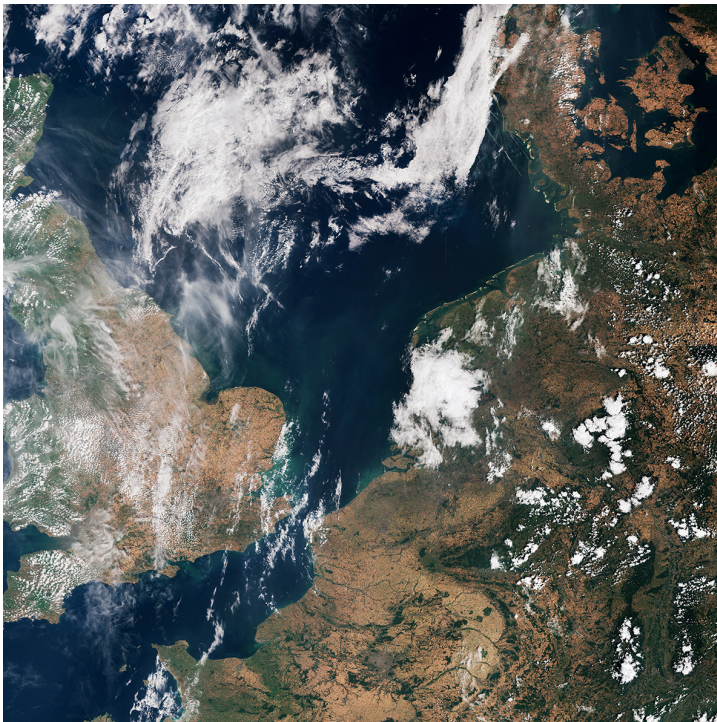


ECMWF Feature article

.....
from Newsletter Number 157 – Autumn 2018

METEOROLOGY

.....
Addressing biases in
near-surface forecasts
.....



Satellite image of Europe – 25 July 2018 © ESA

www.ecmwf.int/en/about/media-centre/media-resources

doi: 10.21957/eng71d53th

This article appeared in the *Meteorology* section of *ECMWF Newsletter No. 157 – Autumn 2018*, pp. 20-25.

Addressing biases in near-surface forecasts

Thomas Haiden, Irina Sandu, Gianpaolo Balsamo, Gabriele Arduini, Anton Beljaars

ECMWF's medium-range forecasts of near-surface weather parameters, such as 2-metre temperature and 10-metre wind speed, have become more skilful over the years, alongside improvements in upper-air scores. There are, however, persistent biases in these forecasts which have proved difficult to eliminate. In-depth investigations carried out at the Centre show that these biases are closely related to the coupling between the atmosphere and the land surface in the Integrated Forecasting System (IFS).

The biases are also related to other processes, such as turbulent mixing, radiation and clouds. In some cases, the representation of these other processes leads to errors which partially cancel some of the errors that can be attributed to the atmosphere–land coupling. See Box A for more details on such 'compensating errors'. A deeper understanding of the underlying causes is necessary to address biases in near-surface weather parameters in a way that ensures increased physical realism and reduces compensating errors. Because of atmosphere–surface feedback mechanisms, an improved representation of surface fluxes may also lead to an increase in medium- and extended-range predictive skill overall.

Compensating errors

A

Increasing the physical realism of surface processes in a model to reduce systematic biases may increase the root-mean-square error (RMSE) because different kinds of errors may no longer partially cancel each other. For example, 2-metre temperature (T2m) is computed diagnostically in the IFS from the temperature at the lowest model level and the skin temperature. There is a limiter in the computation of T2m which becomes active in very stable, low wind situations, and which prevents the T2m from deviating too strongly from the temperature at the lowest model level. Removing this limiter would be physically desirable, but tests have shown that doing so in the current model setup increases the RMSE. This is because errors introduced by the limiter and cloudiness errors partially cancel each other: if the limiter is removed, negative

temperature errors at night are increased in cases where the forecast underestimates cloudiness.

