# How can 2m temperature forecasts possibly go wrong? – an ECMWF perspective

**Tim Hewson**

Principal Scientist, Evaluation Section

ttm.hewson@ecmwf.int

Thanks to the ECMWF Daily Report Team ☺

*One might imagine that forecasting 2m temperatures would be quite easy…*

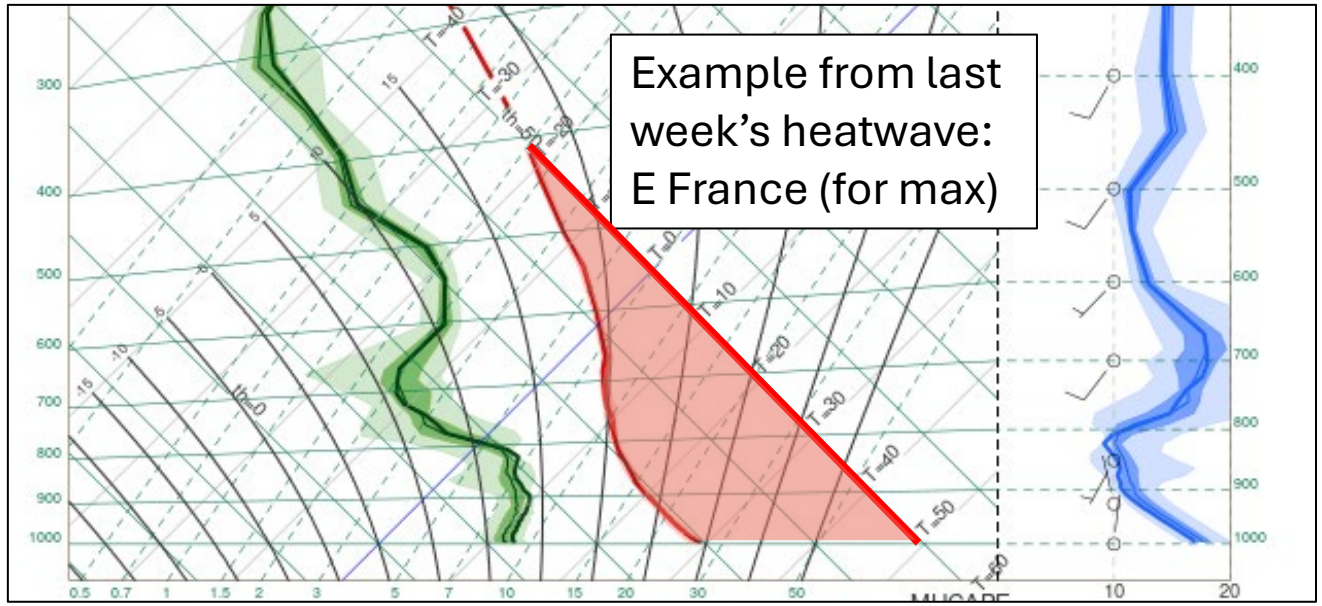
# Layout

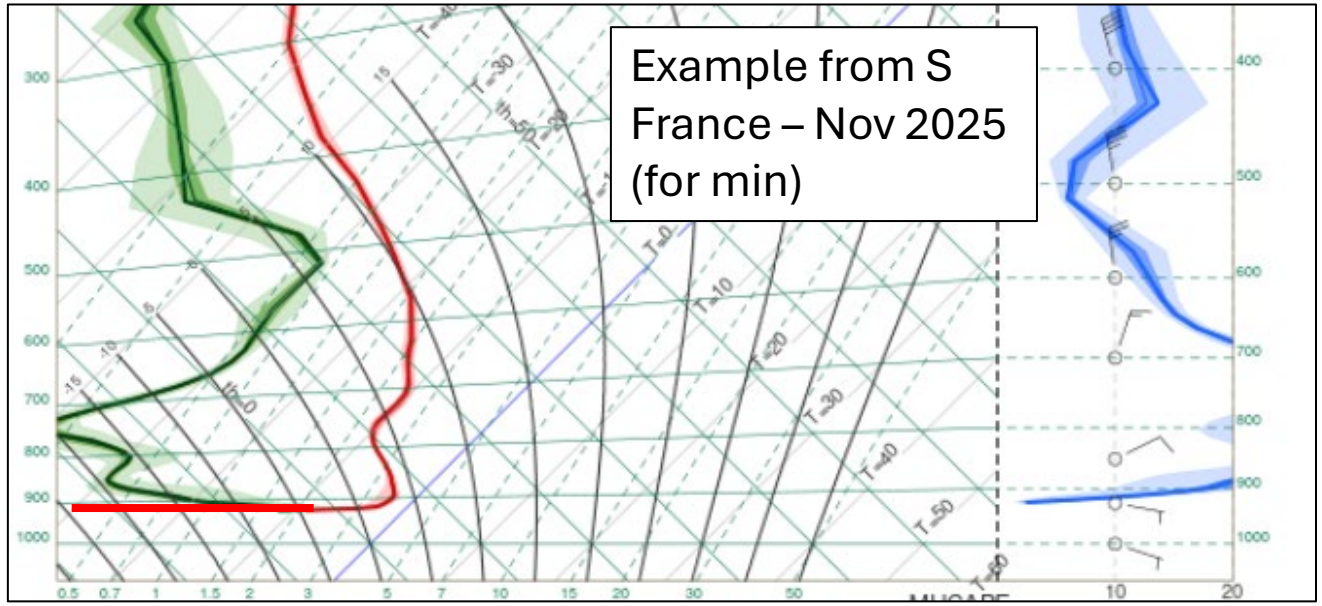
1. Some physical considerations
2. The challenge of extreme low temperatures
3. When the IFS goes wrong – past examples
4. A recent example from the AIFS
5. Concluding remarks

# 1. Some Physical Considerations

Cold versus Warm

First order



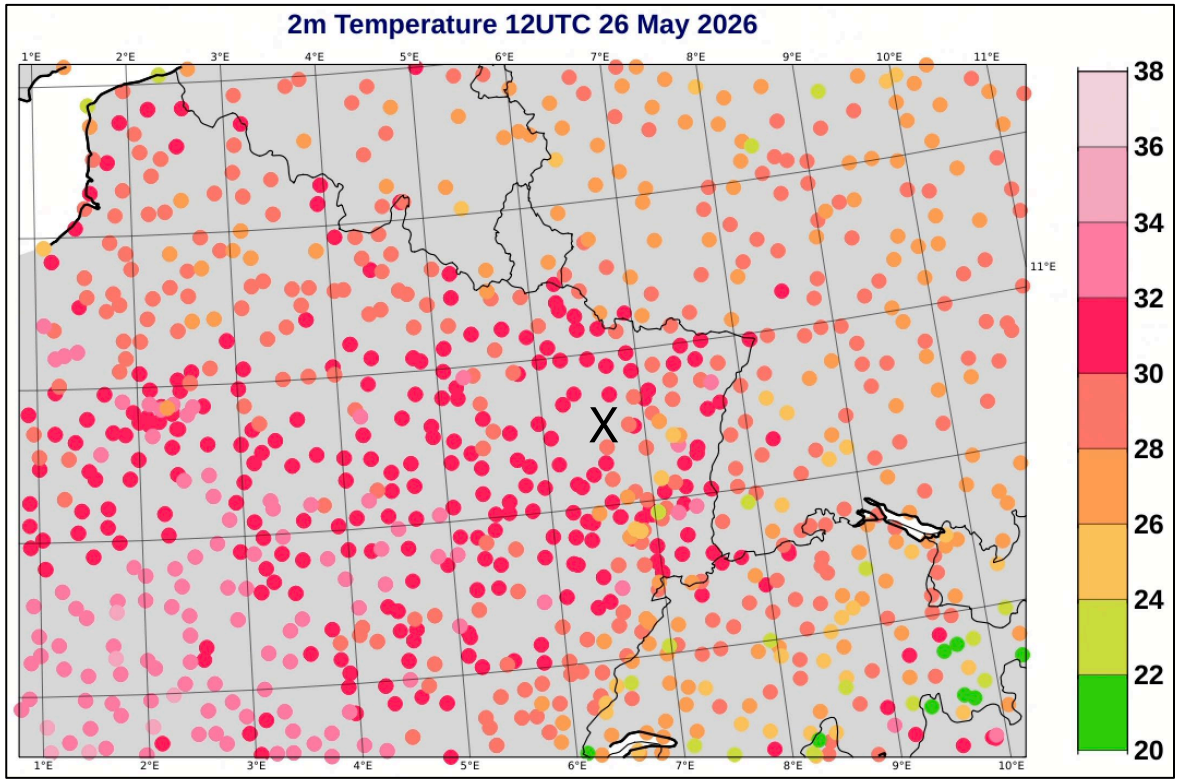

## How much energy needed to change the max/min temperatures by 20C in these cases ?

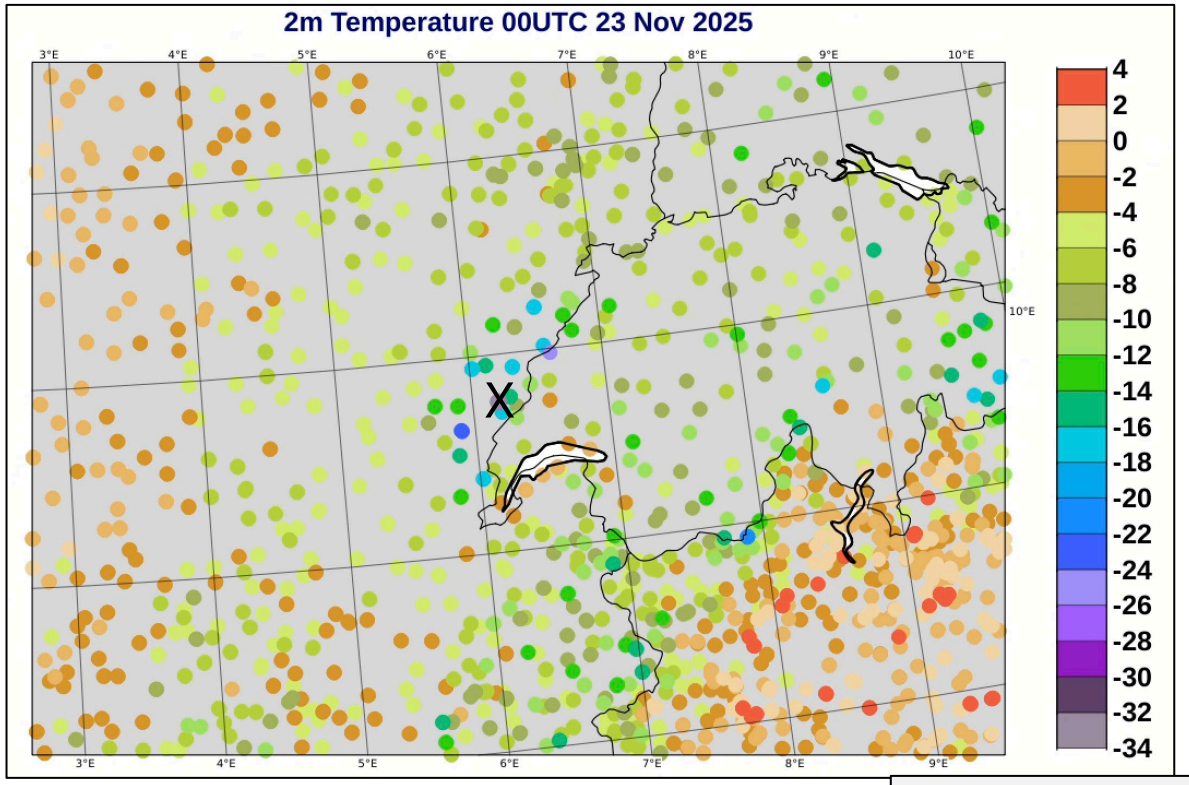
In the warm case the heating necessarily extends over 7000m, say (via convection)

In the cold case it extends over maybe 30m (no convection, minimal mixing)

So even though these estimates are imperfect, it is clear that the energy change (atmospheric storage) in the warm case is orders of magnitude larger

Therefore, other things being equal, max temperature predictions should be much more accurate than min temperature predictions, even more so for extreme values




*Combe Noire Case*

Equally, there is more spatial consistency in maxima on hot days, than in minima on cold days.

