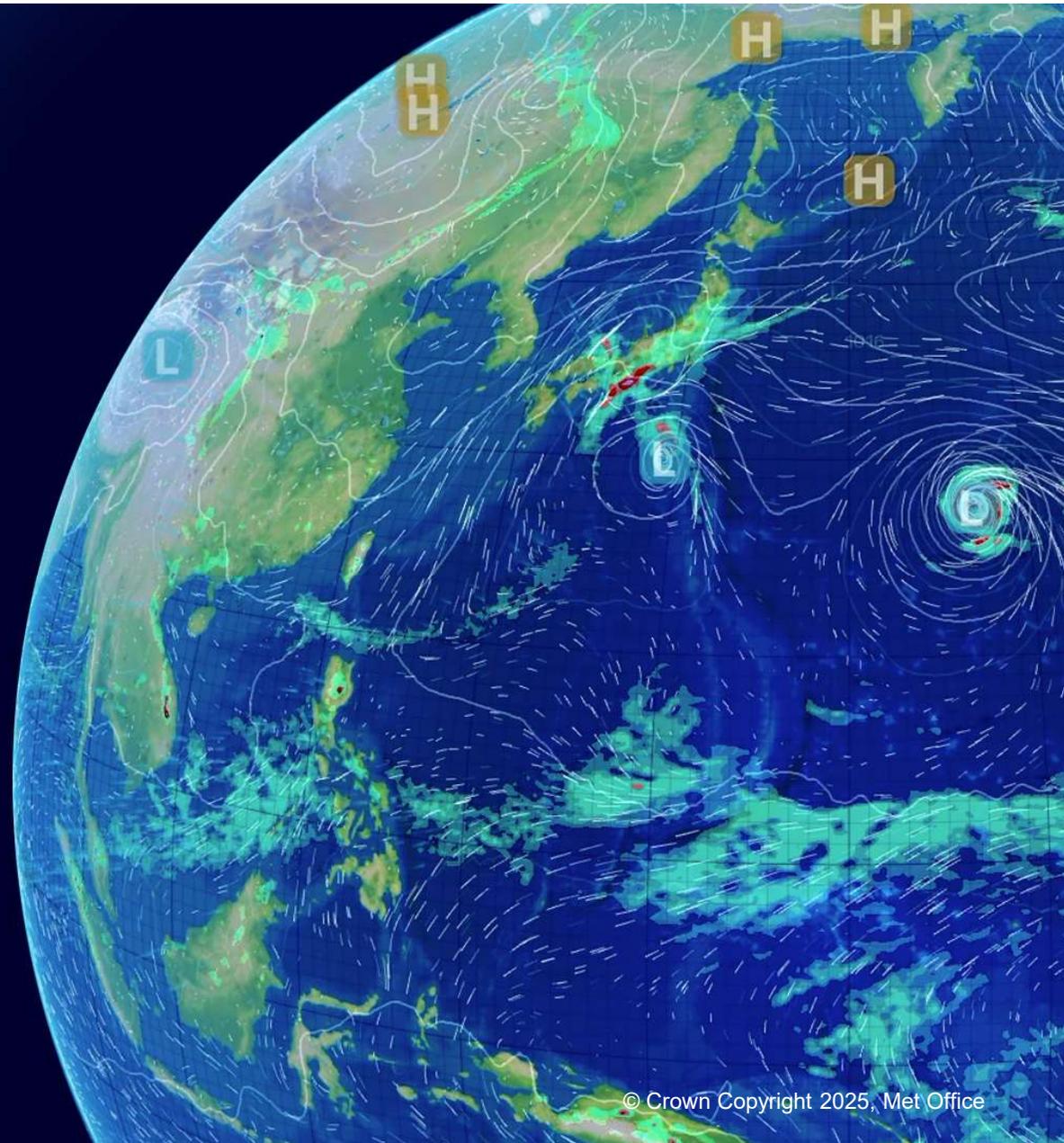# DA vs Leadtime

23/09/2025

Marco Milan and Gareth Dow

Motivations

≋ Met Office

- Estimation of the longevity of the benefits of Data Assimilation (DA) in the Limited Area Model (LAM) forecasts, focused on UKV system and domain.

- Project in collaboration with Satellite and Surface Assimilation (SSA).

- The actual UKV-DA cycling is composed of two main parts:
  1. A LAM one, composed of the UKV 4DVar and the Land Surface DA.
  2. The Large-Scale Blending (LSB), which nudges the large scales of the LAM fields towards the ones of a host model (a downscaler derived from global analysis).