In more general terms, trying to increase the realism in one process can leave the model more exposed to errors in other processes. Another example is the strength of the thermal coupling between the surface (the vegetation canopy) and the uppermost soil layer. Decreasing this coupling allows the surface to cool more strongly and produce stronger surface inversions, more in line with observations. However, it also makes T2m in the model more reactive and increases errors in cases where cloudiness is underestimated. The solution to the problem lies in identifying and properly attributing errors in all contributing processes, and then reducing these errors at the same time.

The investigations presented in this article are part of an ECMWF initiative entitled 'Understanding uncertainties in surface–atmosphere exchange' (USURF), which started in November 2017. USURF addresses the very useful feedback about near-surface issues which ECMWF receives from forecasters in the Member and Co-operating States via the Technical Advisory Committee (TAC), Using ECMWF's Forecasts (UEF) meetings, the 'Green Book' on verification, the 'Known Forecast Issues' page, and various other interactions with forecasters and customers. Key to making progress was the development of a conditional verification methodology, which helped to identify specific processes as likely causes of some of the biases. Work so far has focused on 2-metre temperature (T2m) biases in Europe. However, because of the physical links between T2m, 2-metre humidity and 10-metre wind speed, investigations have also included some aspects of humidity- and wind-related processes.

Biases can be removed statistically to some extent when forecasts are passed on to end users. However, there is considerable value to forecast users in having less biased direct model output as it provides a physically more consistent representation of the atmospheric state. Furthermore, a substantial fraction of the 'random' error may result from state-dependent systematic errors which need to be addressed in the forecast model itself.

Two-metre temperature biases in Europe

Routine verification against SYNOP weather station observations shows that T2m biases in the IFS have diurnal and annual cycles (Figure 1) and a pronounced regional dependence. In winter, for example, there is a night-time cold bias of 0.5–1 K in large parts of Europe, and a warm bias of several K throughout the day in parts of Scandinavia (Figure 2). In summer, there is a general underestimation of the amplitude of the diurnal cycle of temperature and a daytime low-humidity bias. Over recent years, there have been some changes in these biases due to model changes, but they are relatively robust in terms of geographical patterns and annual and diurnal variations.

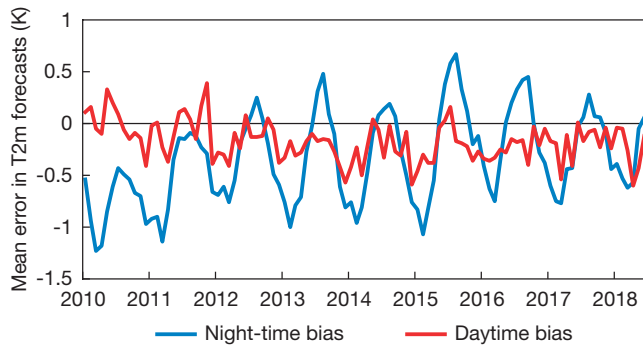


Figure 1 Night-time (00 UTC) and daytime (12 UTC) bias of the ECMWF high-resolution T2m forecast in Europe from verification against SYNOP observations. Stations for which the model elevation differs by more than 100 m from the true elevation, and stations where the nearest grid point is an ocean point have not been included. Lead times are 60 hours and 72 hours, respectively.

