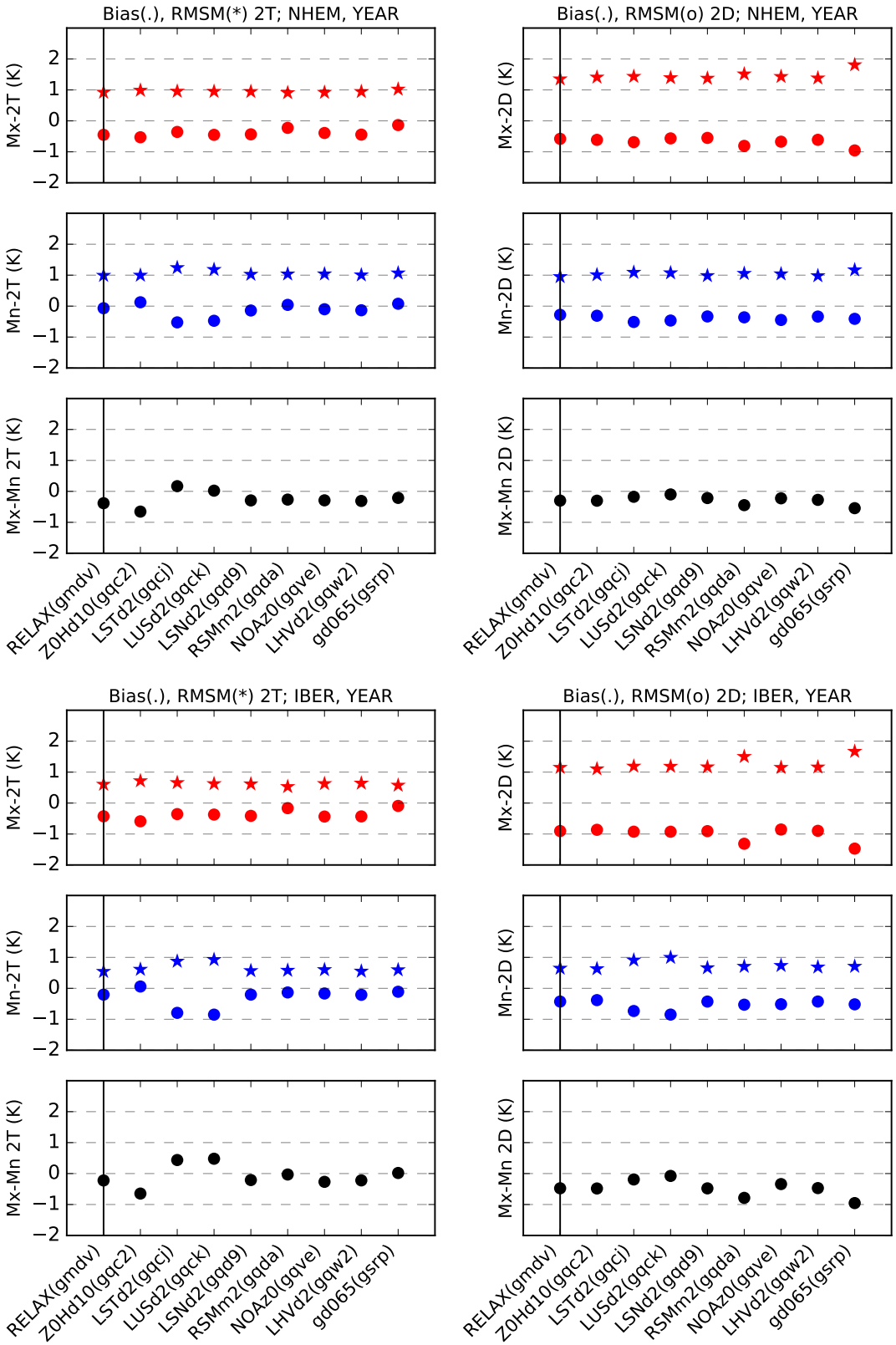

Figure 19: Annual summary statistics for Northern Hemisphere (top 6 panels) and IBERIA (bottom six panels) of the experiments listed in Table 2. The biases are represented by dots, and the RMS errors of monthly means by stars. The left column contains results for 2T, the right column for 2D, with Mx in red, Mn in blue, and Mx-Mn in black. Experiments to the right of the vertical line are sensitivity experiments with relaxation; the vertical line indicates the relaxation control.