- The planned future NWP system consists of three merged parts:
  1. Nowcasting (up to T+2).
  2. JADA regional ensemble data assimilation-based model. An EnKF (up to 1+12-T+24).
  3. MOGREPS-UK regional ensemble forecast model.

- JADA-Regional will concentrate on convective scale DA. The purpose of this study is to define the lead time for JADA

Reproducibility suite (cycle by cycle)

- We started from the Control run of our op. suite, hourly cycling.
- Reproducibility suite. A suite that retrieves the dump files and all increments used in the Control to reproduce its capability. No run of Data Assimilation (VAR), to avoid differences.
- The retrieved files are: Dump file (background) for the initial conditions and all increments files.
- *The background contains information of the observations at T-1.*
- We estimate the value of the increments in the latest hour. No bias in the skill due to cycling DA (smaller differences).
- Different sensitivity tests:

	Control	ALL_OFF	ALL_OFF+LSB	ALL_OFF+VAR+SURF
4D-VAR inc.	ON	OFF	OFF	ON
Surf inc.	ON	OFF	OFF	ON
LSB inc.	ON	OFF	ON	OFF

Graph changes to call for new task and avoid all the tasks (eg OPS, VAR) not used anymore

Retrieve all needed files from Control and create different sensitivity tests

- Dump file
- Land surf inc
- 4DVar inc
- LSB inc
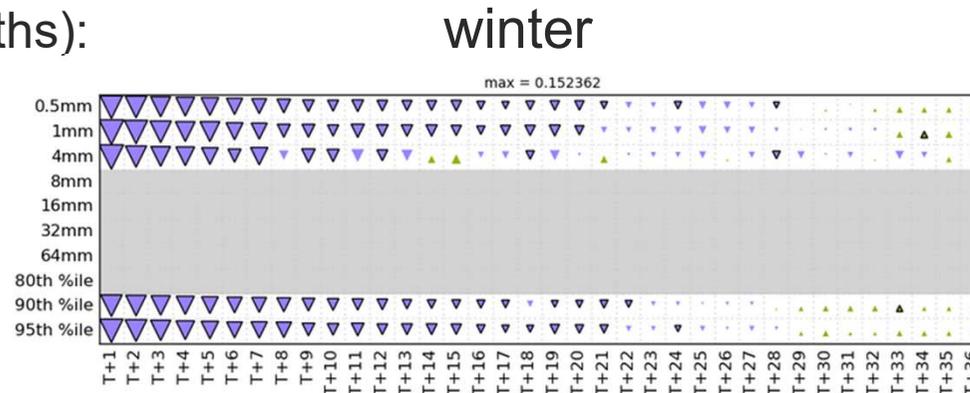
UM model Run IAU 1 step add all increments

Met Office

- Switch off the LAM DA in the Met Office operational set-up.
- The DA's effect accumulates in the background during the cycling.
- HIRA (11 grid points):

winter

- FSS (25 grid lengths):

winter

Met Office

- Switch off the LAM DA in the Met Office operational set-up.
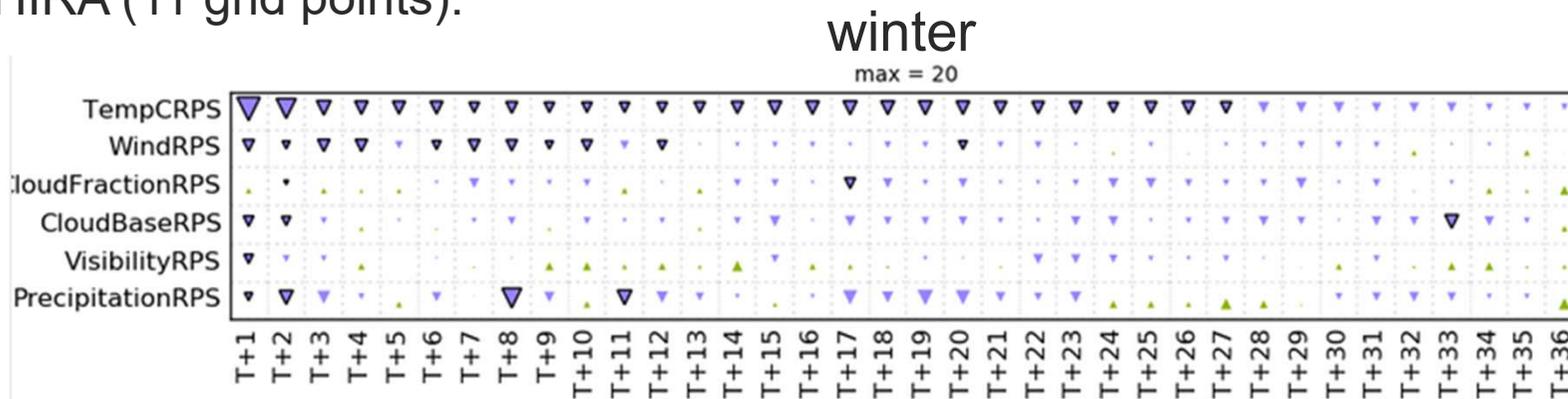- The DA's effect accumulates in the background during the cycling.
- HIRA (11 grid points):