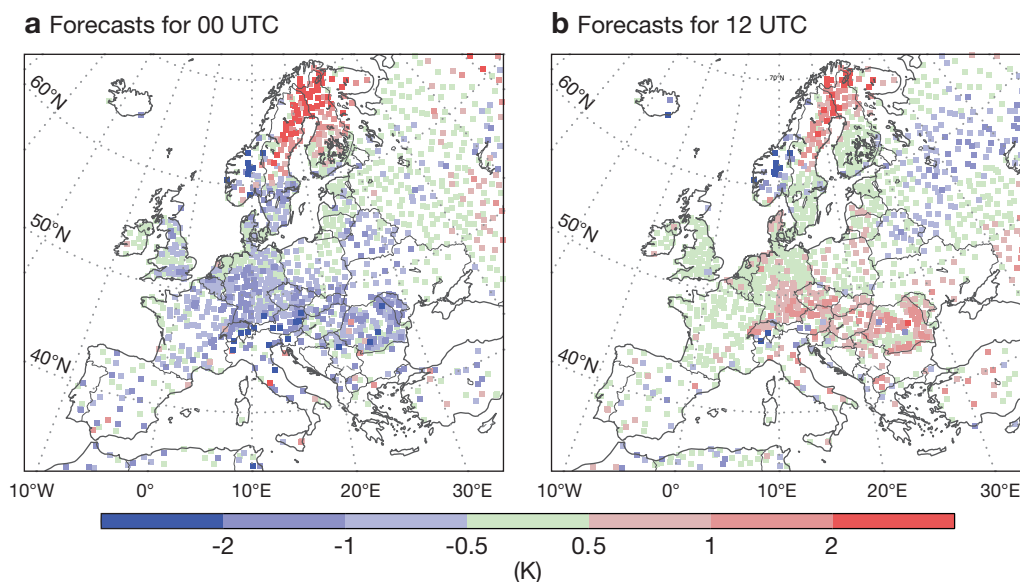


Figure 2 Mean error (bias) of the T2m forecast for day 3 in winter 2017/18 (December–January–February) at (a) 00 UTC and (b) 12 UTC. Verification was performed against the same subset of SYNOP observations as in Figure 1.

The results shown in Figures 1 and 2 are based on a subset of SYNOP stations. They include only those locations where the model orography differs by no more than 100 m from the actual one, and where the nearest model grid point is a land point. This excludes most stations in mountain areas, specifically those on peaks and in small valleys, and many coastal stations. The purpose of this filtering is to focus the verification on larger-scale bias patterns and on areas where the IFS can be expected to represent near-surface weather parameters reasonably well given the limitations imposed by grid resolution.

Night-time cold bias

Conditional verification can be used to quantify relationships between errors in different variables and to disentangle their sources. For example, if we only consider T2m forecasts for days which are (nearly) clear-sky, both in the forecast and SYNOP observations, then the wintertime night-time T2m bias in central Europe is negligible. This suggests that cloudiness plays a role in this bias. In addition to stratifying T2m forecasts according to a quantity like cloudiness, one can also stratify T2m errors according to the forecast error for cloudiness. The left-hand panel of Figure 3 shows that the night-time negative temperature bias in the IFS in central Europe in winter increases roughly linearly with the amount by which total cloud cover is underestimated (against SYNOP observations). However, when weighted by the frequency distribution of the cloud cover errors (shown as green bars in the plot), it turns out that cases where the total cloud cover is underestimated and cases where it is nearly correct contribute about equally to the negative T2m bias (Figure 3b). This indicates that the wintertime negative total cloud cover bias in the IFS over central Europe (on the order of 10% against SYNOP observations, not shown) does not fully explain the negative night-time T2m bias in this region. In cases when the total cloud cover is correctly predicted, the negative T2m bias could be due to other cloud errors, e.g. an underestimation of cloud optical depth, erroneous cloud type or erroneous cloud base height. It could also be due to errors in processes not directly related to clouds, such as vertical mixing or coupling with the surface.