Dry convection puts a cap on the extreme maxima that are possible (implying also that QC of reports on hot days is much more tractable).

No such cap on extreme minima.

# 2. A closer look at the Combe Noire case

**23 Nov 2025**

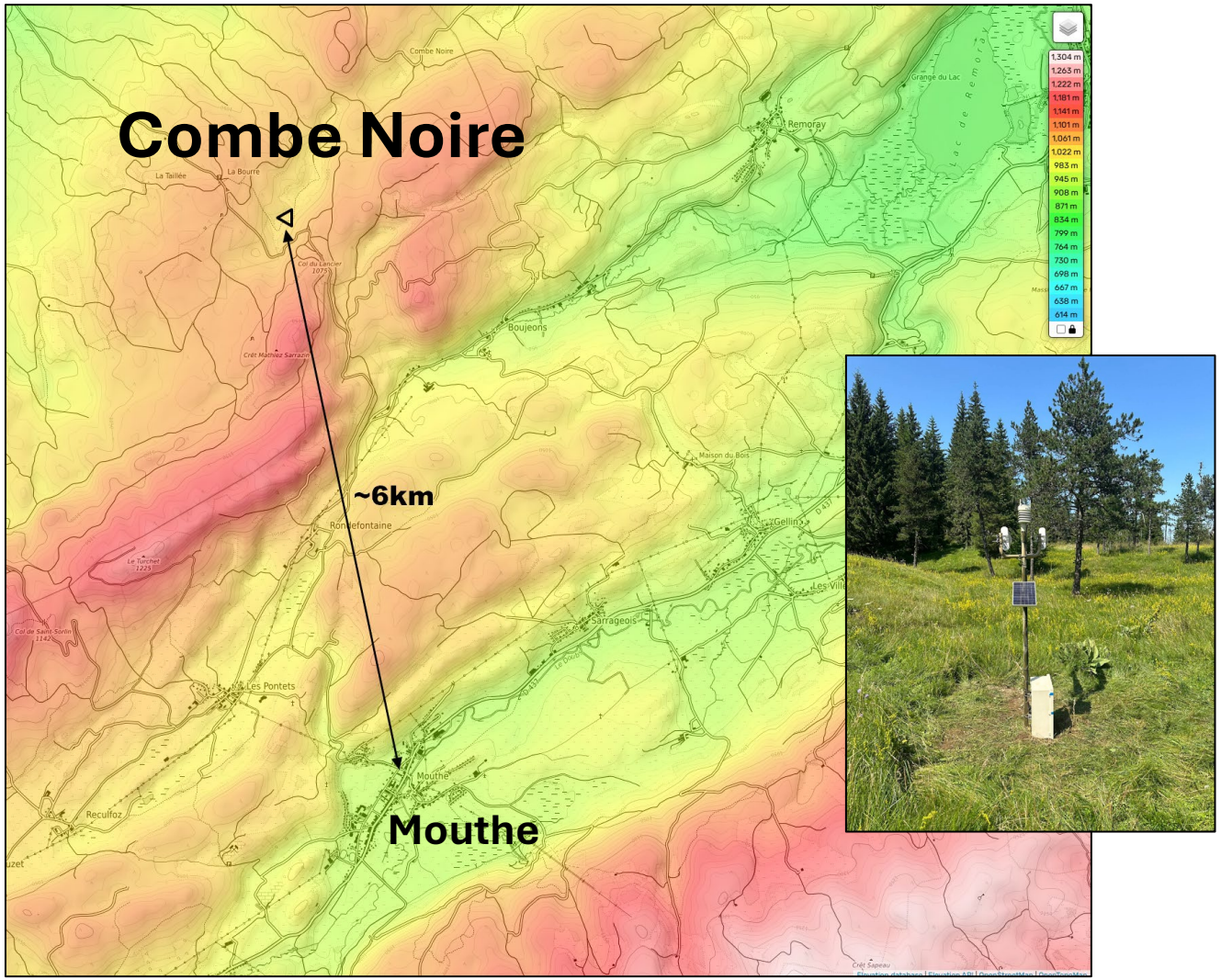

6th place ?

France low temperature record – Mouthe: -41C, 17 Jan 1985

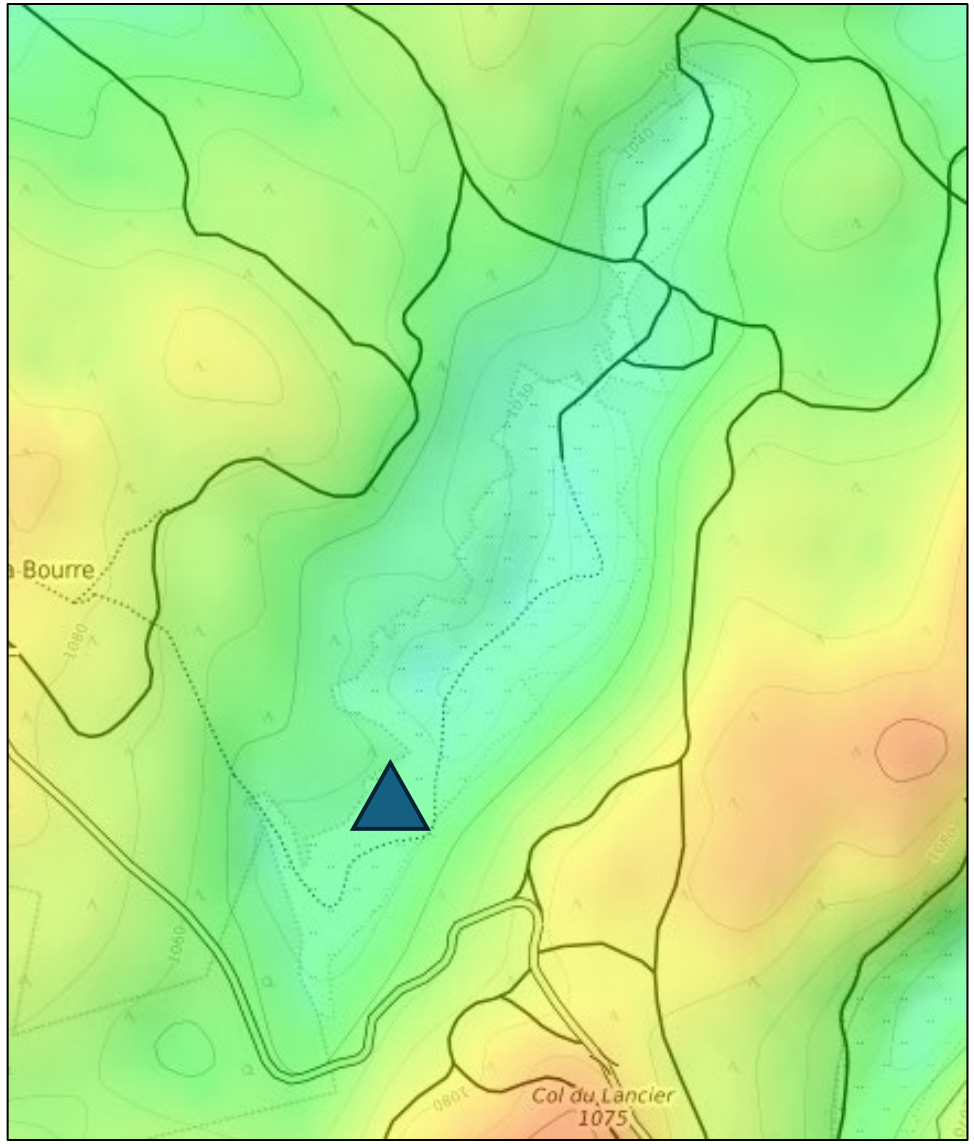
6km away!