summer
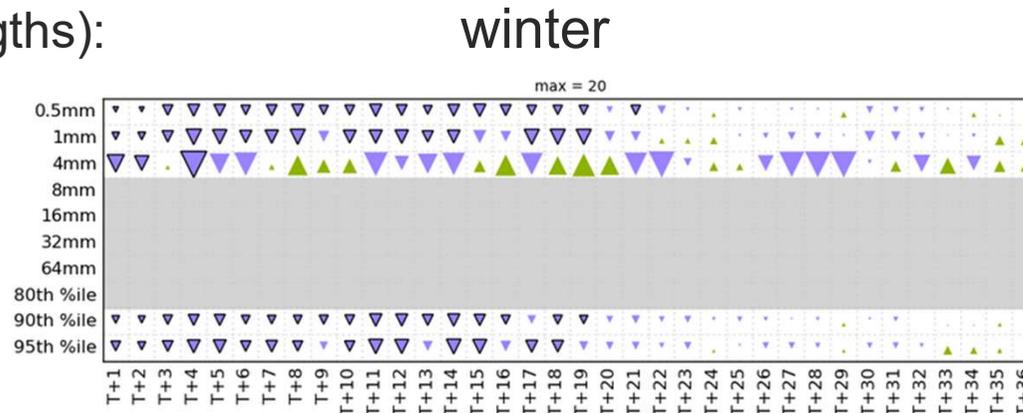

- FSS (25 grid lengths):

summer
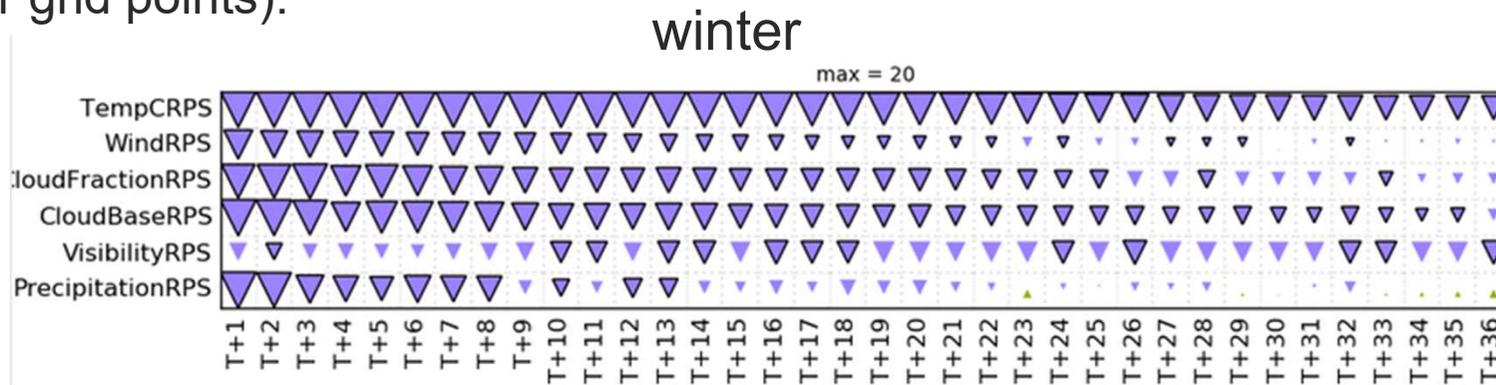
- Switch off the LAM DA in the reproducibility suite.
- Evaluation of the DA's effect cycle-by-cycle.
- HIRA (11 grid points):

- FSS (25 grid lengths):