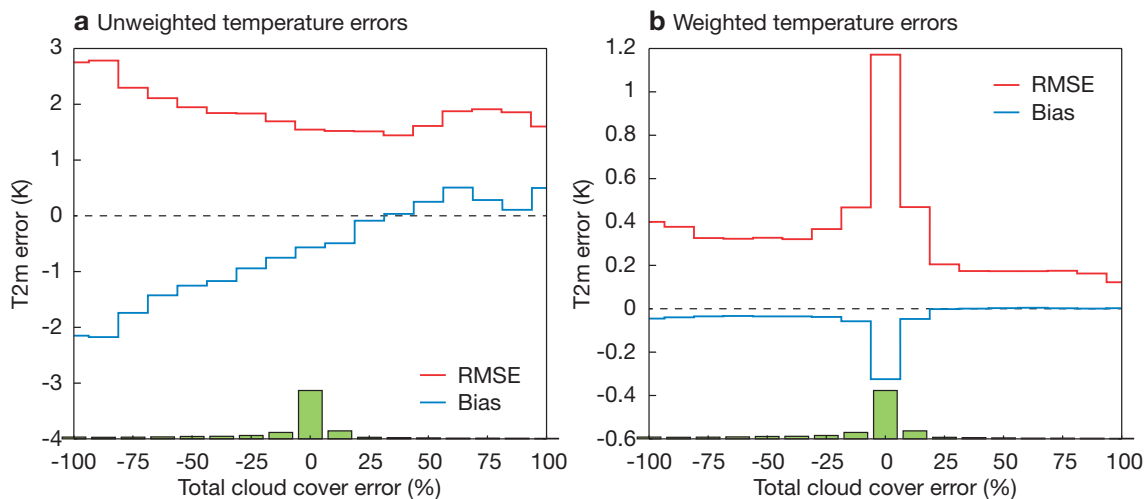


Figure 3 Root mean square error (RMSE) and mean error (bias) for T2m forecasts valid at 00 UTC as a function of the total cloud cover (TCC) error for December–January–February 2016/17 in a central European domain (48–55°N, 0–15°E) at a lead time of 12 hours (a) averaged for each TCC error bin and (b) averaged for each TCC error bin and weighted by the TCC error relative frequency of occurrence. Green bars show the TCC error frequency distribution (arbitrary vertical scale).

To gain further insight into the cloudiness errors, we have also verified downward solar radiation at the surface against both SYNOP observations and a satellite product (Figure 4). Both show an overestimation of downward solar radiation in the order of 5–10 W/m² during wintertime, which corresponds to a relative bias of about 5–10%. These daytime results are consistent with the total cloud cover underestimation against SYNOP observations. The fact that three independent observational datasets indicate very similar forecast biases means that we can have relatively high confidence in these results.

In order to distinguish between different types of cloudy situations, total cloud cover errors can also be stratified according to cloud top height derived from satellite data. Figure 5 shows that a large contribution to the negative cloud cover bias comes from low clouds. Since the satellite identifies the top of the uppermost cloud layer only, the full frequency distribution of cloud top height will be shifted towards lower levels. In central Europe, low stratus with cloud tops typically below 2 km is known to be the main contributor to the negative bias in total cloud cover (Haiden & Trentmann, 2016). Due to recent model upgrades this bias has, however, been reduced significantly and cloud forecast skill has increased accordingly.

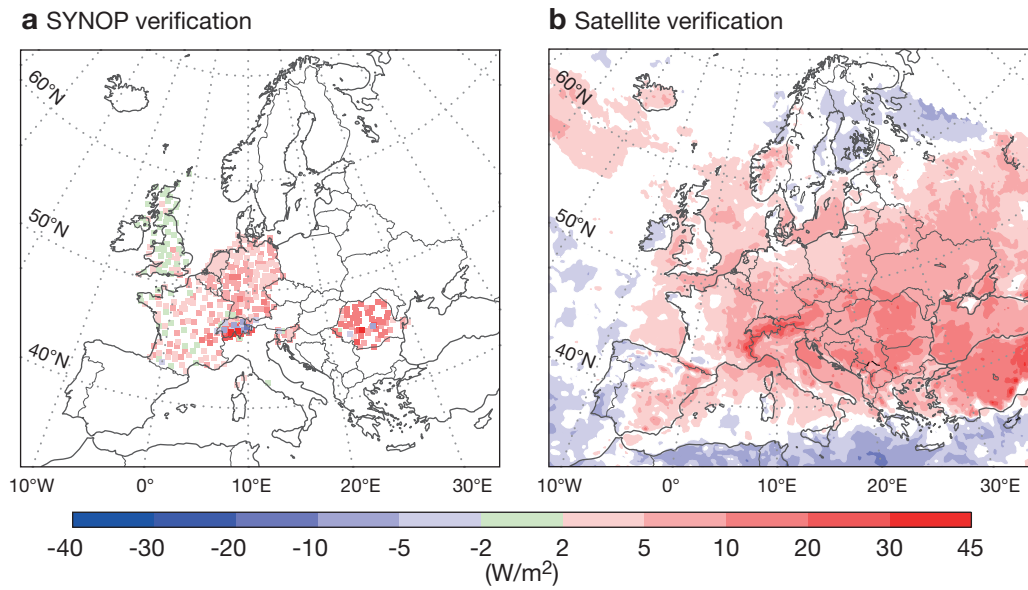


Figure 4 Bias in downward surface solar radiation (24-hour averages) at forecast day 2 in November–December–January 2017/18 from (a) verification against SYNOP and (b) verification against the corresponding satellite product from the Climate Monitoring Satellite Application Facility (CM SAF).