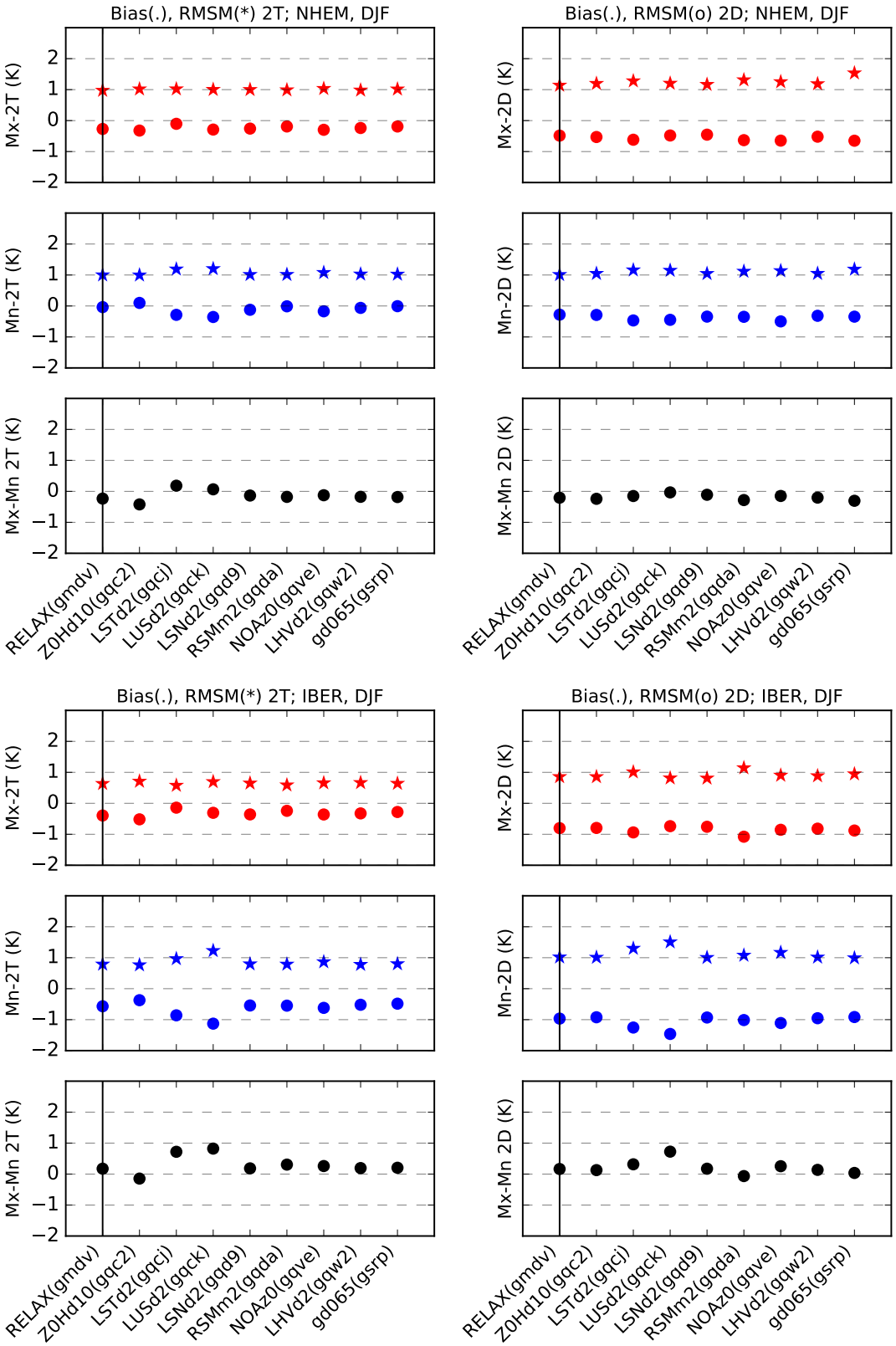

Figure 20: As Fig. 19 for Dec/Jan/Feb.

