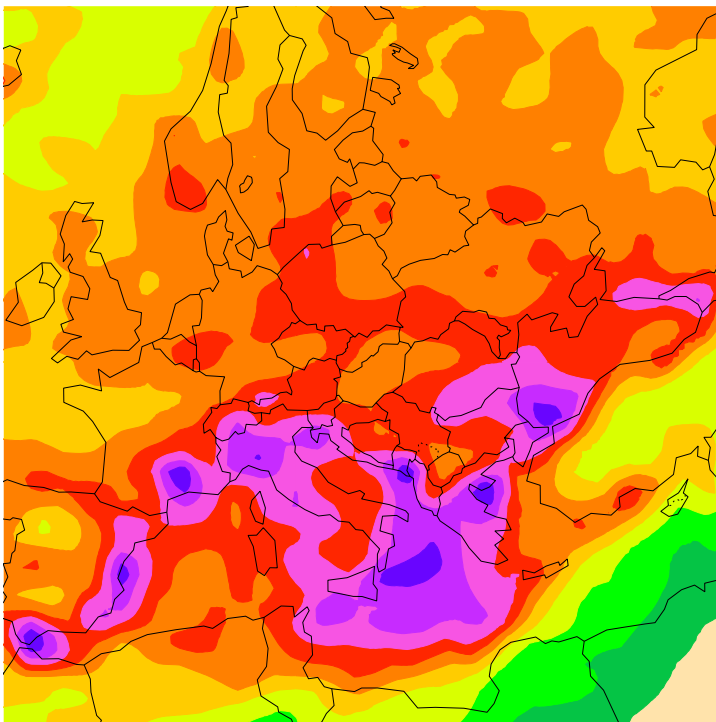




**EARTH SYSTEM SCIENCE**

Improved two-metre  
temperature forecasts in the  
2024 upgrade



This article appeared in the Earth system science section of ECMWF Newsletter No. 178 – Winter 2023/24, pp. 24–29.

# Improved two-metre temperature forecasts in the 2024 upgrade

Bruce Ingleby, Gabriele Arduini, Gianpaolo Balsamo, Souhail Boussetta, Kenta Ochi (Japan Meteorological Agency), Ewan Pinnington, Patricia de Rosnay

Two-metre temperature (T2m) is a key forecast variable. Here we describe how changes to ECMWF's Integrated Forecasting System (IFS), which are expected to become operational as part of IFS Cycle 49r1 in 2024, improve short-range forecasts of T2m. Physics changes in IFS Cycle 49r1 will include extensive improvements to surface vegetation fields and changes to make the interpolation of the temperature to 2 m more realistic. There have also been major upgrades to the data assimilation system, which establishes the initial conditions of forecasts (the 'analysis'). Currently, T2m ('screen temperature') and two-metre humidity ('screen humidity') from SYNOP weather station reports and METAR aviation weather reports are assimilated in a separate land surface analysis – primarily to update initial soil moisture and temperature conditions. However, only daytime screen humidity is assimilated in the atmospheric component of the data assimilation system, 4D-Var. In IFS Cycle 49r1, T2m will be assimilated in 4D-Var. This proved beneficial after extensive testing/tuning. Improvements in the separate snow data assimilation system as well as in T2m and soil moisture land data assimilation are also applied. Overall this results in better T2m forecasts, especially in northern hemisphere winter.