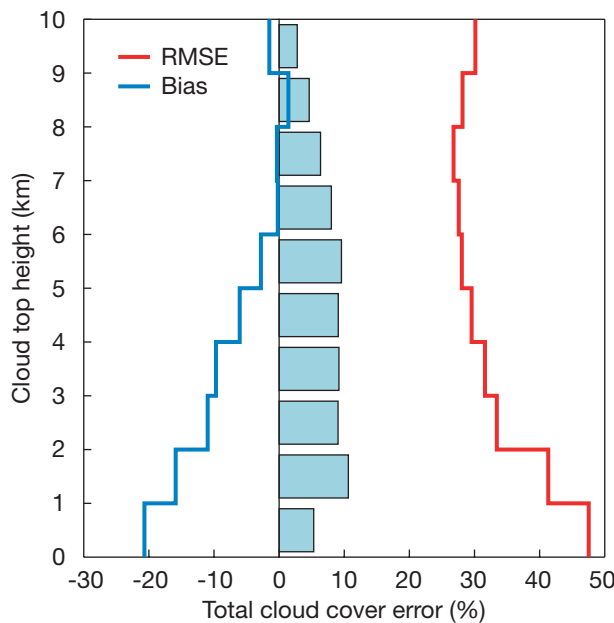


Figure 5 Dependence of short-range wintertime total cloud cover RMSE and mean error (bias) on cloud top height. The bars show the frequency distribution of cloud top heights. The results shown are for central and eastern Europe in December–January–February 2016/17.

Warm bias in Scandinavia

Cloudiness errors are also a factor contributing to the wintertime positive T2m bias in parts of Scandinavia, although the sign of the errors is different: for reasons that are not fully understood, in this region cloudiness tends to be overestimated rather than underestimated. Another factor appears to be the representation of snow cover in the IFS. In clear, calm nights in the real atmosphere, the uppermost layers of the snow cool rapidly, and a strong vertical temperature gradient is established within the snowpack. The skin temperature of the snow drops substantially and T2m decreases accordingly. The single-layer snowpack used operationally in the IFS at present reacts more slowly, due to its larger thermal inertia. The result is a delay in the drop in skin temperature. This delay cannot be fully compensated for by reducing the coupling between the skin temperature and the snow layer because that would lead to an overestimation of the daytime warming of the snow surface. Preliminary results from tests performed with an experimental multi-layer snow scheme

show substantial improvement in the form of 2–3 K stronger night-time T2m drops when conditions are undisturbed (Figure 6). However, some adverse effects on daytime T2m (increase of the warm bias) have been noted which require further study.