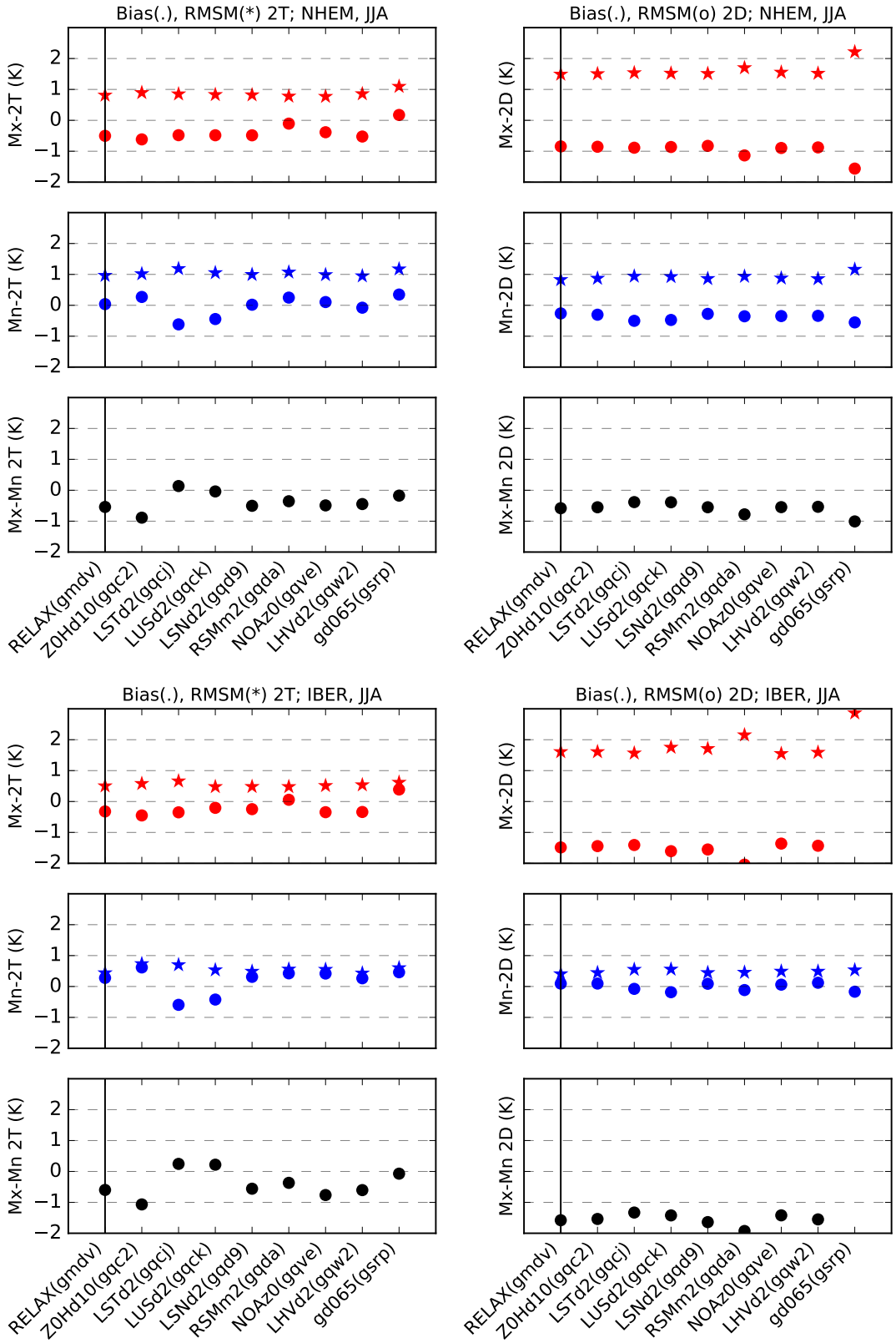

Figure 21: As Fig. 19 for Jun/Jul/Aug.

Table 3: List of data assimilation experiments to investigate interaction of land data assimilation with model bias. The experiments run from 1 March 2015 to 9 August 2015 with CY42R1 at T511 resolution.

Short name	Description	Resolution	DA Exp ID	FC Exp ID
CONTR	Control CY42R1	T511	ge4b	gpgy runs from ge4b
OLDCNVC	Old conv. diurnal cycle	T511		gpm2 runs from ge4b
OLDCNV	Old conv. diurnal cycle	T511	ge4d	gph1 runs from ge4d
OLDOBS	Old observation errors	T511	geud	gpgz runs from geud

A Appendix: Model error and soil moisture data assimilation

On 12 May 2015, model library CY41R1 was introduced in operations. Among other changes, two changes were made with impact on the soil moisture analysis: (i) a change in the convection closure to improve the diurnal cycle of convective precipitation, and (ii) a reduction of the observation errors of SYNOP temperature and humidity in the analysis of soil moisture and soil temperature. The latter error settings were reduced from 2 to 1 K for 2m temperature and from 10% to 4% for 2m relative humidity (similar experiments were performed by Fairbairn et al. (2019)). The convection change reduced the day-time dry bias over land, and since the soil moisture analysis relies heavily on atmospheric moisture observations, it is possible that the two changes interact. To investigate, three different data assimilation experiments were performed: (i) a control experiment with CY42R1[1](CONTR), (ii) one with the convection changes reverted (OLDCNV), and (iii) an experiment with the old observation errors (OLDOBS) (see table 3 for more details). Separate forecasts were run for each experiment to have 6-hourly post-processing. An additional forecast only experiment with the old convection scheme was run from the control to see the impact of the convection change in forecast mode. It turns out that the forecasts with the old convection scheme look very similar irrespective whether they were initialized from the control data assimilation experiment or from its own data assimilation experiment.

The summary plot for the NHEM (Fig. A.1) shows that compared to OLDOBS (i) the average soil moisture increments were increased (as expected), (ii) on the average soil water is deleted, (iii) RMS 2T and 2D errors are reduced mainly through the data assimilation, as also shown by Fairbairn et al. (2019), (iv) the error reduction is dominated by the systematic part of the RMS errors, and (v) the NHEM average cold and dry bias reduced mainly through the convection change. The regional signals e.g. WEUR and IBER (Figs. A.2 and A.3) are more variable. Although the new cycle (compared to OLDOBS) systematically increases the soil water increments, the impact on 2T and 2D is not always positive; it depends on the relative weight of temperature and moisture observations (Jacobians and error estimates) and on the consistency of biases in 2T and 2D. IBER is an example where the day time temperature is too low and the air too dry. Soil moisture can not explain such an error structure, so there must be another reason or combination of reasons to explain this error.