## New vegetation maps and soil moisture stress representation

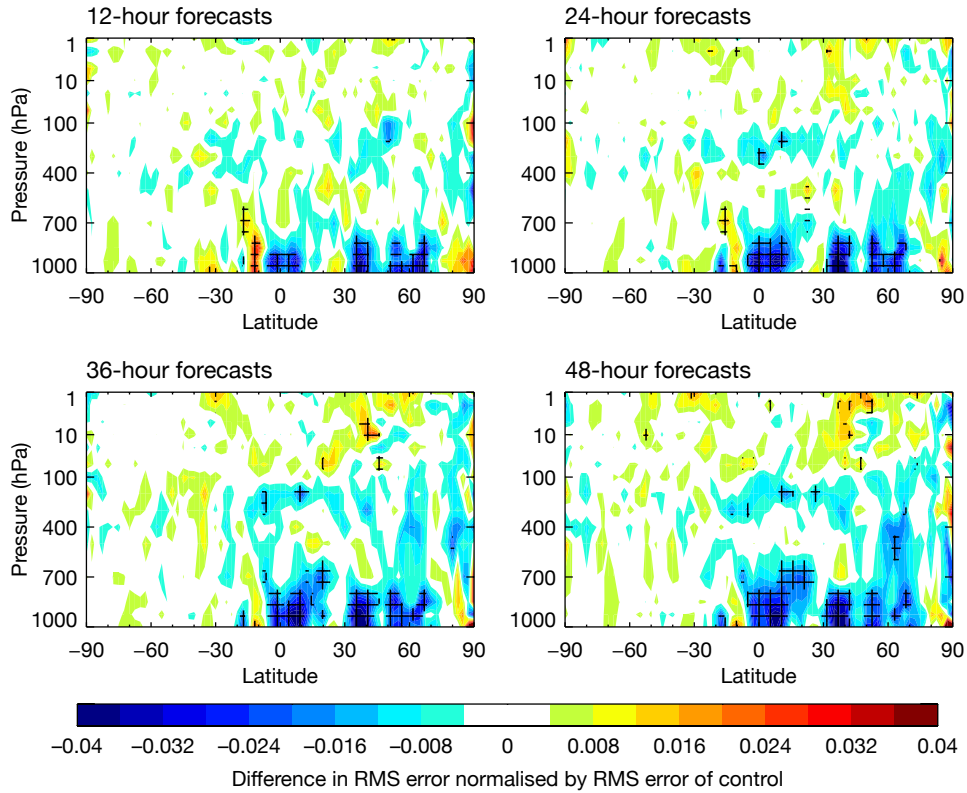
The new land-use and land-cover (LULC) maps of the European Space Agency (ESA) Climate Change Initiative and the new Copernicus Global Land Services-based leaf area index (LAI) maps have been successfully tested and included in Cycle 49r1. As well as using the new maps, several model changes were made towards a more realistic physical representation. The first change is the implementation of a new LAI observation operator, which enables a conservative disaggregation of the observed LAI into high and low vegetation components. This resulted in a more accurate seasonal variability of vegetation. The second step is the implementation of a more realistic soil moisture stress function. The updated stress function mimics the behaviour of the soil matric potential, a measure of the relative availability of water in the soil for plant uptake, and allows plant roots to better extract water, especially in relatively dry conditions. The third change affects mainly winter and spring seasons by explicitly allocating the surface albedo for forests with snow underneath from processed satellite data rather than relying on tabulated values.

With these changes, the near-surface atmospheric variables still needed additional adjustment for acceptable NWP skill. Model parameters were first optimised for individual grid points based on in-situ observations (FLUXNET) and in surface standalone mode. This enabled a better match of the surface fluxes with in-situ observations. The second level of model parameter adjustments uses global offline surface simulations for which the near-surface parameter errors are stratified by vegetation type. The minimum stomatal resistance (a plant characteristic that represents the degree of opposition of the plant leaf to release water vapour to the atmosphere) is used to adjust T2m and relative humidity, mainly for spring and summer. In Figure 1, the improvements brought by the land-use and land-cover package and associated model changes are illustrated for temperature cross-sections for different latitudes and pressure levels at several forecast ranges.