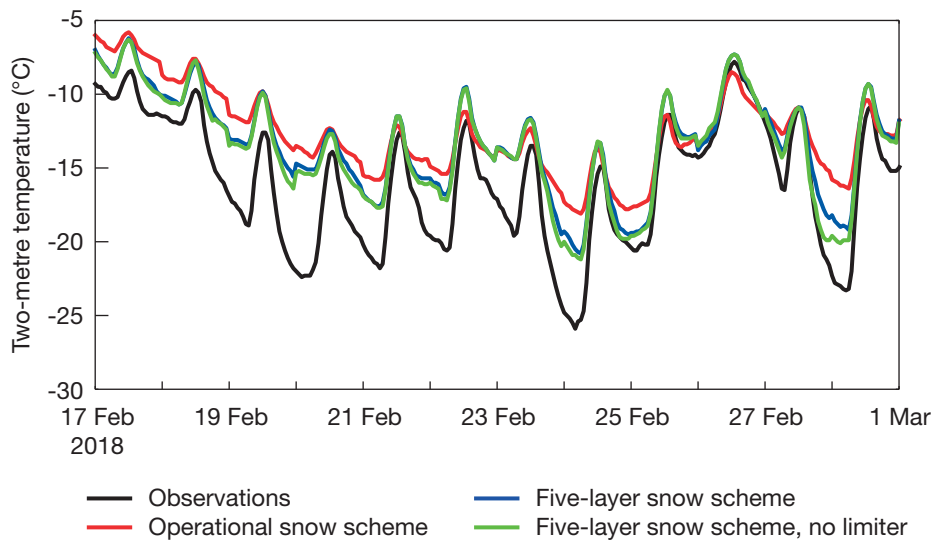


Figure 6 Observed and predicted T2m averaged over northern Scandinavia (64–70°N, 15–30°E). The forecasts are a control experiment with the operational snow scheme and two multi-layer snow scheme experiments in which a five-layer vertical discretization is used. In one of them the T2m limiter is active as in the operational model, in the other the T2m limiter is deactivated. Verification is against SYNOP stations for the period 17 February – 1 March 2018.

Summertime biases

The main systematic T2m forecast error in summer is an underestimation of the diurnal cycle by about 1–2 K, with a cold bias during the day and a warm bias during the night. Cloudiness errors do not appear to play a major role in this case. Another factor that influences interactions between the surface and the atmosphere is surface type, such as soil, vegetation, snow, water or ice. The sub-grid heterogeneity underneath the atmospheric model grid-point is represented by means of a mosaic or tiling that makes it possible to solve the energy balance for each component separately. This is necessary because each land/water element has different properties. Key land surface-related factors controlling the near-surface temperature are soil moisture and soil temperature in a warm climate and snow density and snow/ice temperature in a cold climate. Over land, forests, grassland, and bare soil interact differently with the atmosphere. The interaction also depends on local physiographic conditions, which influence surface drag, aerodynamic resistance, and canopy resistance to evapotranspiration.

Looking first at the night-time warm bias in summer, it is worth noting that night-time cooling in the model is sensitive to the parametrized strength of the thermal coupling between the vegetation canopy and the soil. A reduction of this coupling reduces the night-time bias in summer, but it also leads to stronger night-time cooling in winter, potentially increasing negative T2m biases in central Europe. In IFS Cycle 43r3, implemented in July 2017, the coupling was slightly reduced, which led to a reduction in the night-time warm bias in summer. Reducing it further would have degraded the winter T2m. The night-time warm bias in summer can be reduced further only if the cloud bias in winter can be improved. This example illustrates how a combination of simultaneous changes can be required to reduce near-surface biases without worsening forecast scores.

Turning next to daytime biases in summer, daytime T2m and humidity forecasts for Europe have an overall cold and dry bias. This could be an issue just in the surface layer (up to about 50–100 m) or it could be due to problems in a deeper layer of the atmosphere or both. Figure 7 shows that the model underestimates the difference in temperature and humidity across the surface layer compared to radiosonde observations. This underestimation is particularly pronounced at lower latitudes and contributes to the negative biases there. It means that part of the daytime cool/low humidity bias in summer is likely due to the surface layer in the model being too strongly mixed.