QC investigation at ECMWF suggested this was quite plausible

**Bowl 1.5km long, 30m deep**





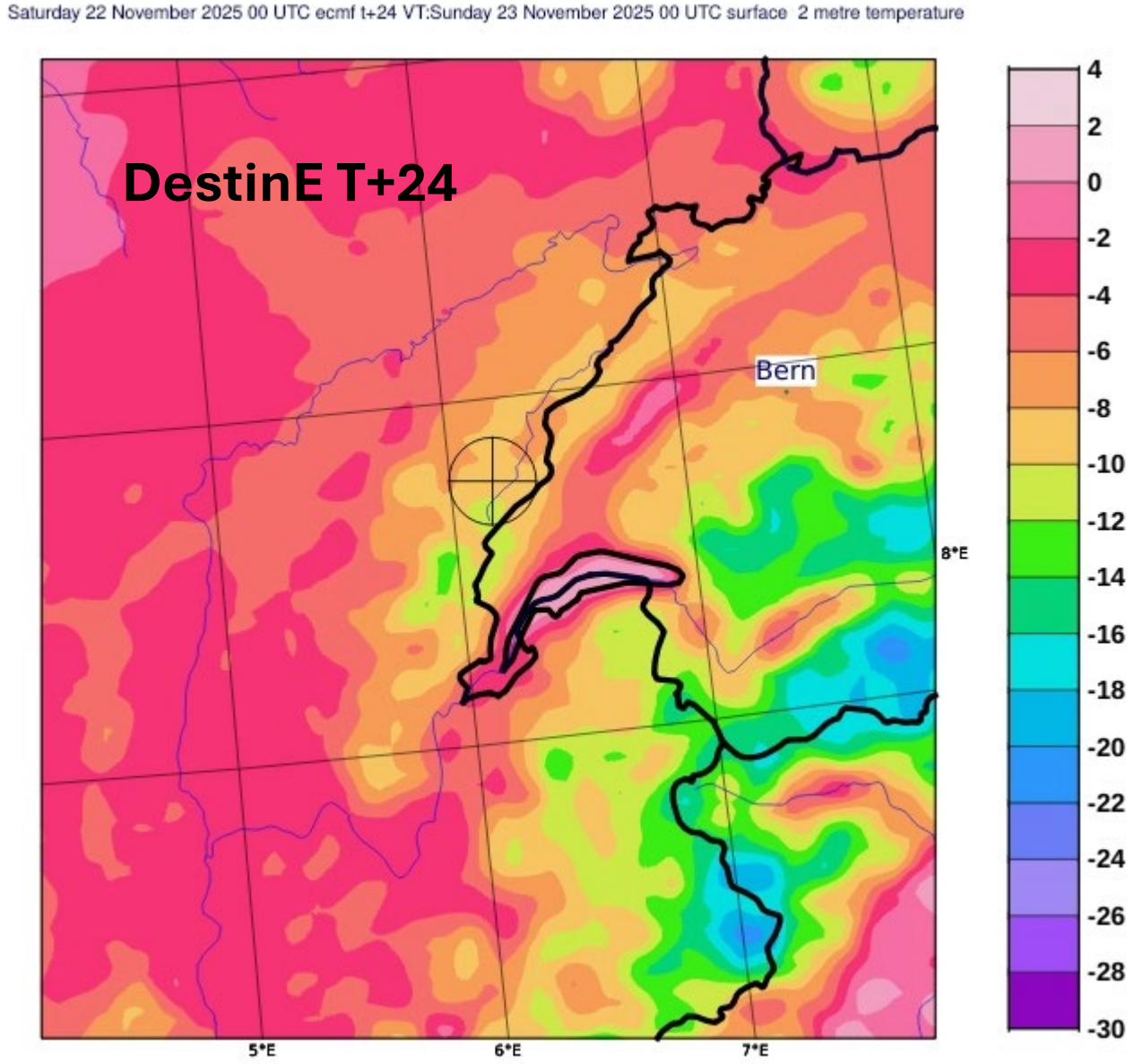

Sensor site seems to have been judiciously chosen!

# What are the ingredients for extreme low temperatures ?

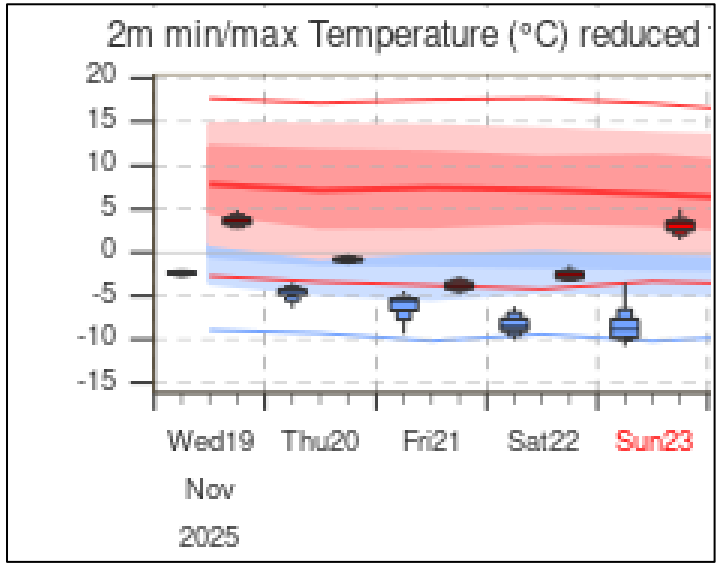
- *Bowl* topography (stops cold air escaping by drainage) ✗
- High altitude (colder anyway) ✓
- Deep cover of low density snow (insulates from upward heat flux) ?
- Cold airmass (obvious?) ✓
- Dry air in depth (promotes radiative cooling – strong net OLR) ✓
- Light winds (less mixing) ?
- Clear skies (stops most downward long-wave radiation) ✓
- Dry air at surface (less chance of freezing fog) ?
- Static weather situation – with cold previous day (cold starting point) ✓
- Obscuration of daytime sun by trees or mountains (colder starting point) ✗
  - But not too much!

***Pretty much all the above were satisfied in this case***

*How good is the IFS at representing these factors?*




20-25C errors in model forecasts (too warm!)




But close to the min temperature extreme (1st percentile) in the model climate

# So how can 2m temperature forecasts possibly go wrong, for extreme minima?

- In models, very easily!
- Because of intrinsic predictability challenges, linked to:
  - Tiny amounts of energy involved
  - Model resolution limitations (IFS), meaning that representation of critically important local factors is not possible
  - Extreme local (sub-grid) variability
- There are also important (interlinked) secondary factors:
  - Low level cloud / fog representation
  - Wind speeds rarely drop to zero in the IFS
  - Data assimilation very challenging near the surface (vertical profiles and 2m temperatures)

*The France scenario shown is similar to many seen over N Europe in winter – this hurts the ECMWF day 5 CRPS-based headline score for 2m temperature (>5C is considered a bust)*

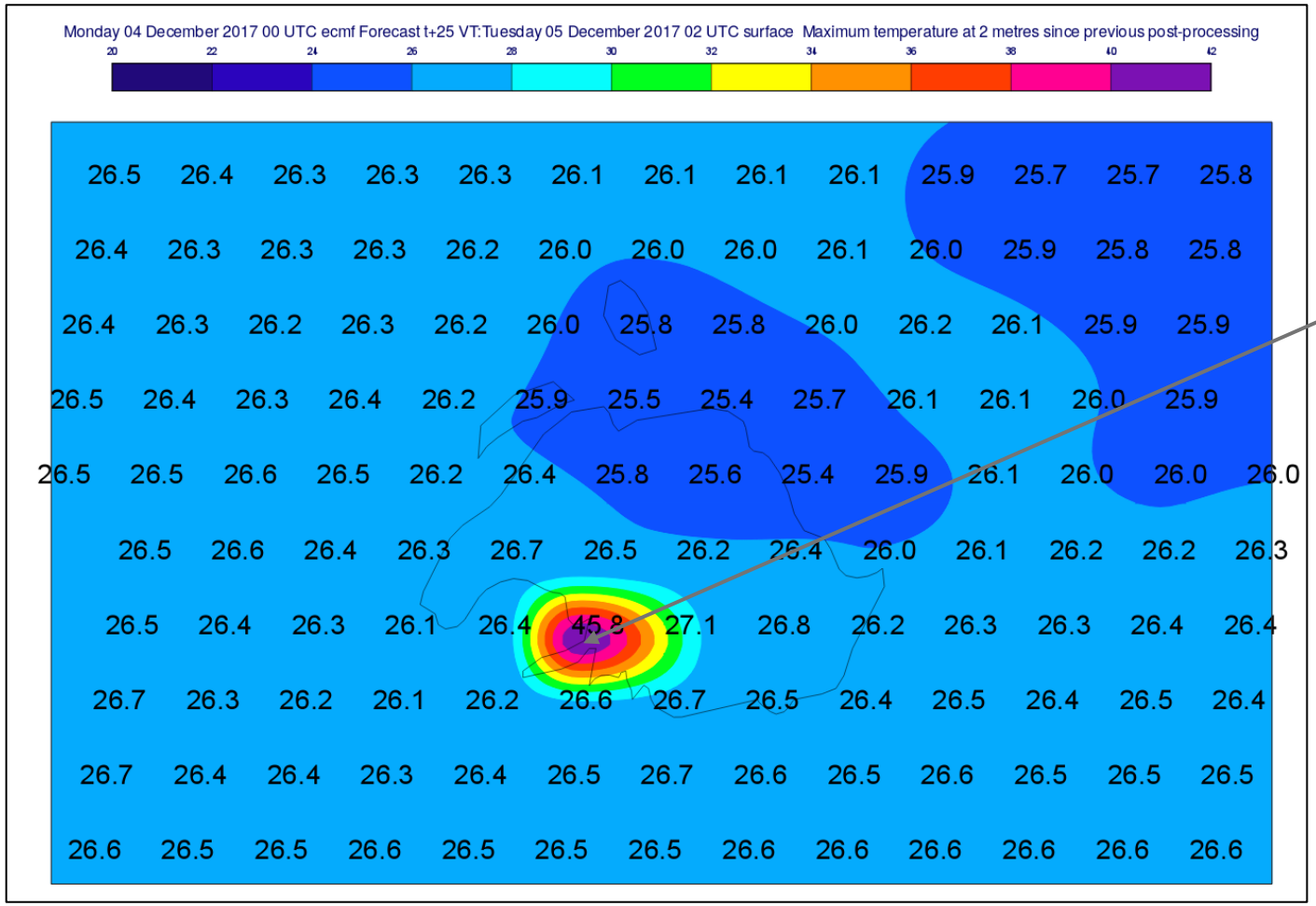
# 3. Erratic IFS physics behaviour – for maxima and minima

- Interactions between components of the physics in an increasingly complex model can be problematic
- Issues can occur very infrequently, can be confined to one model timestep, and can be very hard to spot
- Clever monitoring tools need to be designed – “expect the unexpected”