It is interesting that the current land/atmosphere/data assimilation configuration of the IFS (De Rosnay et al. 2013) behaves differently from the one described by Drusch and Viterbo (2007). The old combination of TESSEL and OI systematically added soil moisture to prevent a dry drift of the model. The current configuration (HTESSEL with Kalman filter) does the opposite. We know that the water holding capacity of the land surface has increased with HTESSEL (Balsamo et al. 2009) but perhaps unstressed evaporation is too large? Additional evidence for an overestimation of evaporation is provided by a study with a carbon model, in which canopy stress is derived from carbon uptake (Boussetta et al. 2013), and by an evaluation of ERA5 with FLUXNET data (Martens et al. 2020).

[1]For technical reasons, CY42R1 was used instead of CY41R1. Both cycles are the same from the scientific point of view.

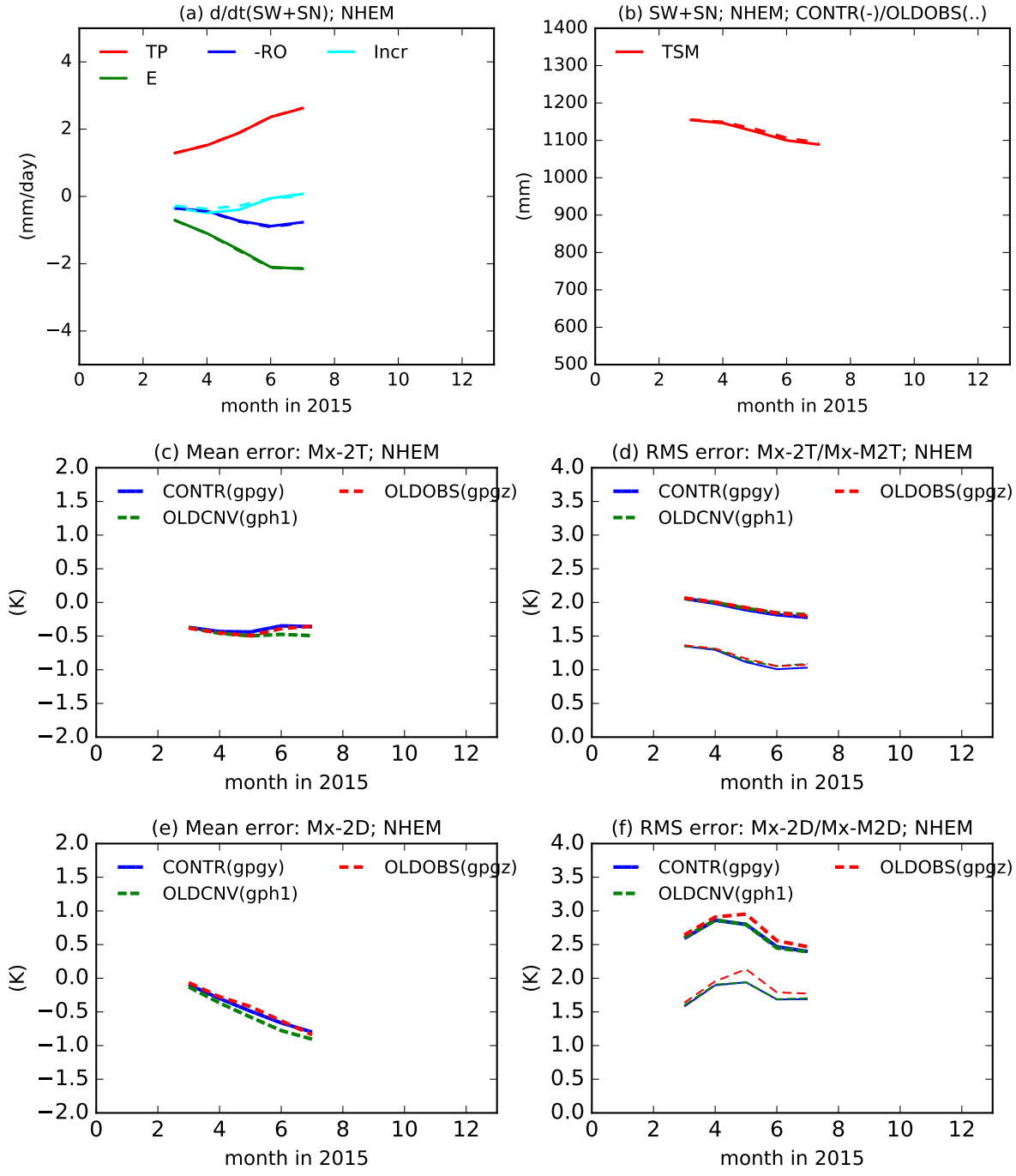

Figure A.1: Monthly soil water budget for CONTR (solid) and OLDOBS (dashed) (a), total soil water + snow evolution (b), mean Mx-2T error (c), RMS-2T (thick lines) & RMS-M2T (thin lines) (d), mean Mx-2D error (e), RMS-2D (thick lines) & RMS-M2D (thin lines) (f) for the Northern Hemisphere (NHEM; $20^o - 90^oN/180^oW - 180^oE$*).*