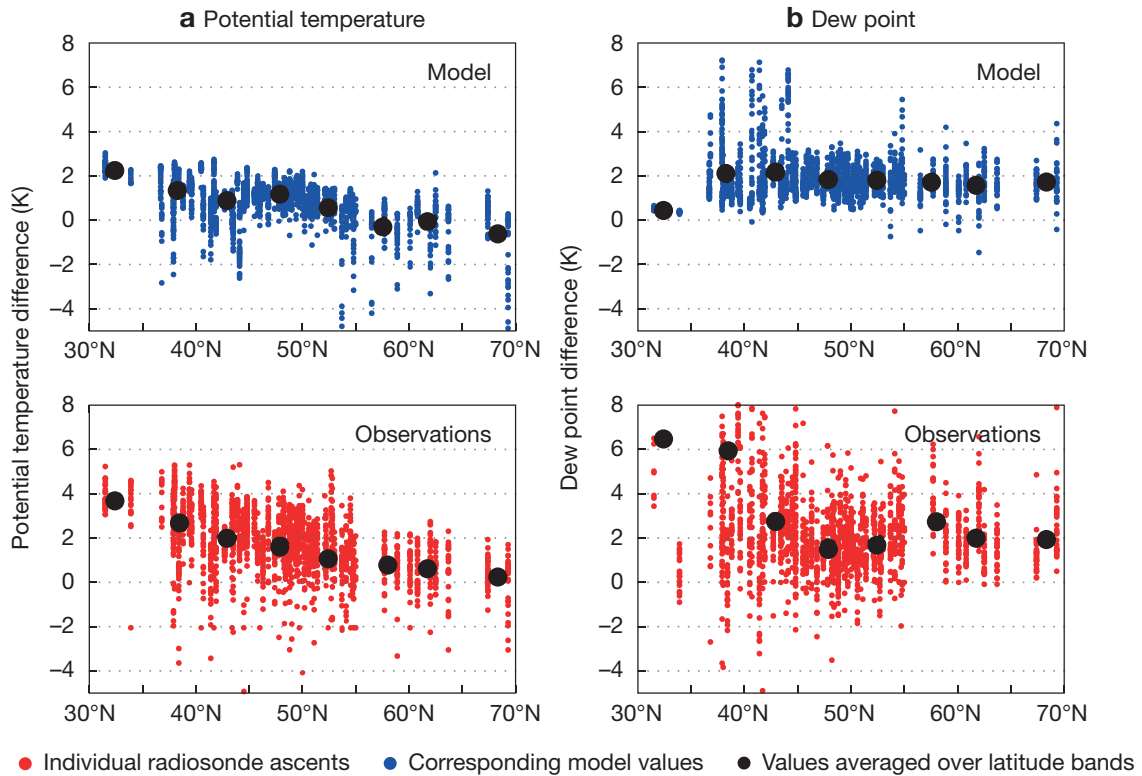


Figure 7 Differences in (a) potential temperature and (b) dew point between the heights of 2 m and 200 m above ground in the ECMWF model and in radiosonde observations as a function of latitude in Europe (10°W–28°E). The verification period is July 2015.

A method that is complementary to conditional verification is to investigate the sensitivity of modelled near-surface weather parameters to changes in boundary-layer physics. Such experiments indicate the extent to which observed biases can be attributed to specific parametrization choices in the turbulent mixing in the boundary layer. Figure 8 shows that summertime T2m is quite insensitive to major changes in the mixing profile, whereas 2-metre dew point does show some sensitivity. Increased vertical mixing reduces the 2-metre dew point by about 0.5 K. Combined with insights from conditional verification of the dew point for clear-sky and cloudy cases, these experiments suggest that the overall low humidity bias during summertime is partly associated with an overestimation of mixing in cloudy cases associated with summertime shallow or deep convection. If only clear-sky cases are evaluated, the forecast has a slight moist bias.