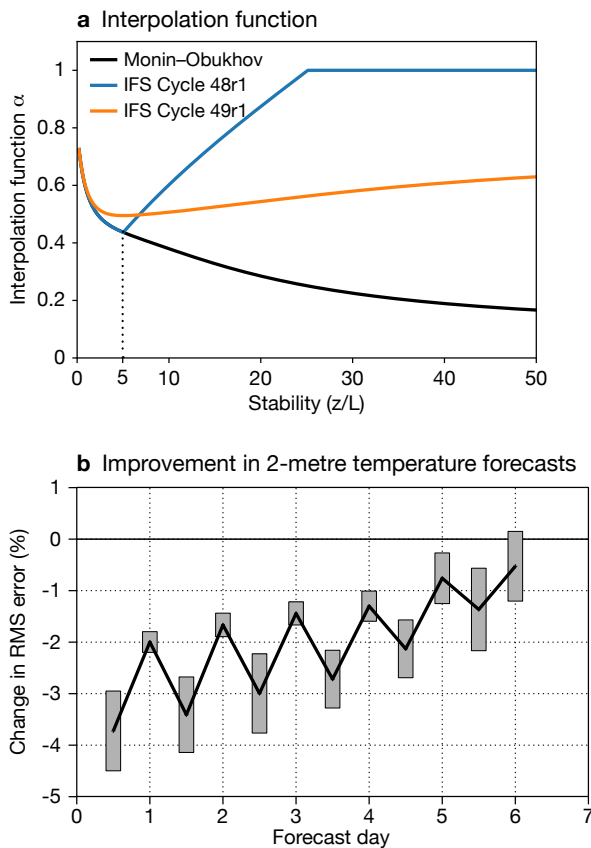
## Modified 2-metre temperature diagnostic formula

Two-metre temperature is a diagnostic variable in the IFS, which means that it is not predicted directly by the model but is derived from other variables. This involves a vertical interpolation between surface (skin) temperature and the lowest model level temperature using Monin-Obukhov similarity theory. To avoid possible runaway cooling, the interpolation function used in Cycle 48r1 and earlier was tapered to the lowest model level temperature for relatively high stability (see Figure 2a). In Cycle 49r1, a modified interpolation function is introduced that is closer to the theory and smoother, avoiding the abrupt jump to the lowest model level temperature at a fixed value of stability.

The introduction of the revised diagnostic formula shows a significant improvement in root-mean-square (RMS) error in the northern hemisphere at short lead times (see Figure 2b), due to a more realistic evolution of T2m in stably stratified conditions.



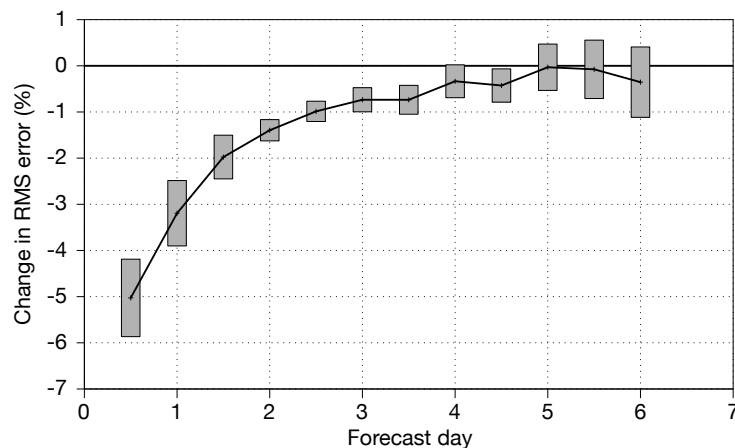
**Figure 1** Temperature forecast improvements related to the vegetation changes and associated modifications are shown in the blue areas, which correspond to a reduction in the root-mean-square (RMS) error of the temperature forecasts. The tests were carried out from 5 June 2020 to 31 August 2020. Forecasts were verified against each experiment’s own analysis, and cross-hatching indicates statistical significance at the 95% confidence level.



**Figure 2** We show (a) the interpolation function,  $\alpha$ , as a function of stability, which is defined as height above ground/Obukhov length ( $z/L$ ).  $\alpha=1$  indicates that T2m equals T at the lowest model level;  $\alpha=0$  indicates that T2m equals the skin temperature. The black line shows the function based on Monin-Obukhov theory, the blue line indicates the current interpolation formula (Cycle 48r1), and the orange line indicates the new function to be introduced in Cycle 49r1. We also show (b) the improvement (negative) in T2m forecasts due to changes in the interpolation function of T2m. This relates to forecasts for December 2021 to February 2022, for the region 20°–90°N. We show the percentage change in root-mean-square (RMS) error, verified against SYNOP observations. The grey rectangles show 95% confidence intervals.