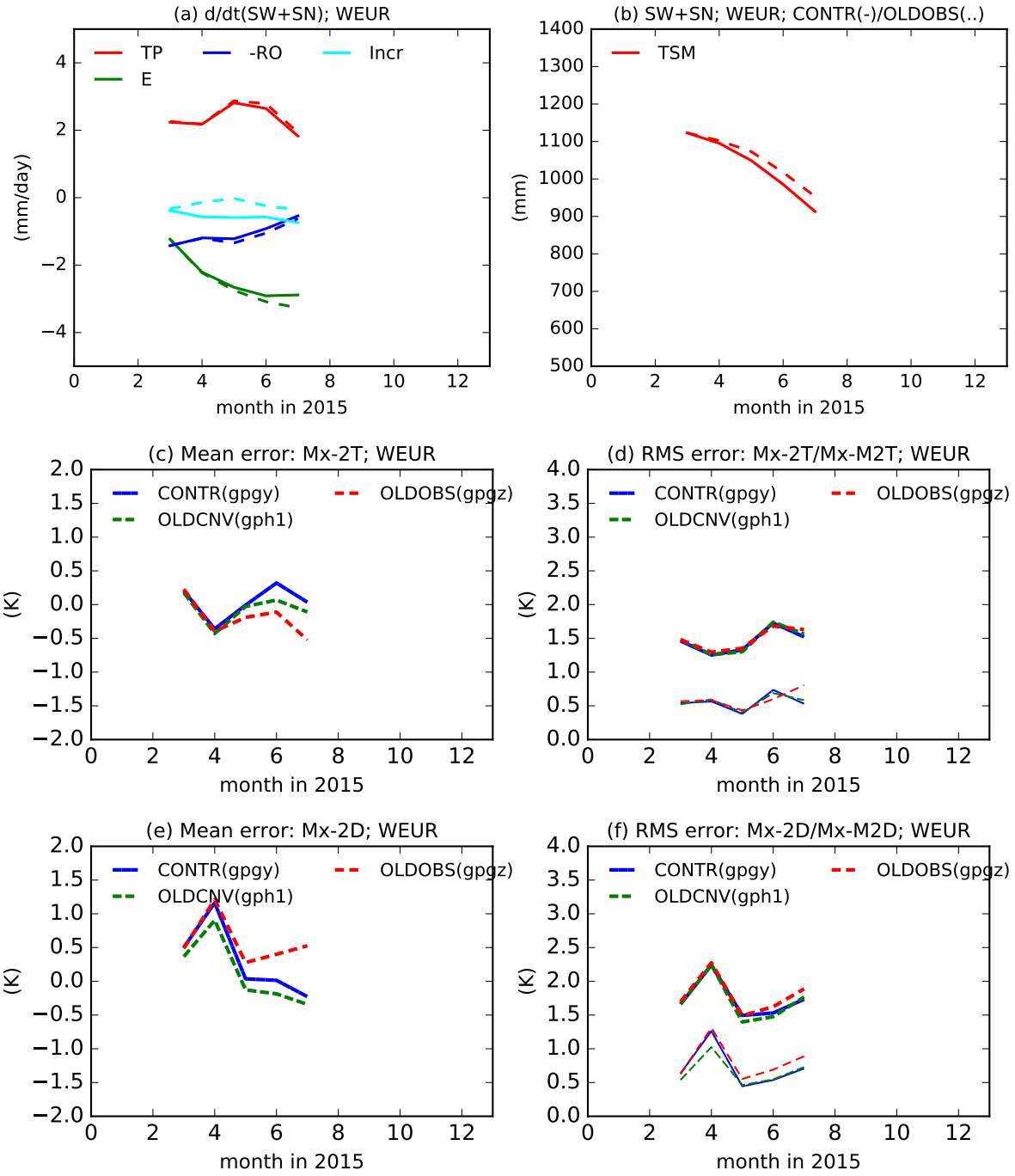

Figure A.2: Monthly soil water budget for CONTR (solid) and OLDOBS (dashed) (a), total soil water + snow evolution (b), mean Mx-2T error (c), RMS-2T (thick lines) & RMS-M2T (thin lines) (d), mean Mx-2D error (e), RMS-2D (thick lines) & RMS-M2D (thin lines) (f) for the area Western Europe (WEUR; $47^o - 52^oN/6^o - 12^oE$*).*