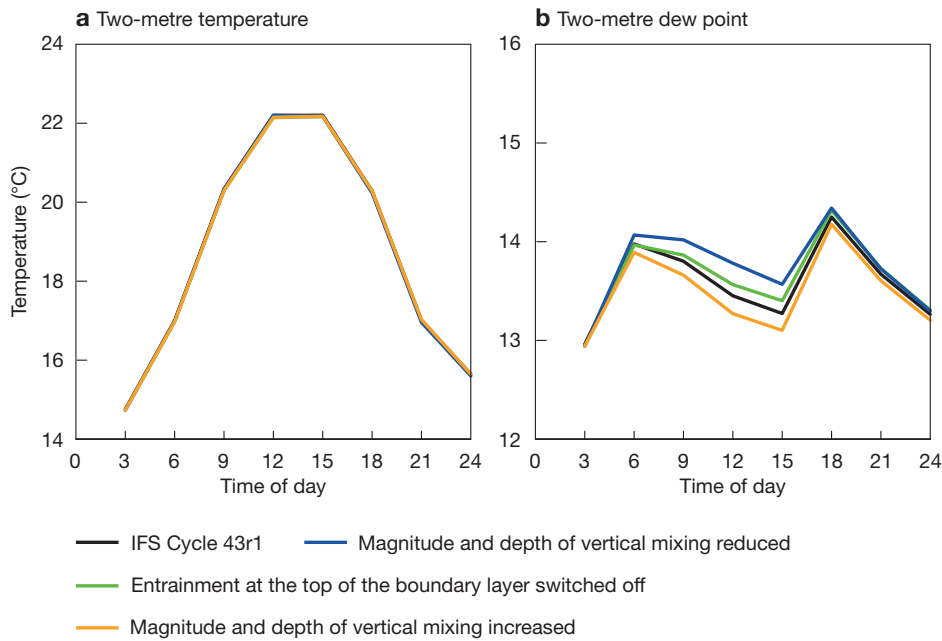


Figure 8 Effect of different degrees of mixing across the daytime planetary boundary layer on (a) T2m forecasts and (b) 2-metre dew point forecasts in a central European domain as a function of the time of day. Results for increased strengths of turbulent vertical mixing show negligible sensitivity for temperature but a noticeable effect on dew point. Black lines correspond to IFS Cycle 43r1, operational from November 2016 to July 2017. The forecasts are short-range forecasts at 25 km resolution, aggregated over the month of July 2016.

Summary of findings

Near-surface weather parameters such as T2m are governed by a range of processes, such as advection, boundary-layer turbulent mixing, the strength of the land–atmosphere coupling, radiation fluxes, the state of the soil and vegetation, and the presence of snow or orography. The large number of factors involved complicates forecast error attribution. Significant progress has recently been made by using conditional verification and by running sensitivity experiments to explore the impact of parametrization changes on near-surface parameters. The main findings of the ongoing ECMWF project focusing on these biases are that (a) biases are easier to address if one focuses first on non-coastal stations outside major mountain areas; (b) the night-time cold bias in most of Europe in winter is partly related to an underestimation of the cloud cover, but some of it is present even when the cloud cover is correct; (c) the warm bias in Scandinavia in winter is partly due to the use of a single layer in the snow scheme; (d) the underestimation of near-surface temperature and humidity in summer over land is at least partly due to an insufficient temperature and dew-point gradient in the lowest 200 m; (e) daytime near-surface temperature in the model is resilient to changes in atmospheric mixing, while humidity is moderately sensitive to atmospheric mixing; and (f) the low humidity bias in summer appears to be mostly related to an overestimation of turbulent mixing in cloudy boundary layers.

Further work

One of the next steps will be to perform a more in-depth verification against datasets from meteorological masts, such as the Lindenberg site mast (run by the German national meteorological service, DWD), which is now available to ECMWF in near-real time, and the Cabauw mast (run by the Dutch national meteorological service, KNMI). This will show to what extent biases in near-surface temperature and dew point are representative of biases over a deeper layer and how this changes with the time of day and with season. It will also make it possible to concurrently examine errors in temperature in the lowest 100 m of the atmosphere and in the soil, as well as errors in the surface energy budget. It is hoped that this will help to further pin down the cause of biases in the operational forecast. The reasons for the different kinds of cloud errors found in forecasts for Scandinavia and central Europe will also be investigated, notably using data from the Sodankylä mast in Finland.

Work towards the operational implementation of a multi-layer snow scheme will continue. The scheme will be calibrated on in-situ measurements within the ESM-snowMIP experiment to optimise the underlying parameters so that the observed snow depth and density are reproduced. Evaluation of the scheme will particularly focus on its ability to reproduce the observed near-surface temperature amplitude of the diurnal cycle.

A simpler framework for moist processes is being developed, relying on consistent assumptions and improved coupling between the turbulent diffusion, convection and cloud schemes and the dynamics. This work, together with planned improvements to the representation of warm- and cold-phase microphysical processes, should help to further reduce systematic errors in cloudiness and precipitation and thereby reduce biases in near-surface weather parameters.

Further reading

Haiden, T. & J. Trentmann, 2016: Verification of cloudiness and radiation forecasts in the greater Alpine region. *Meteorol. Z.*, **25**, 3–15.

© Copyright 2018

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England

The content of this Newsletter is available for use under a Creative Commons Attribution-Non-Commercial-No-Derivatives-4.0-Unported Licence. See the terms at <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.