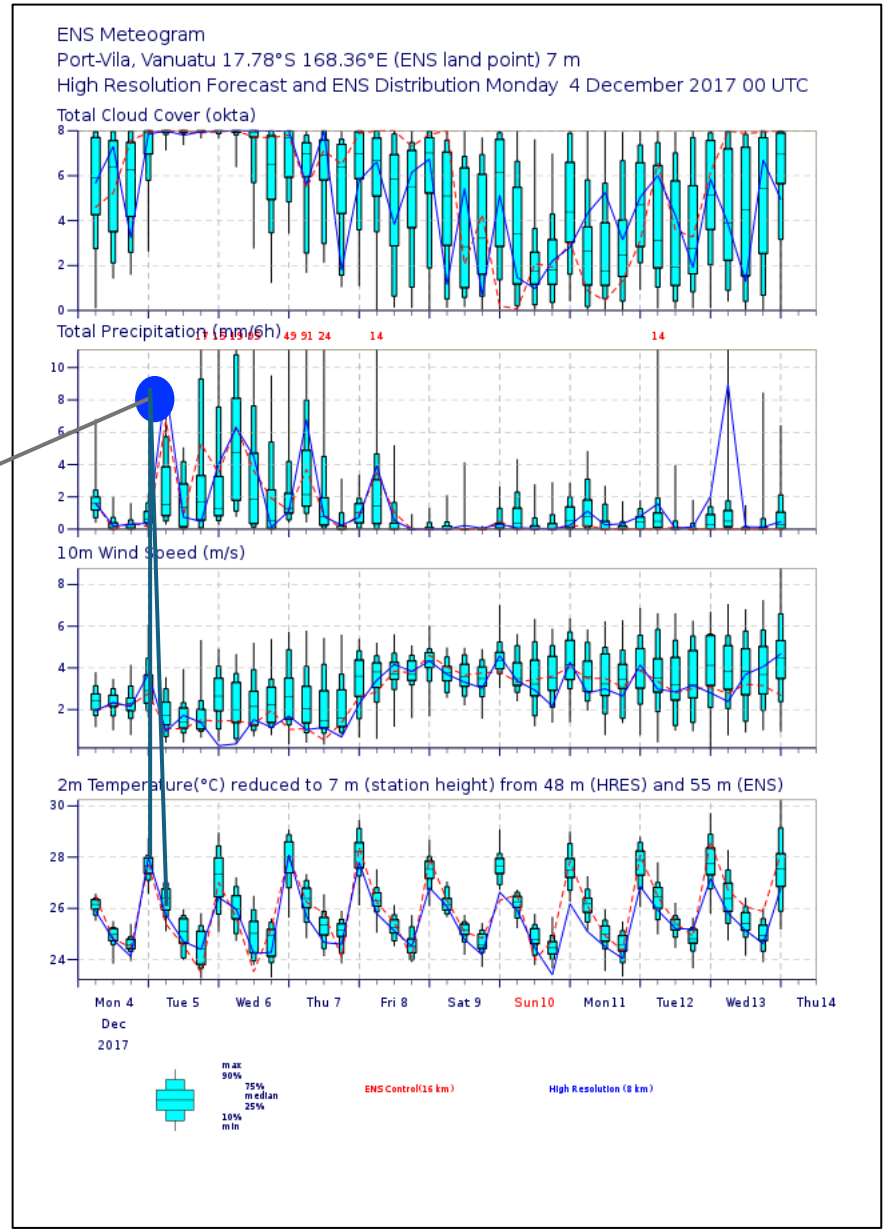
- The land surface scheme uses tiles to represent different types of surface within a gridbox (low vegetation, bare soil, etc.)
- The state of the tiles can vary, due to rainfall
- Switchovers from dry to wet (and vice versa) has caused instability problems for 2m temperatures in the past

- Users are encouraged to report any particularly odd forecast behaviours to ECMWF, but ideally after checking the known model issues pages first

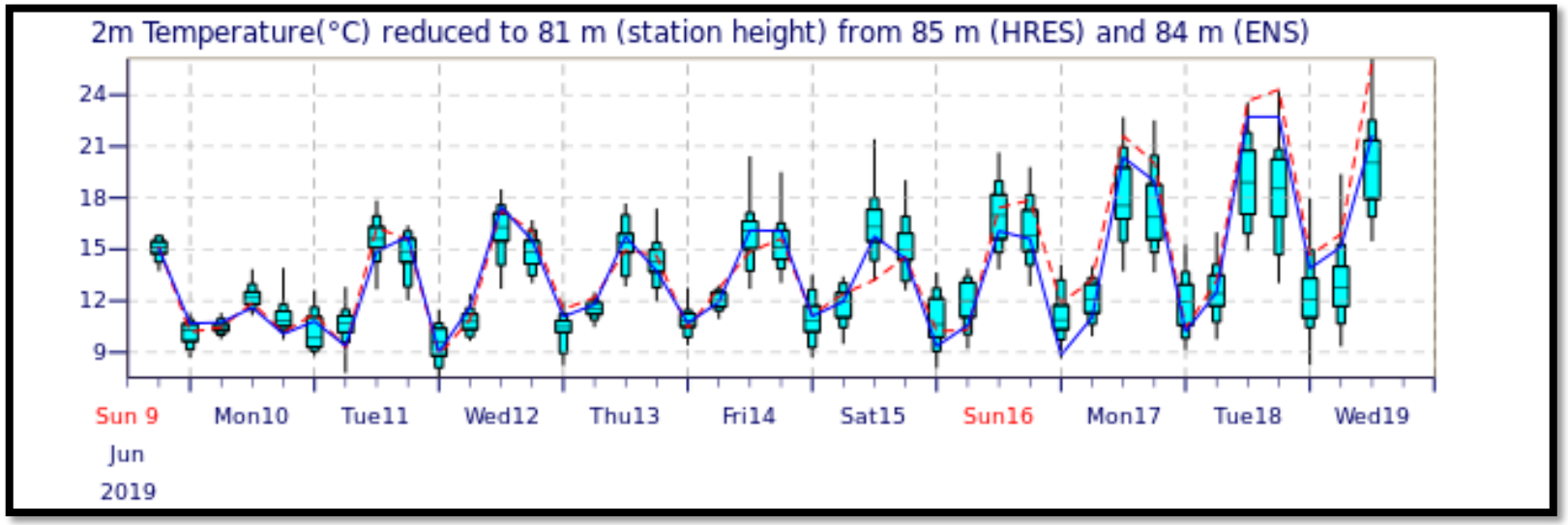
# Short-lived “Hot Spikes” (2017)