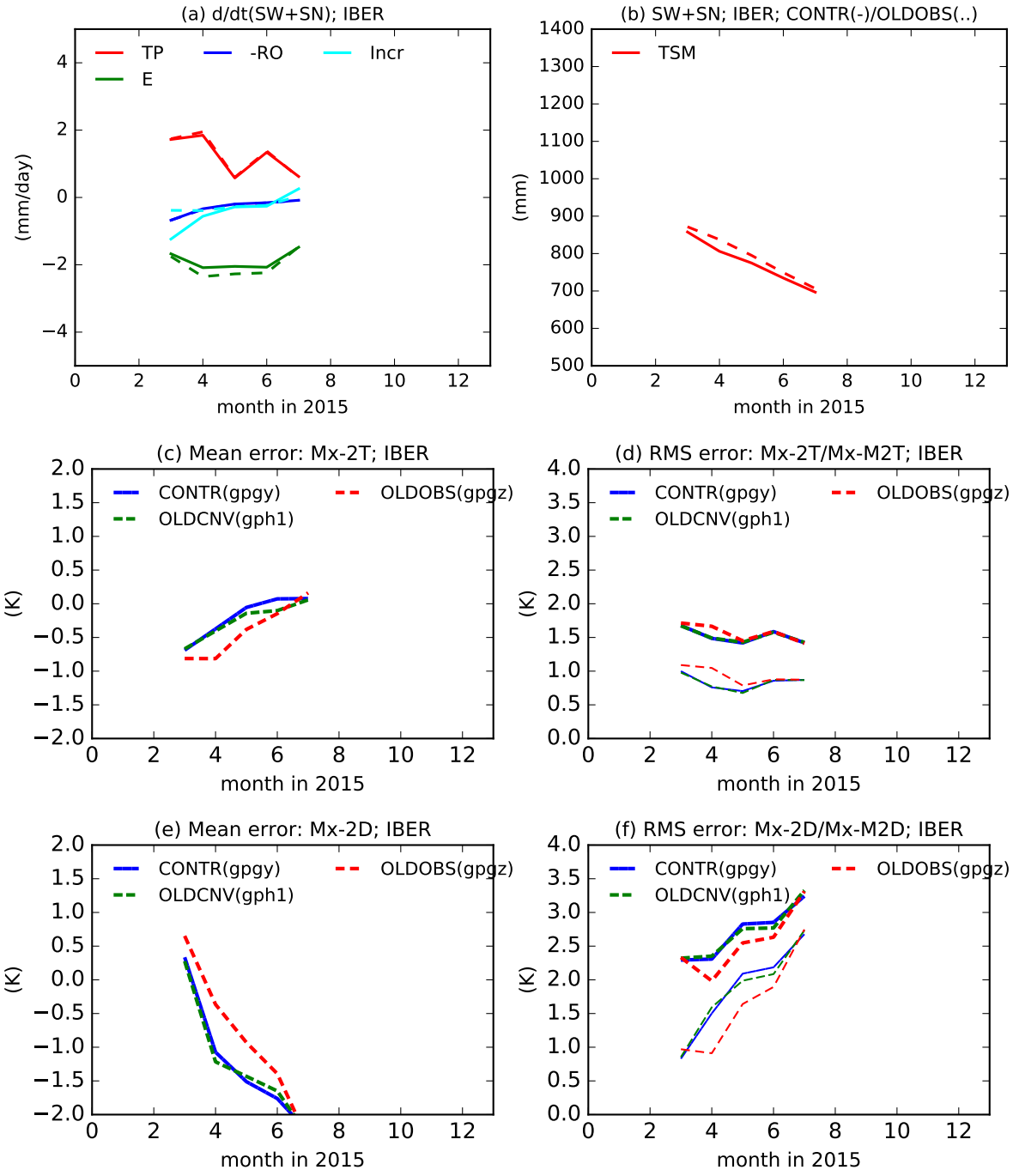

Figure A.3: Monthly soil water budget for CONTR (solid) and OLDOBS (dashed) (a), total soil water + snow evolution (b), mean Mx-2T error (c), RMS-2T (thick lines) & RMS-M2T (thin lines) (d), mean Mx-2D error (e), RMS-2D (thick lines) & RMS-M2D (thin lines) (f) for the area Iberia (IBER; $37.5^o - 43.5^oN/1^o - 8^oW$*).*

B Appendix: Meso-scale variability and resolution effects

In section 4 it was shown that the night time temperature errors have a resolution dependent bias over orography. The low resolution forecasts show a lower night time temperature over orography than the high resolution forecasts. This behaviour was also discussed in Sandu et al. (2014, internal memo RD14-315), which showed that temperature differences over orography between high and low resolution forecasts are mainly due to differences in the resolved orography. Here we explore the hypothesis that high resolution forecasts have more meso-scale variability resulting in more wind shear over orography. The turbulent diffusion in stable situations is driven by wind shear and therefore more wind shear results in more mixing and less cooling at night. The top panel of Fig. B.1 shows profiles of wind shear over the Himalayas area from forecasts with the Tco1279, Tco639, and Tco319 models. In the height range of 0 to 500 m above the surface, we see typically 10% more shear at Tco1279 than at Tco639. Comparing Tco1279 to Tco319 the difference is about 20%, but also the height range increases to about 750 m. As expected the additional shear comes from the high wave numbers. The wave numbers below truncation 21 do not show this signal. The same profiles for the relatively flat area over Russia (bottom panel of Fig. B.1) does not show the resolution dependence of shear, which suggests that the additional resolved shear at high resolution is related to orography, which is perhaps not surprising because orography generates a lot of variability at all scales. It is not clear why the vertical scale changes from 500 to 750 m when changing resolution from Tco639 to Tco319. It may be related to the horizontal scale of truncation, but also to the increase of amplitude of the subgrid orography towards lower resolution (see caption of Fig. B.1).

Obviously, missing meso-scale variability should be parametrized in a resolution dependent way. Subgrid scale orography effects are currently accounted for in the momentum equations with parametrization schemes for blocking and gravity wave generation (Lott and Miller 1997) and turbulent orographic form drag (Beljaars et al. 2004). However, these schemes ignore heat / moisture transport, and research on this topic is still in its infancy. A possible way forward is to analyse shear from observations and from very high resolution simulations.