### Atmospheric assimilation of 2-metre temperature and other surface variables

For many years, assimilation of SYNOP temperature and humidity data was standard in limited-area models, but not in global systems. In 2008, the UK Met Office global model started to assimilate these data, and in 2022 the German weather service (DWD) followed suit. At ECMWF, tests assimilating T2m in 4D-Var gave large improvements in short-range T2m forecasts, but in some winter regions 850 hPa temperatures were degraded, partly due to unrealistic coupling in stable conditions. Remedies aimed at retaining the T2m benefits, while reducing the problems, were explored. Reducing the weight given to T2m data with very large departures from the short-range forecast used to establish the analysis (the 'background'), and limiting the assimilation of data to the first six hours of the 12-hour 4D-Var window to produce more localised increments, both helped. Temperature differences of more than 7.5 K are not used. Following these modifications, a large benefit is still seen in northern hemisphere winter (Figure 3): a 5% reduction in RMS error for 12-hour forecasts, decreasing with lead-time when verified against SYNOP data. The benefit in summer and in the tropics is smaller, but still useful.

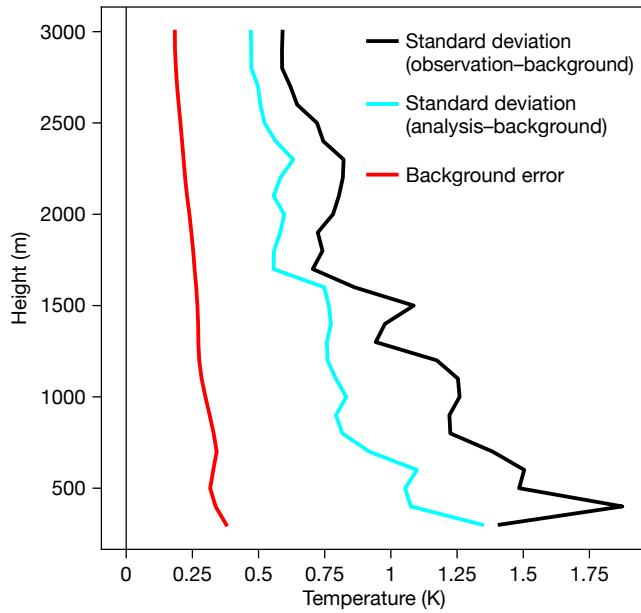


**Figure 3** The improvement (negative) in T2m forecasts due to 4D-Var assimilation of T2m. The plot shows results for the period of December 2021 to February 2022, for the region 20°–90°N. We show the percentage change in root-mean-square (RMS) error, verified against SYNOP observations. The grey rectangles show 95% confidence intervals.

On average, temperature decreases with height in the troposphere – this is known as the lapse rate. To take account of the difference between station height and model height in 4D-Var, an adjustment is calculated from the height difference using a lapse rate of 5.5 K/km and applied to T2m. This fits the data slightly better than 6.5 K/km, which is often used, including in verification. Stations are used between 400 m below and 200 m above the model height. On average, stations are slightly lower than the model height as they are more likely to be in valleys. There is no bias correction applied, because of the complexity of observation–background T2m biases, and in many cases the background biases are larger than observation biases. The verification shown here uses SYNOP stations within 250 m of the model height, and observations more than 15 K from the analysis are rejected. The assimilated screen humidity variable changes from relative humidity to specific humidity, and it will be used both day and night after the change in Cycle 49r1.

Work at the UK Met Office and elsewhere (see Pauley and Ingleby, 2022) suggests that it is difficult to get a positive impact from assimilating 10 m wind over land. SYNOP wind assimilation was tested with the other changes, but the impact was marginal and the winds will not be assimilated at ECMWF for now. Winds from ships and moored buoys continue to be assimilated. Surface pressure remains the most important surface variable for global NWP.

The background errors used in 4D-Var are derived using the ensemble of data assimilations (EDA). In common with other ensemble systems, the EDA has long been underspread near the surface (especially over snow). There is a change to the stochastic physics in Cycle 49r1 that partially addresses this, but as shown in Figure 4, the background error estimate still appears too small in the lowest levels. Larger spread near the surface would tend to increase the size of analysis increments from T2m assimilation at the lowest model level and to reduce them a few levels up, which would be more realistic for winter cases.