**Vanuatu (Efate) – maximum temperature in previous hour**

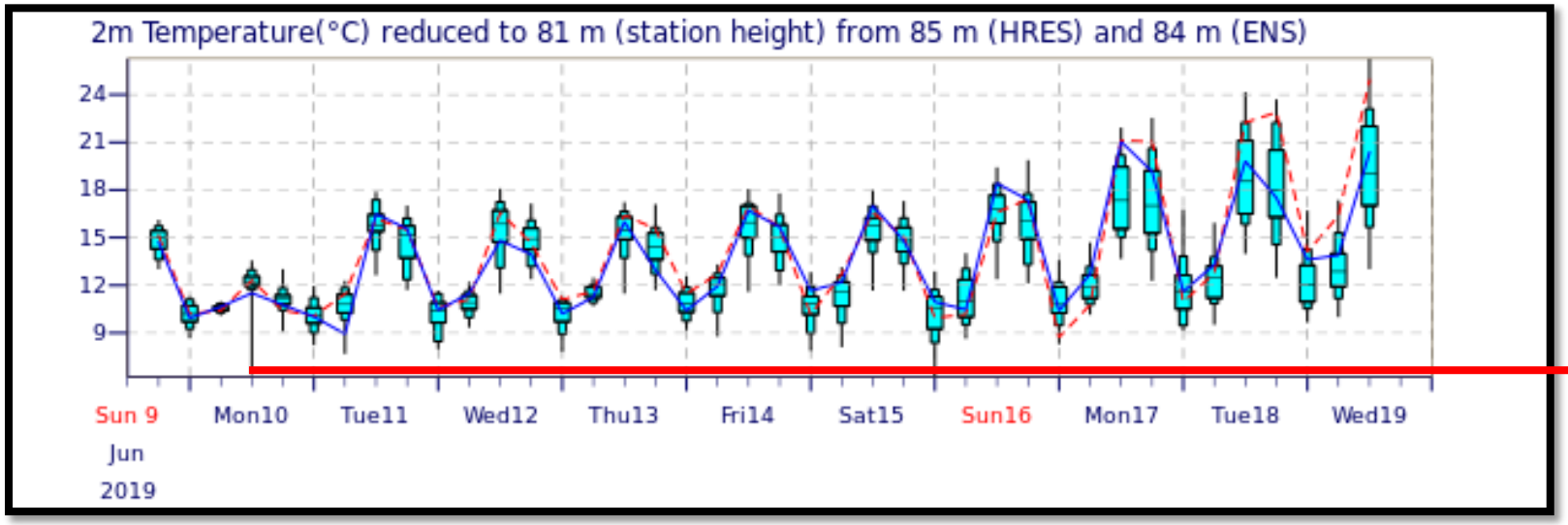


# “Cold Bubbles” in cycle 46r1 (June 2019)



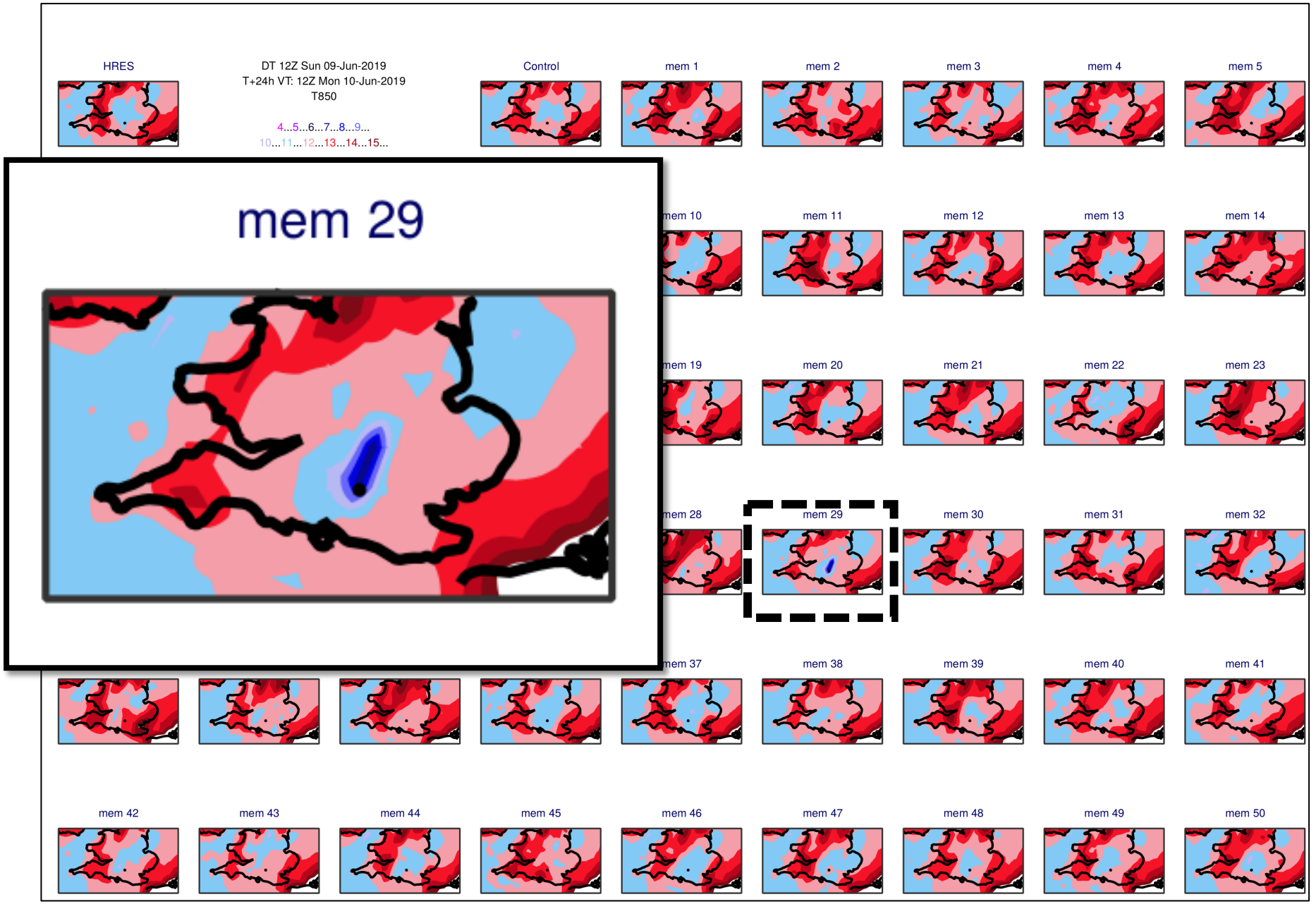

**Reading**
**BT: 9 June 2019 12 UTC**

O-suite




E-suite

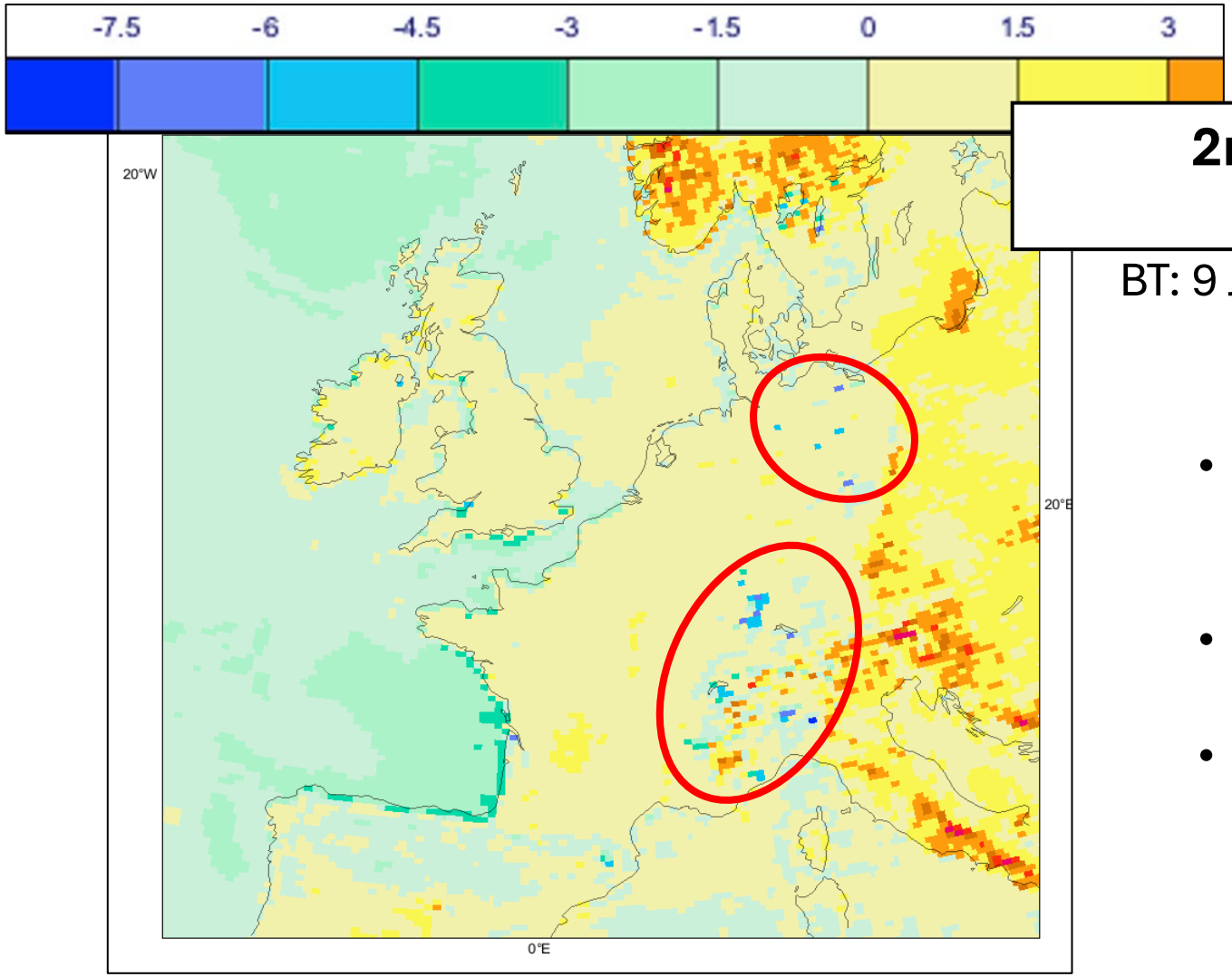
~ 6 degrees




2 m temperature stamp maps (ENS members)

24h forecast for 00UTC

# 12h later in the same forecast…



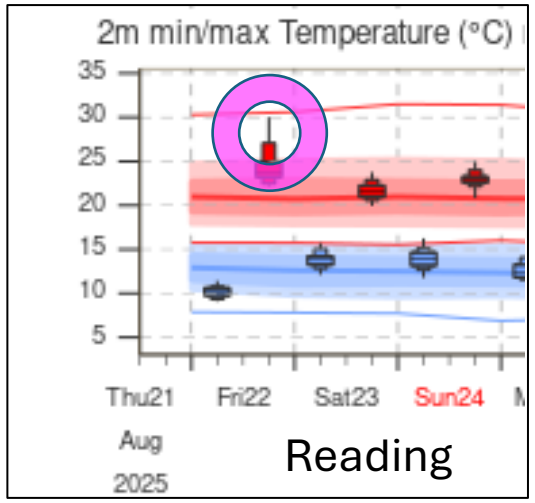

**2m temperature – skin temperature**

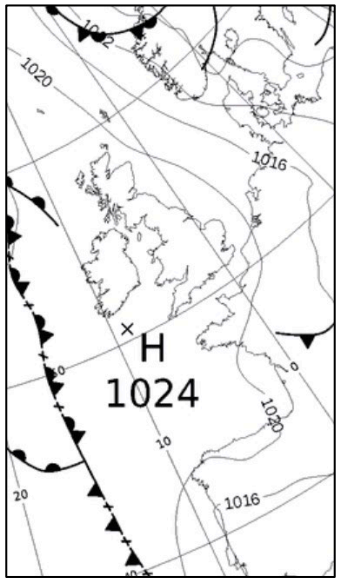
BT: 9 June 2019 00 UTC VT: t+36h

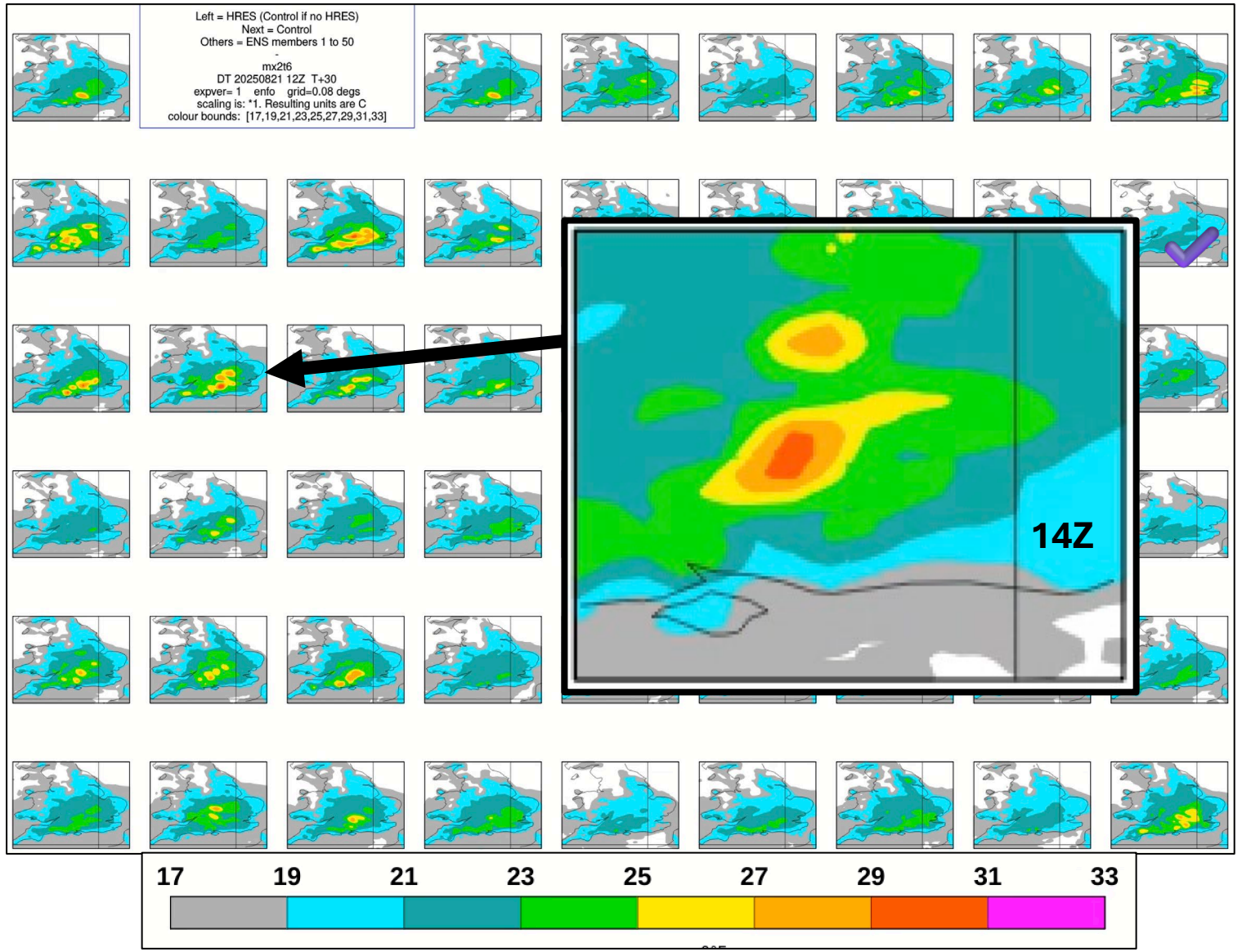
- 2m temperature up to ~10C lower than skin temperature!
- Very locally
- Found to only occur in areas with recent rain