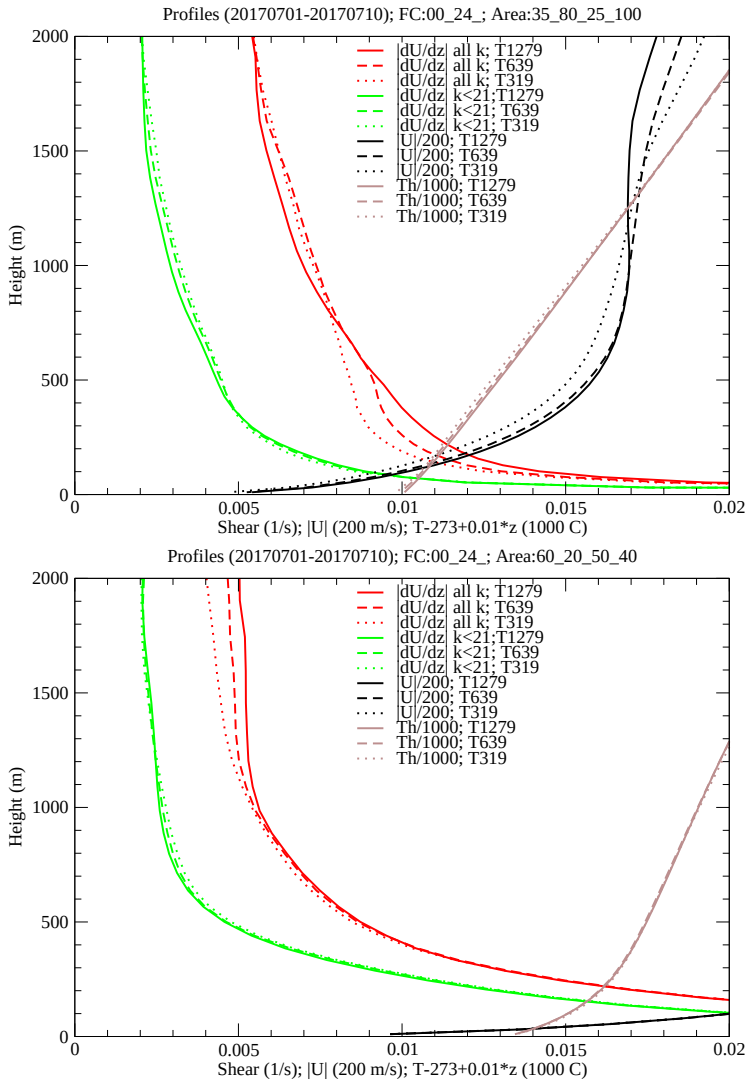

Figure B.1: Profiles of absolute shear based on all wave numbers (red) and wave numbers below 21 (green) from the IFS at Tco1279 (solid), Tco639 (dashed), and Tco319 (dotted). The shear is the RMS of instantaneous wind gradients verifying at 0 UTC, with area averaging and averaging over 10 forecasts (20170701-20170710). The black curves represent the mean absolute wind speed and the brown curves the mean potential temperature. The top figure shows the Himalayan mountains ($25^o - 35^oN/80^o - 100^oE$) and the bottom figure shows a large area over Russia ($50^o - 60^oN/20^o - 40^oE$). The Himalayan area is very mountainous with standard deviations of subgrid orography of 102, 165 and 228 m for the Tco1279, Tco639 and Tco319 model configurations. The Russian area has standard deviations of 12, 8 and 5 m.

C Appendix: Low level profile shape of temperature and dewpoint

In section 4 it was shown that in summer the model has a dry bias at the 2m level combined with a cold bias. This feature is most pronounced at latitudes from $25^o - 40^oN$ and can not be explained by errors in the Bowen ratio, and therefore soil moisture increments may reduce either temperature or dew point errors, but never both. However, the low level departures from radio sondes in data assimilation do not appear to show the same signals as the SYNOP verification. This raises the question whether the model has errors in the vertical gradients of temperature and dewpoint.

To explore further, the difference of temperature and dewpoint between the 2m level and model level 130 (about 200 m above the surface) has been compared with observations. The latter is done by taking the difference at sonde locations between the SYNOP temperature and the sonde temperature at the same level as model level 130. This level has been selected because it is about the height where we expect the sonde to provide reasonable observations after the initial shock at launch. Two examples of time series of daily observations at 12 UTC and their model equivalent are given in Fig. C.1, namely for Madrid and for Lindenberg.

The data from Madrid indeed shows that the difference between the 2m and 200m levels is larger than observed for dewpoint and smaller than observed for temperature. However, if we look at other stations, e.g. Lindenberg (right hand panels of Fig. C.1) the signal is not confirmed. To have more statistics, all the sondes in the European area are put together for all days in July 2015 and stratified according to wind speed and latitude (see Fig. C.2). Stratification against longitude, model minus station height and subgrid orography was also considered (not shown), but nothing systematic could be demonstrated.