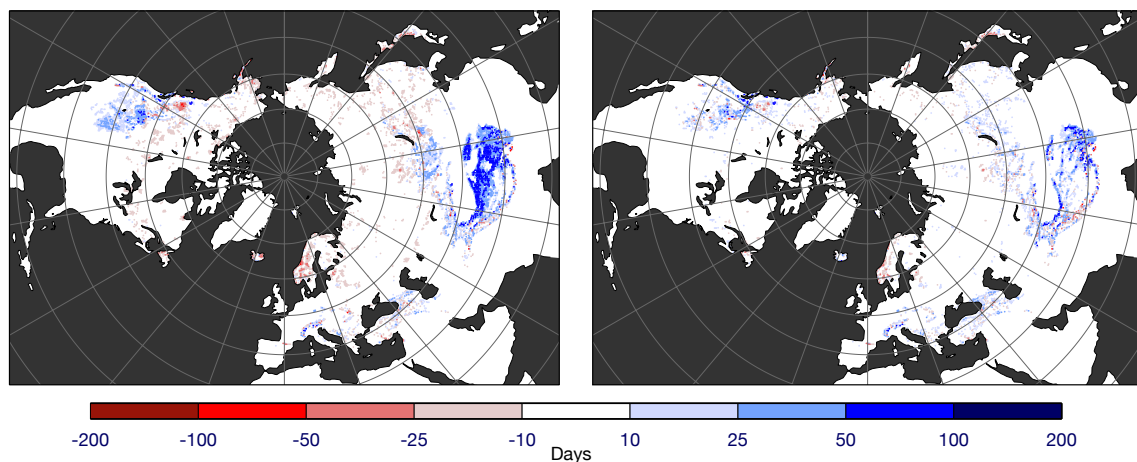
**Figure 4** Low-level temperature statistics from the radiosonde station at Sodankylä, Finland, for December 2022. The standard deviations of observation-minus-background (O-B) and analysis-minus-background (A-B) are shown, as is the mean estimate of background error based on the EDA.

**Snow data assimilation and improved snow cover**

Several major developments were made to improve the snow analysis. This uses both synoptic snow depth reports and the US National Oceanic and Atmospheric Administration’s National Environment Satellite, Data, and Information Service (NESDIS) Interactive Multi-sensor Snow and Ice Mapping System (IMS). The use of IMS data has been revised, enabling the exploitation of satellite-based snow cover information in mountainous areas. It reduces IFS positive biases in snow cover and snow depth, with a strong positive impact on atmospheric forecasts. Other changes include an increase in the maximum allowed snow depth value from 1.4 m to 3 m in the analysis, and a reduction of the vertical structure function scales from 800 m to 500 m. The latter change increases the weight of observations that have the lowest altitude difference with the model. In addition to data assimilation changes, snow cover parametrization in the snow model was also revised to improve the fraction of ground covered by snow. Both the snow changes were extensively evaluated for winter and summer as well as spring, which is key for snow model and data assimilation (see Figure 5). Other results (not shown) indicate a strong positive impact on T2m, especially in the Rocky Mountains and the Tibetan Plateau areas, with widespread improvements in the troposphere in geopotential height, vector wind and humidity forecasts in the northern hemisphere in winter.

**a** Snow duration bias – old

**b** Snow duration bias – new

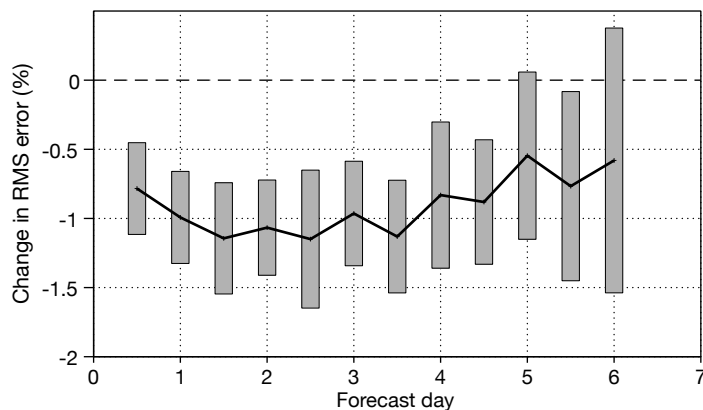


**Figure 5** Biases of snow duration (number of days when snow depth is more than 1 cm in the model) compared to IMS in March–May 2022 (a) before and (b) after the snow data assimilation changes. Positive biases are reduced in mountainous areas.