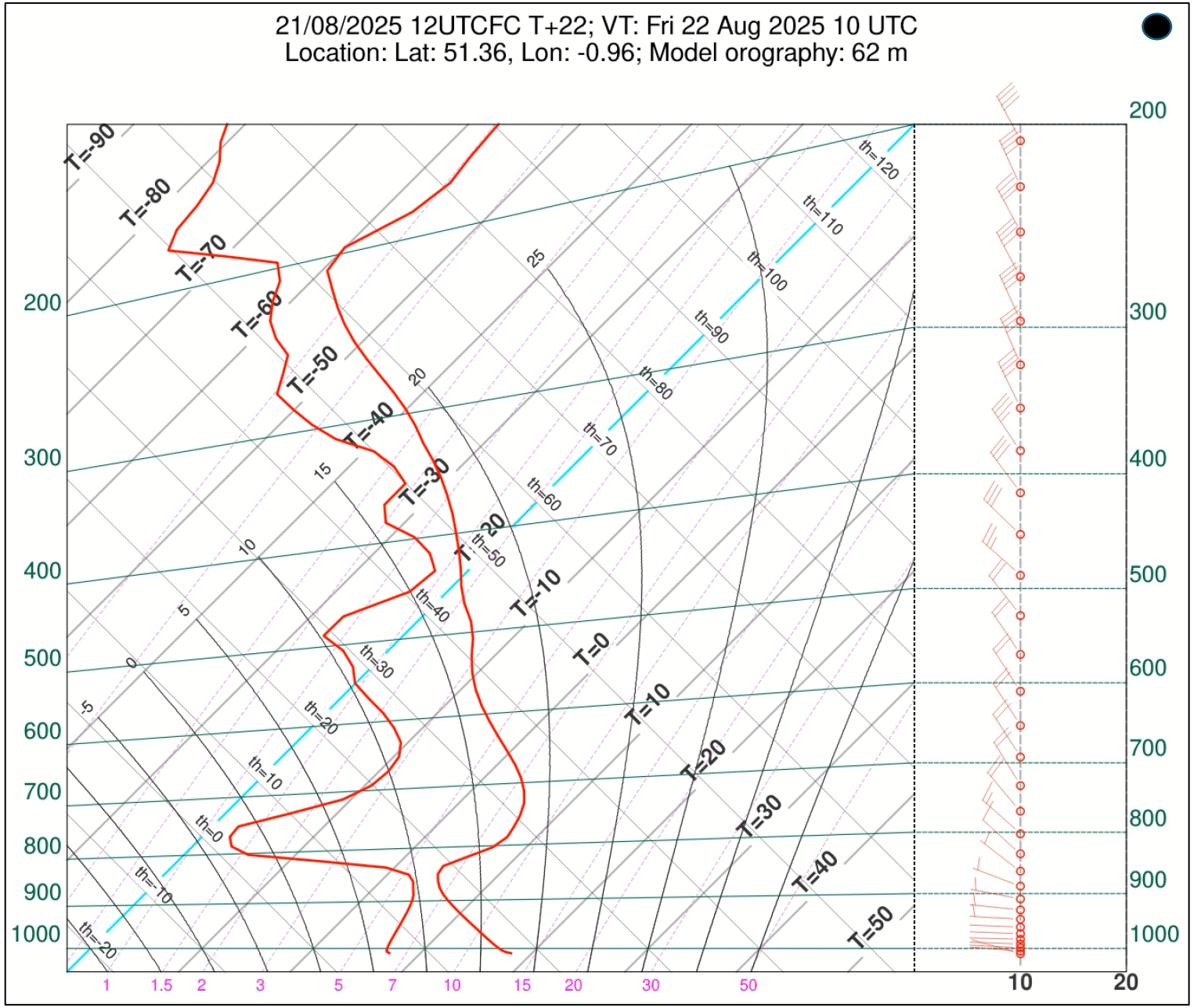
# “Boundary Layer Oscillations” (2025)

Max temp 12-18Z on Fri 22 Aug 2025



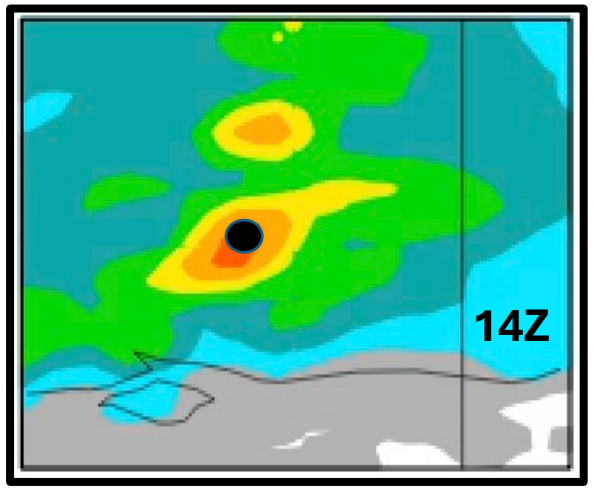


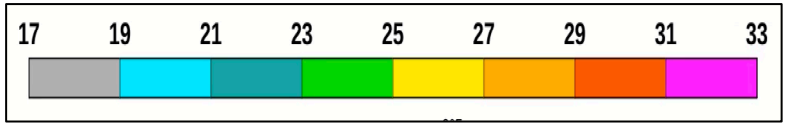

Forecast Sounding Sequence

Member 16

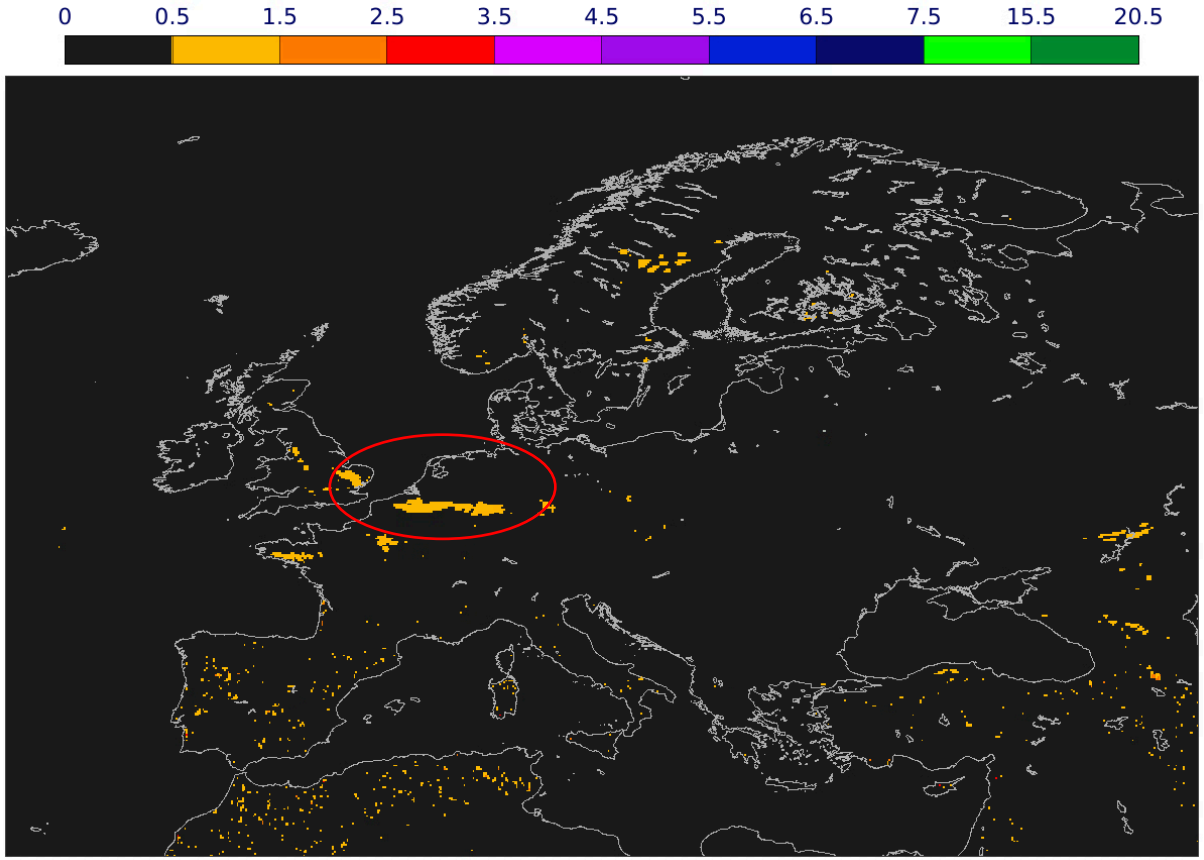
Reading

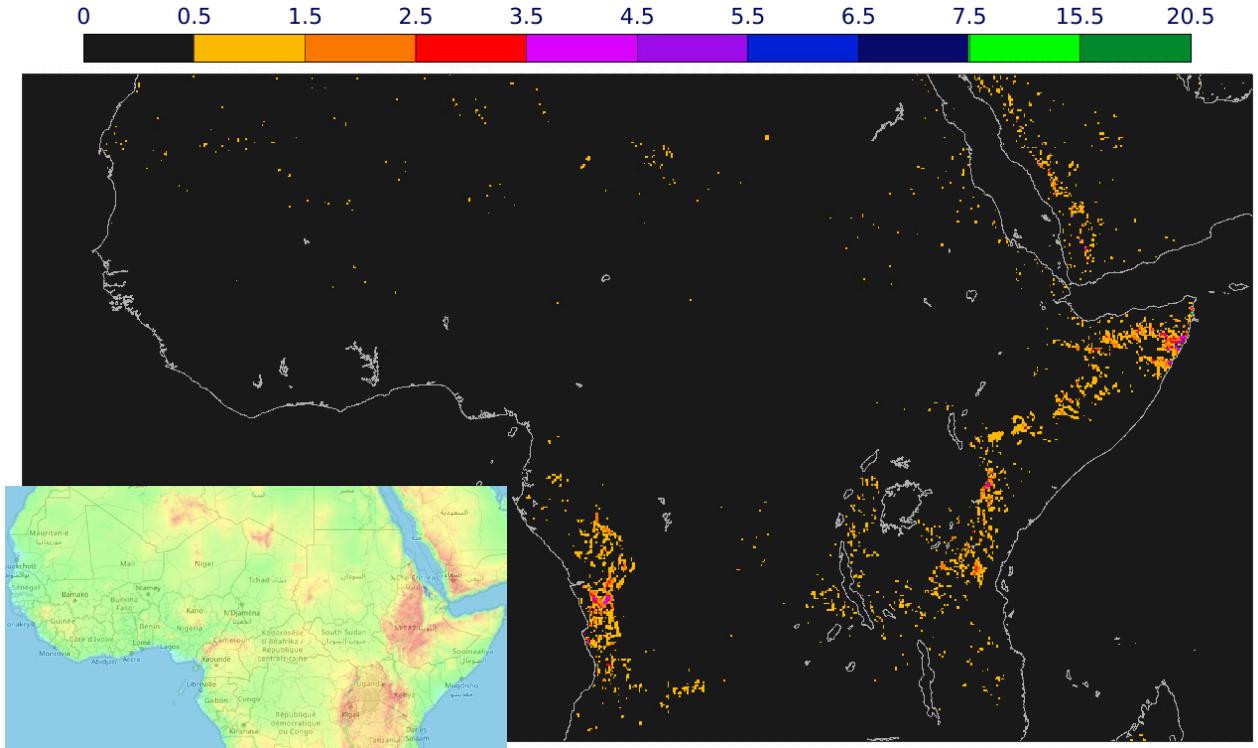
10-18UTC, 1h intervals



# Super-adiabatic lapse rates – number of days in August 2025




- Usually scattered around but sometime organised structures can be identified.
- Some areas, particularly, but not only, of complex orography seem to be more prone to super-adiabatic lapse rates.