The two left columns of Fig. C.2 show that the temperature and dewpoint differences do not depend strongly on wind speed, except at very low observed winds. The model has less often low winds for reasons that are not clear. Unfortunately the observation processing utilities round the SYNOP wind to integer values in m/s, so it is hard to draw conclusions. In the 0 to 1 m/s class, observations indicate a mean dewpoint difference of 4.5 K whereas the model has 2 K. The right columns of Fig. C.2 show the latitude dependence. The strong increase of observed dewpoint difference at low latitudes is not reproduced by the model.

It is too early to draw conclusions, but there is a suggestion that the model does not reproduce the vertical gradients correctly, and is too well mixed in low wind speed situations and at lower latitudes. Mixing processes at the boundary layer top are another source of uncertainty, and are well known to have impact on near surface boundary layer moisture and its diurnal cycle (e.g. Beljaars 1994). Another possibility is that the SYNOP observations have a representativeness issue. They tend to be in valleys of the undulating or mountainous terrain. These areas may keep soil moisture longer than the elevated areas and therefore evaporate more and show higher near surface humidity.

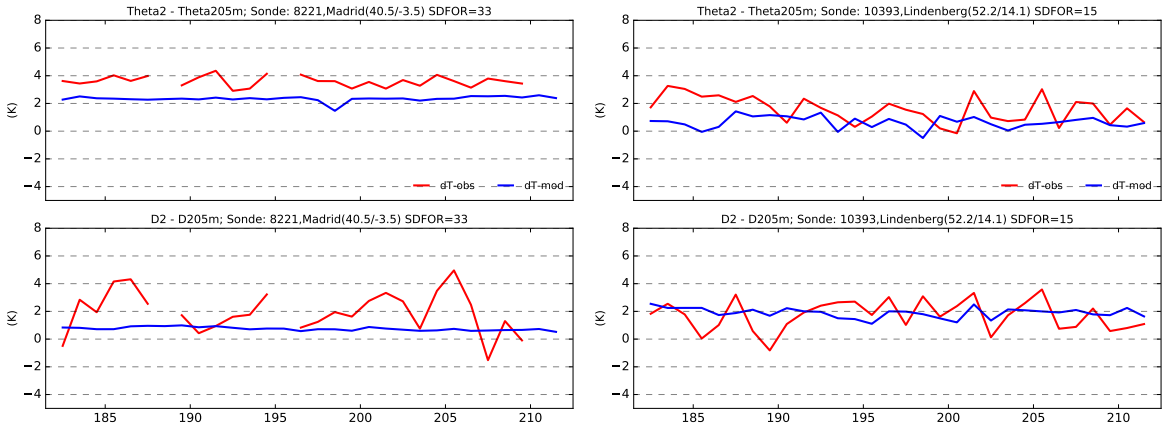

Figure C.1: Time series (July 2015) of temperature difference between 2m and 200 m in the top panels, and similarly for dewpoint in the lower panels for the sonde stations Madrid(left) and Lindenberg(right) 12 UTC. The blue curve represents the model; the red curve the surface SYNOP combined with the sonde observation at 200m.

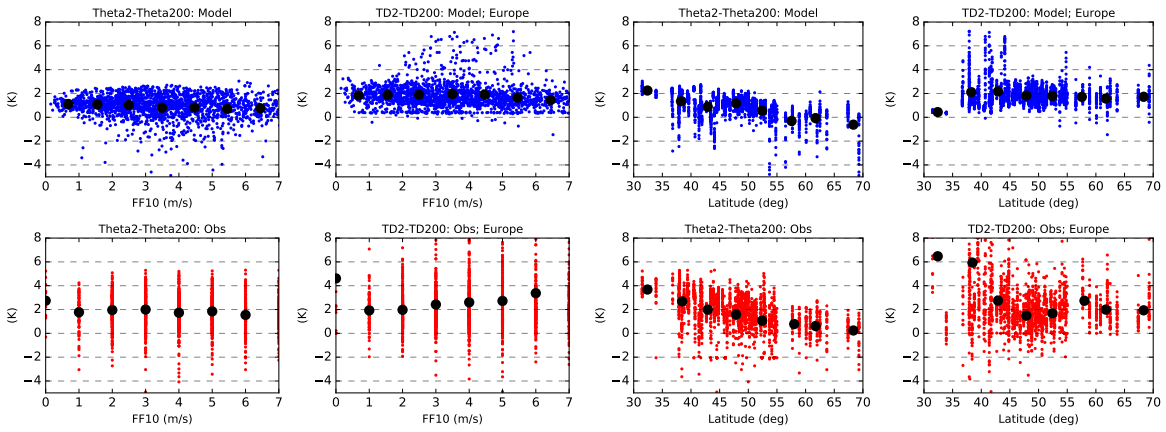

Figure C.2: Difference between temperature at 2m and 200 m above the surface in the left panels, and similarly for dewpoint in the right panels, for all radio sonde stations at 12 UTC in the area ($25^o - 70^oN/10^oW - 28^oE$) in July 2015. The top panels are for the operational high resolution model. The lower panels represent observations. The black dots are bin averages of the individual data points. The left Figure shows the dependency on wind speed at 10m. The right Figure shows the dependency on latitude.