In the last decade, the World Meteorological Organization (WMO) Snow Watch team has worked to improve the exchange of in-situ snow observations and to assess the quality of satellite snow products. The IFS Cycle 48r1 upgrade, implemented in June 2023, included a change from a single layer snow model to a multi-layer snow model, which makes snow melt more realistic (see Newsletter No. 176: <https://www.ecmwf.int/en/newsletter/176/>).

### Soil moisture and 2-metre temperature land assimilation

The soil moisture analysis relies on a Simplified Extended Kalman Filter (SEKF). We assimilate soil moisture products from ASCAT (Advanced Scatterometer) and SMOS (Soil Moisture and Ocean Salinity) satellite instruments along with screen level T2m and relative humidity information. There is a linked 2D Optimal Interpolation analysis of T2m (de Rosnay et al., 2022). These components have been revised, bringing improvements in both winter and summer forecasts, in T2m and near-surface forecasts. The revisions include the implementation of a lapse rate correction in the T2m analysis to account and correct for altitude differences between the model grid points and the SYNOP observations. An adjustment of 5.5 K/km is introduced, consistent with the lapse rate correction used in the 4D-Var T2m analysis. In the SEKF soil analysis, the background error standard deviation has been doubled, to 0.02 m<sup>3</sup> m<sup>-3</sup>, giving more weight to the observations used in the SEKF to constrain the soil moisture analysis. Both changes have a substantial positive impact in the tropics and northern hemisphere during summer. Figure 6 shows the effect of these changes, plus the snow changes described above, on the forecast fit to SYNOP temperature in winter: an improvement of about 1% approximately, constant out to day six. The improvement of the forecast fit to our own analysis is larger (not shown).



**Figure 6** Impact of snow, soil moisture and lapse rate changes on verification of predicted temperature against SYNOP temperature 20°–90°N, December 2021 – February 2022. We show the percentage change in root-mean-square (RMS) error. The grey rectangles show 95% confidence intervals.

### Combined impact

Figure 7 shows the scorecard against observations for the changes described above. The results are mostly very positive. One of the few exceptions to this is that, in the tropics at 850 hPa, there is a detriment to short-range temperature forecasts. This stems from the interaction of the LULC and T2m assimilation changes: before tuning the assimilation change, the problem was more widespread. The land-surface changes also give slight benefits beyond day 10 (not shown). Figures 8 and 9 show verification of T2m for the whole Cycle 49r1 against SYNOP observations for northern hemisphere winter. Most benefit is seen north of 45°N. Figures 7–9 illustrate the widespread positive impact of the changes on T2m and atmospheric forecasts. Another change included in Cycle 49r1 is the assimilation of integrated water vapour information from ground-based GPS stations (predominantly over Europe and China for now). This has little impact on T2m forecasts, but it slightly improves short-range screen humidity forecasts. In the northern extratropical winter, the headline result is that Cycle 49r1 improves the 12-hour T2m forecast by more than 9% (Figure 9).

The results presented here all used forecast/analysis grid spacings of 28/80 km. The whole Cycle 49r1 package is now being tested at grid spacings of 9/40 km. The results confirm the previous tests.

Parameter	Level (hPa)	Northern hemisphere		Southern hemisphere		Tropics	Europe	North America	East Asia	Australia, New Zealand	Arctic	Antarctic
		RMS error		RMS error		RMS error	RMS error	RMS error	RMS error	RMS error	RMS error	RMS error
		Forecast day		Forecast day		Forecast day	Forecast day	Forecast day	Forecast day	Forecast day	Forecast day	Forecast day
Geopotential	100											
	250	▲	▲		▲			▲	▲		▲	▲
	500	▲	▲				▲	▲	▲		▲	▲
	850	▲	▲				▲	▲	▲		▲	▲
Temperature	100											
	250		▲					▲	▲			
	500	▲	▲				▲	▲	▲		▲	▲
	850	▲	▲				▲	▲	▲		▲	▲
2 m temperature		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
Vector wind	100											
	250		▲		▲			▲	▲		▲	▲
	500		▲		▲		▲	▲	▲		▲	▲
	850		▲		▲		▲	▲	▲		▲	▲
10 m wind speed		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
Relative humidity	250											
700		▲										
2 m dew point		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
Total cloud cover		▲	▲			▲		▲	▲		▲	▲
Total precipitation				▼	▼	▼		▼	▼		▼	▼
Significant wave height												

Symbol legend: for a given forecast step...