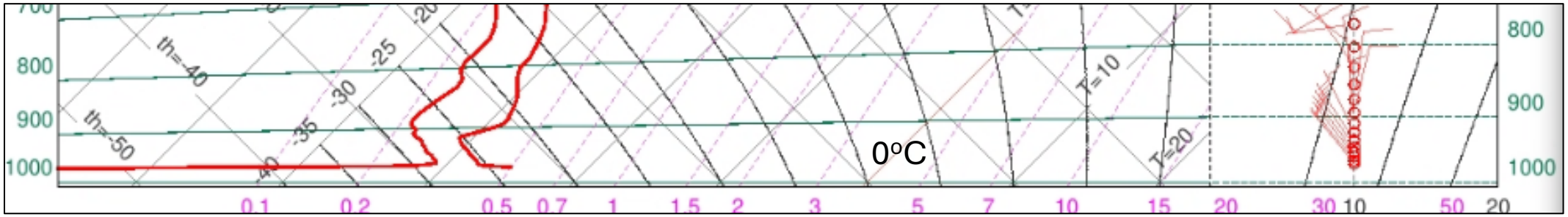
# Fixes

- A model fix for the “**Hot Spikes**”, related to the skin conductivity for wet tiles, was implemented many years ago
  - This was a complex problem involving considerable experimentation
  - Not all instances of undesirable behaviour were fixed…

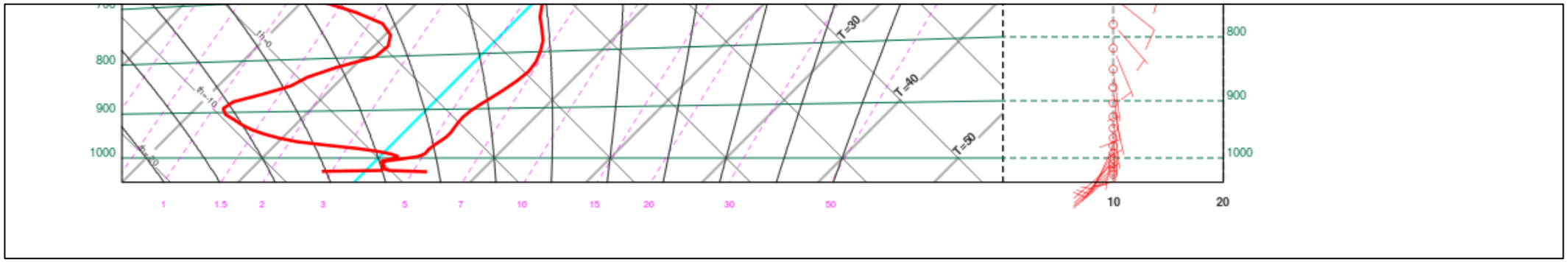
- The “**Cold Bubble**” problem was discovered on model implementation day (!!)
  - The fix in this case was a simple, quick reversion of a newly implemented physics change
  - This related to definition of the dominant low vegetation tile

- The “**Boundary Layer Oscillations**” issue related to interplay between various IFS physics components
  - Mainly occur in the presence of just an isolated single cloud layer, though contrary examples also found
  - A partial fix was implemented in cycle 50r1 (May 2026), but work continues
  - Instances also found in ERA6, potentially more common than in 49r1. Not fixed.