- ▲ experiment better than control statistically significant with 99.7% confidence
- △ experiment better than control statistically significant with 95% confidence
- experiment better than control statistically significant with 50% confidence
- not really any difference between control and experiment
- experiment worse than control statistically significant with 50% confidence
- ▼ experiment worse than control statistically significant with 95% confidence
- ▼ experiment worse than control statistically significant with 99.7% confidence

Figure 7 Scorecard of root-mean-square (RMS) error with respect to observations for winter and summer combined (six months) for the changes described in the text.

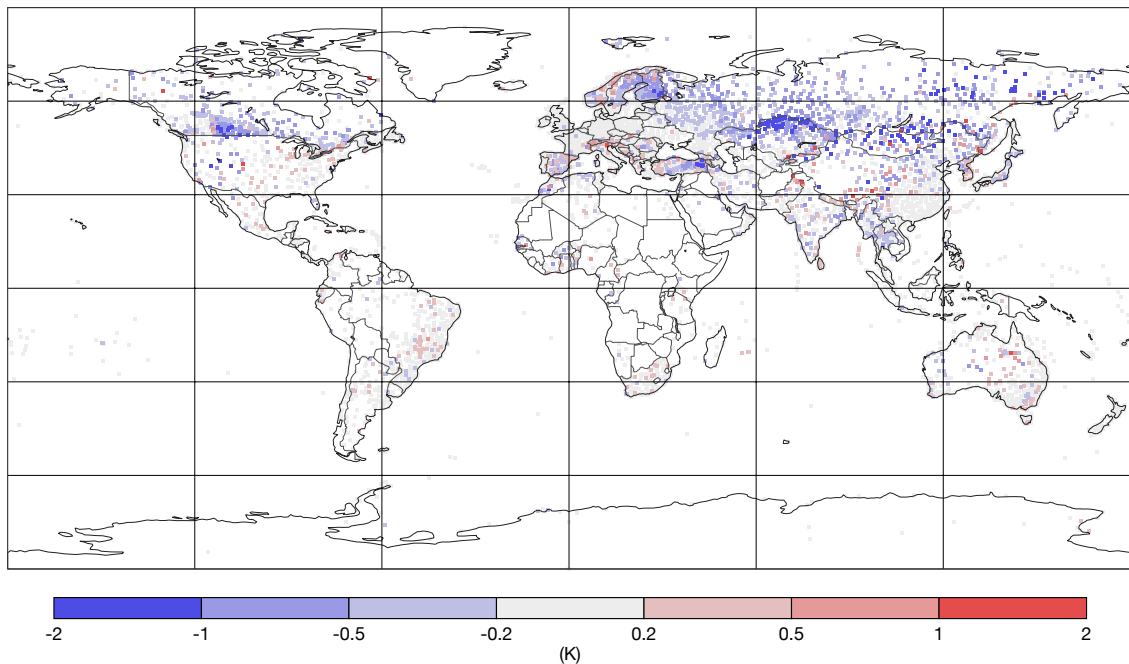
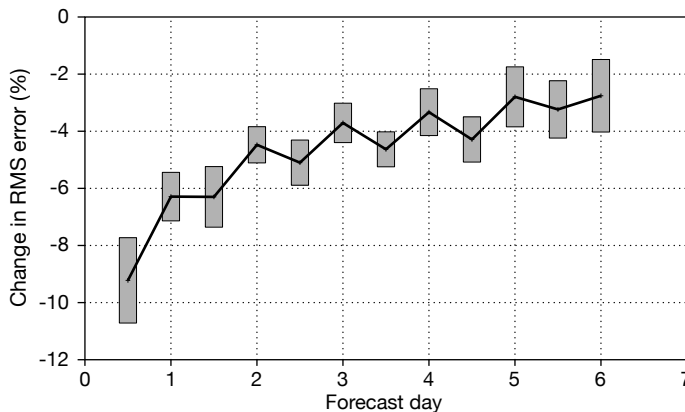


Figure 8 Impact on 12-hour verification of T2m against SYNOP observations of the changes to be implemented in Cycle 49r1 from December 2021 to February 2022 at 12 UTC. Blue colours indicate an improvement.



**Figure 9** The impact of Cycle 49r1 on T2m forecasts verified against SYNOP observations (percentage change in root-mean-square error) for December 2021 to February 2022, 20°–90°N. The grey rectangles show 95% confidence intervals.

## Outlook

ECMWF will continue to work on improving T2m forecasts. One aspect is increasing the ensemble spread of near-surface temperatures, especially in winter, as mentioned above. 4D-Var has several outer loops, starting at coarse resolution and then moving to finer scale. We will try putting more weight on the T2m observations in the later outer loops, as the model representation of near-surface temperature improves. Above the boundary layer (i.e. above the part of the atmosphere which is directly influenced by the surface), there is less sensitivity to model resolution. In recent years, partly due to EU policy, some European countries have started providing data from more stations, and this trend looks set to continue. There is separate work evaluating crowd-sourced observations. In a global NWP system, we are well aware that some parts of the world are sparsely observed. With the WMO and others, efforts to improve data exchange and quality will continue. Also, ongoing developments of coupled land–atmosphere assimilation at ECMWF will enhance the exploitation of surface-sensitive satellite observations and contribute to improving T2m forecasts.

On the land-modelling side, several new developments aim to improve the surface representation and coupling with the atmosphere. These are facilitated by a new climate field generation suite, which enables timely testing of new global vegetation and land input fields; and the prospective Multi-Parameter Regionalisation (MPR) system, which is intended to enable an automated physically-based optimisation of model parameters. Specifically, the T2m forecast would benefit from the improvement of the water, energy and carbon cycles representation. Similar optimisation has already shown benefit for T2m in the DWD system. Future work will also focus on improved vertical soil discretisation and coupling with the hydrological model. It would also benefit from a better representation of snow intercepted by vegetation and improved wind-induced water flux from snow to the atmosphere (sublimation). Over glaciers, a new sub-grid ice representation with a better coupling between snow and ice would also improve the near-surface atmosphere. Finally, anthropogenic impacts are being explored under the very-high-resolution framework of the EU’s Destination Earth (DestinE) initiative, in which ECMWF is one of the partners.

## Further reading

**Pauley, P. & B. Ingleby**, 2022: Assimilation of in-situ observations. In: Park, S.K. & L. Xu (eds.) *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, Vol. IV, Springer.

**de Rosnay, P., P. Browne, E. de Boissésou, D. Fairbairn, Y. Hirahara, D. Schepers et al.**, 2022: Coupled data assimilation at ECMWF: current status, challenges and future developments, *Q. J. R. Meteorol. Soc.*, **148**(747), 2672–2702.

**Boussetta, S., G. Balsamo, G. Arduini, E. Dutra, J. McNorton, M. Choulga et al.**, 2021: ECLand: The ECMWF Land Surface Modelling System, *Atmosphere*, **12**, 723. <https://doi.org/10.3390/atmos12060723>

**Arduini, G., G. Balsamo, E. Dutra, J.J. Day, I. Sandu, S. Boussetta et al.**, 2019: Impact of a multi-layer snow scheme on near-surface weather forecasts, *Journal of Advances in Modeling Earth Systems*, **11**, 4687–4710. <https://doi.org/10.1029/2019MS001725>



© Copyright 2024

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

The content of this document, excluding images representing individuals, is available for use under a Creative Commons Attribution 4.0 International Public License. See the terms at <https://creativecommons.org/licenses/by/4.0/>. To request permission to use images representing individuals, please contact [pressoffice@ecmwf.int](mailto:pressoffice@ecmwf.int).

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.