# Other Cases



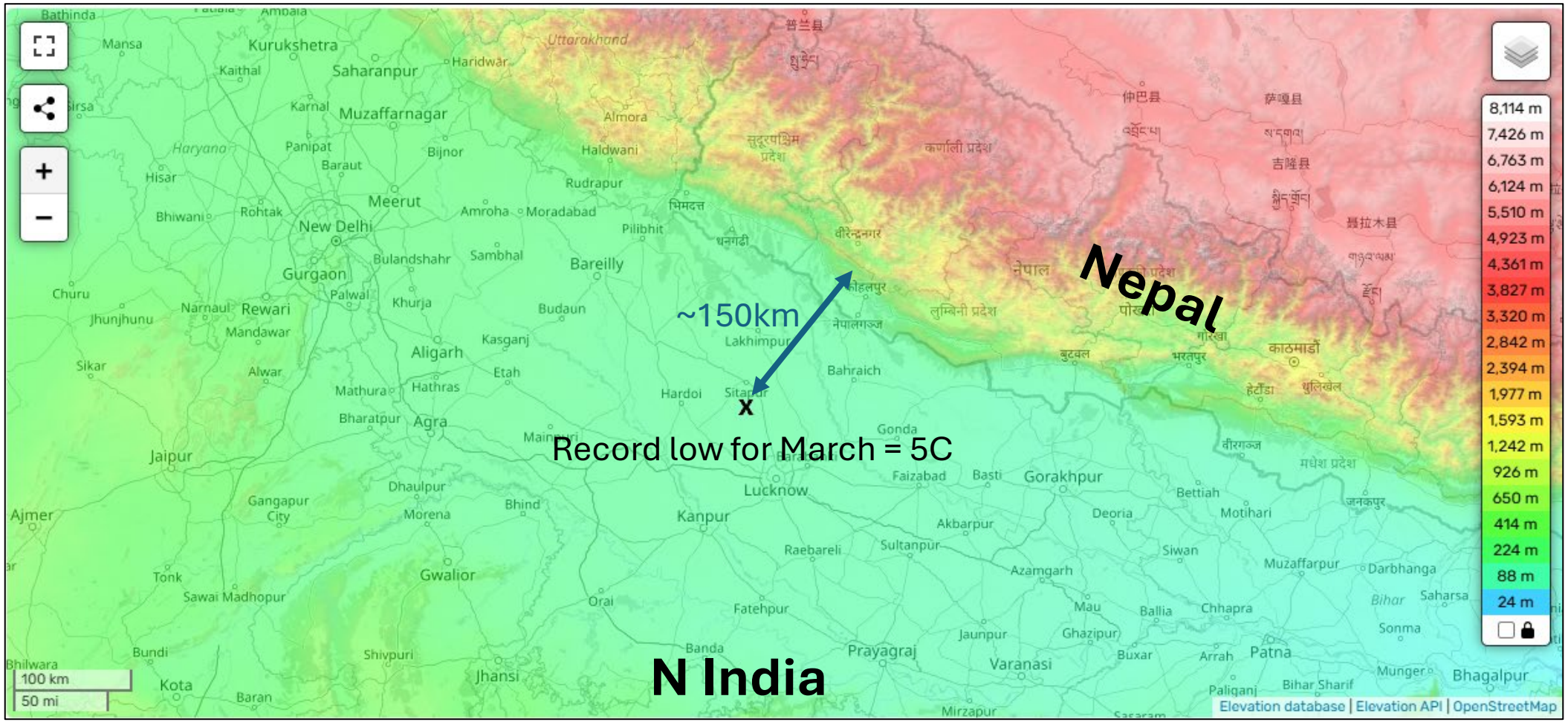

Huge volumes of spurious snowfall



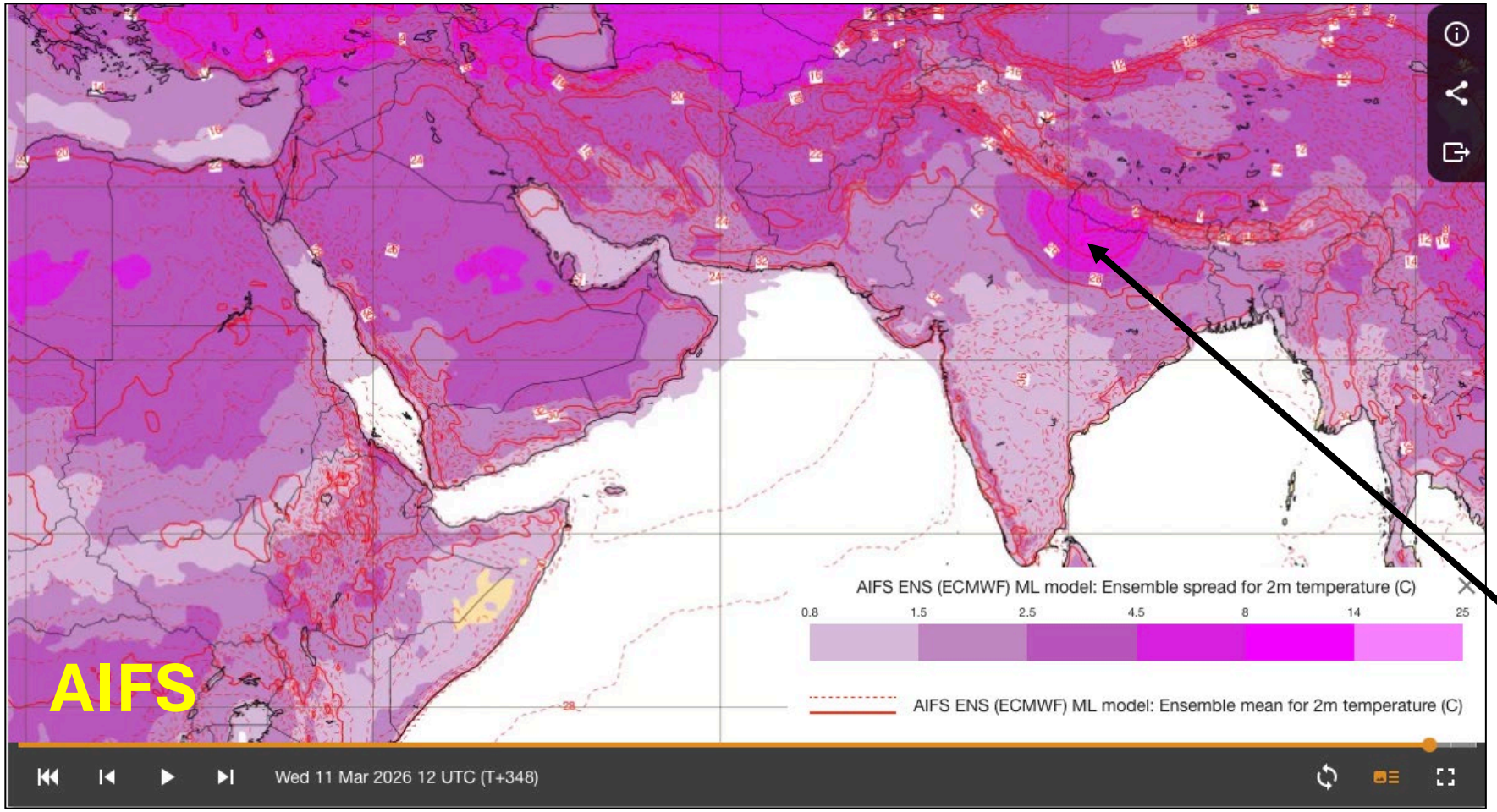

Runaway fog

**2m temperature problems can also be “symptomatic” of other model issues !**

# 4. AIFS-ENS “cold explosion”



# AIFS versus IFS - 2m temperature spread

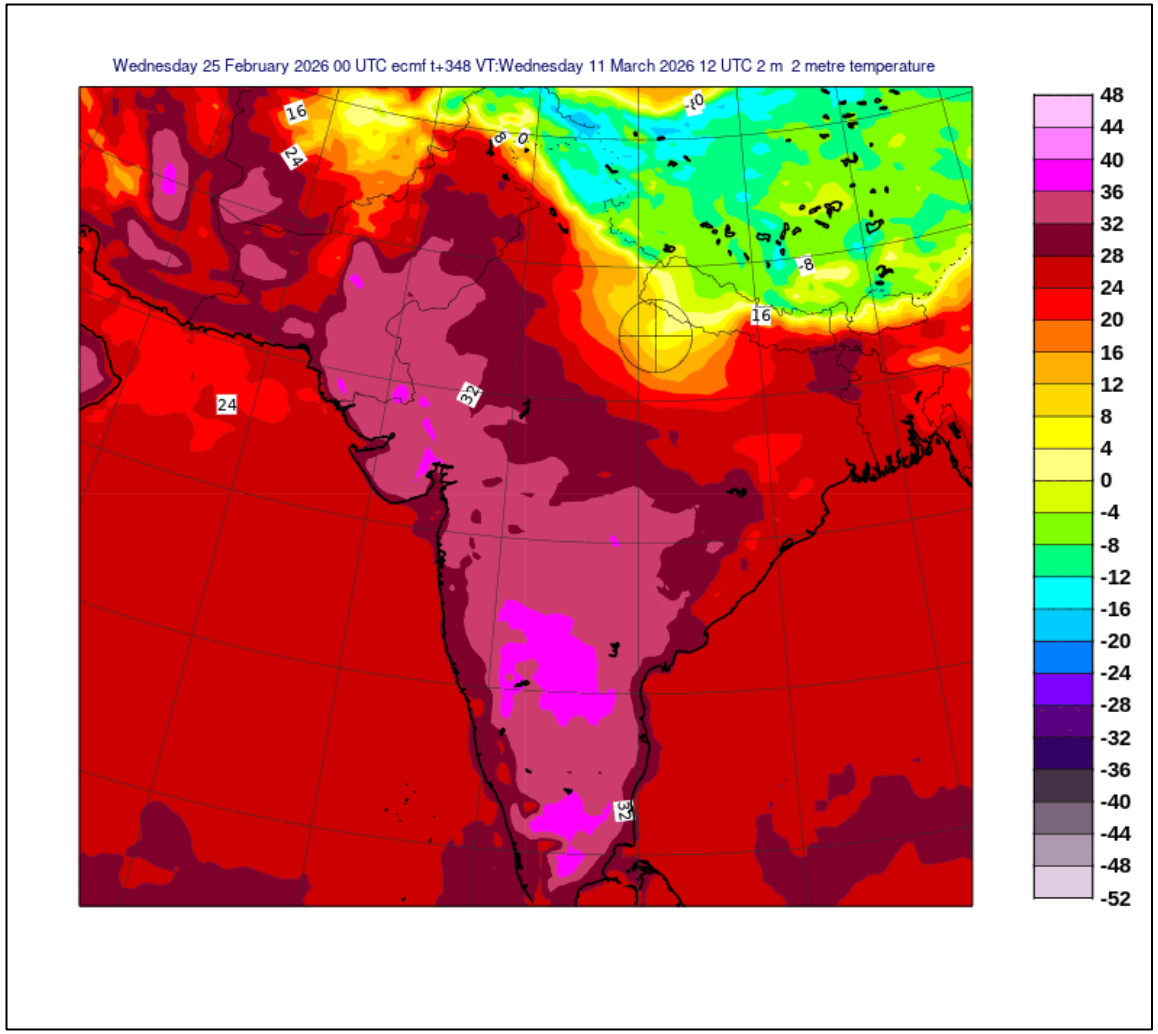



**Day 15**

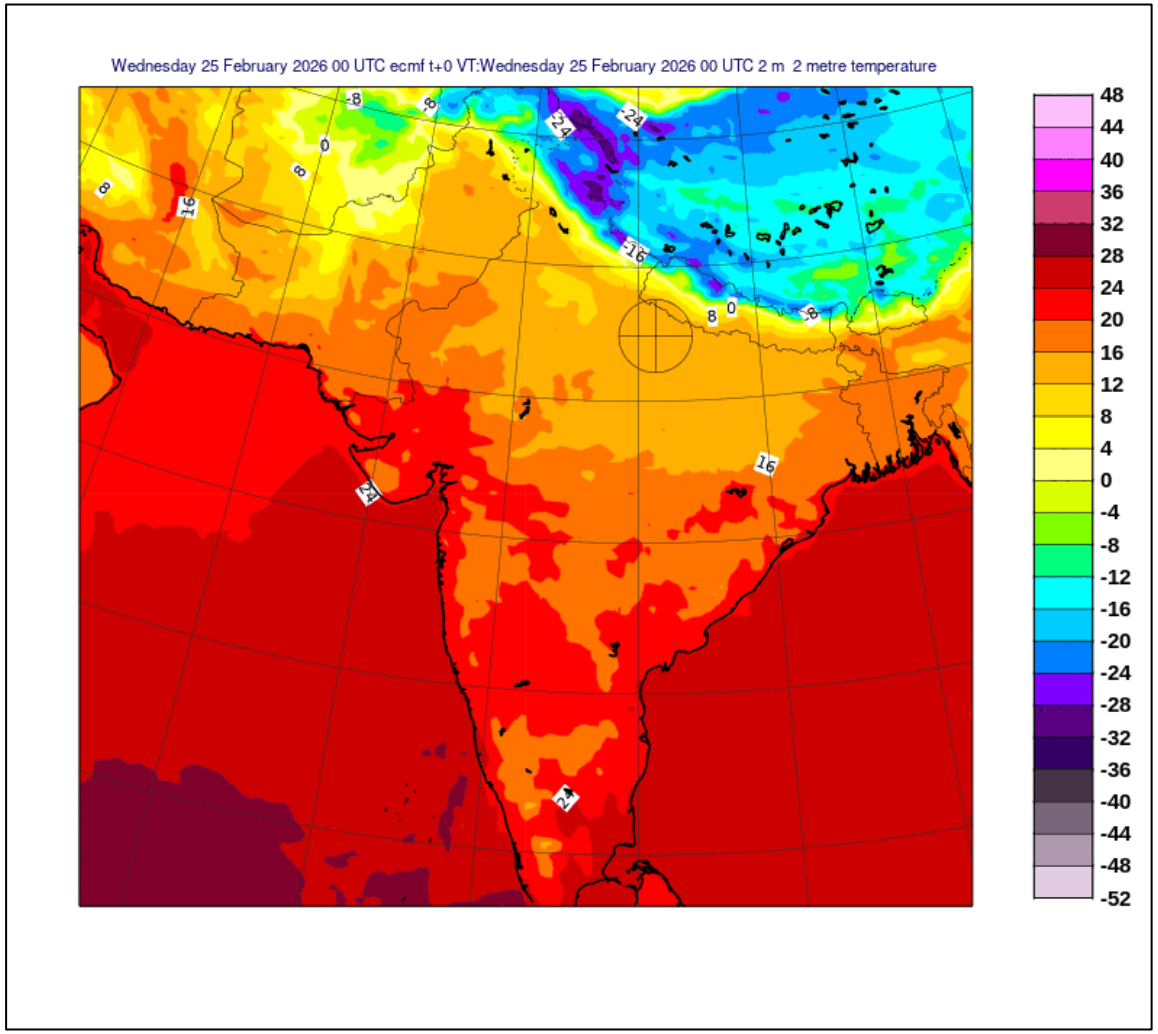
**Valid 12UTC 11 Mar 2026**

**>8C**

# 2m temperatures in AIFS ENS members




All members at T+348 = VT 12UTC 11 Mar




Member 45 full 15-day evolution

# Remarks

- In the most erroneous member (45)
  - No cloud of note
  - Sub-zero temperatures in late afternoon at the end of the run
  - Cold explosion physically unrealistic (though had a diurnal cycle superimposed)

- This was the old AIFS-ENS V1. We now have V2.
- V2 may not exhibit this behaviour, although such cases are hard to trap
- Finding this case was just by chance!
- However we had already seen anomalous cooling in other areas, like Iran, E Africa, Red Sea

- Without a full understanding, pollution of any AIFS re-forecasts, or longer sub-seasonal integrations (both to come) must be a concern

# 5. Summary

- From simple energy consideration we expect night-time minima to be more difficult to predict than daytime maxima.
  - Particularly extreme values.
  - Evidence bears this out. Forecasters understand. Huge spatial variability is possible.
  - Should we be targeting physics changes to address these; would bias correction be a more expedient option ?
- In model world the picture can be more complex, due to a multitude of interactions between physical components.
  - On occasions when minimum temperatures do seem to be predictable, bizarre errors can still occur
  - And for clearcut maxima there are also very implausible outcomes at times
- In general, the cause of errors in classical physics-based models can be tracked down, and in time addressed
- AI models are still in their infancy. We have identified some even more bizarre behaviours. These tend to be more difficult to understand and address. More work on “explainable AI” seems to be needed.
- ECMWF encourages users to keep reporting to us any strange model outputs they come across ☺

*“Expect the Unexpected”*

# Thanks for listening !

**Tim Hewson**

Principal Scientist, Evaluation Section

tim.hewson@ecmwf.int

www.ecmwf.int