Met Office

- Switch off the LAM DA in the reproducibility suite.
- Evaluation of the DA's effect cycle-by-cycle.
- HIRA (11 grid points):

summer

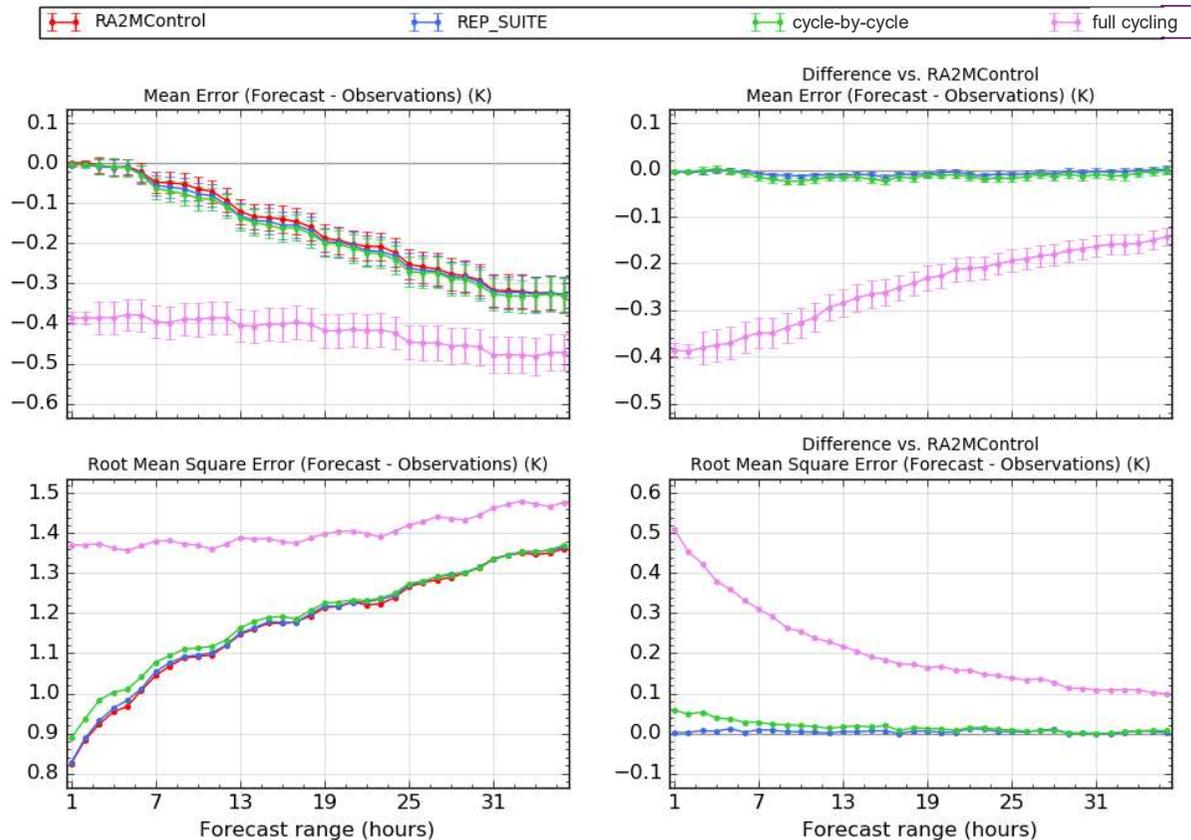- FSS (25 grid lengths):

summer

Surface (1.5m) Temperature, WMO Block 03 station list,
Equalized and Meaned between 20211201 18:00 and 20220118 00:00, Surface Obs

RA2MControl   REP_SUITE   cycle-by-cycle   full cycling

Mean Error (Forecast - Observations) (K)

Difference vs. RA2MControl
Mean Error (Forecast - Observations) (K)

Root Mean Square Error (Forecast - Observations) (K)

Difference vs. RA2MControl
Root Mean Square Error (Forecast - Observations) (K)

Forecast range (hours)

± 1 standard error bars calculated assuming independent observations

- ALL_OFF as in the previous analysis.

- Comparison on the cycles starting at 00Z, 06Z, 12Z, 18Z.

- BIAS in screen temperature strongly reduced in the cycling DA, over all forecast ranges.

- DA switch OFF in full cycling approach, BIAS nearly constant over the lead time.

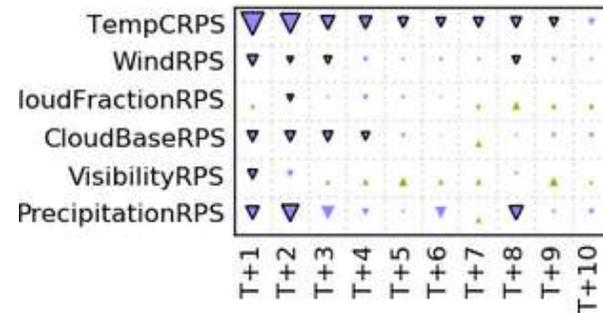- Much lower RMSE using cycling DA as well.

LSB impact

- LSB has impact in the later leadtime (T+12-T+24).

- To understand the impact of the high-resolution observation on the forecast lead time, we can compare ALL_OFF+LSB vs CONTROL. Where both runs has the LSB switched on. (e.g. Theta at 1000m)

- The two increments are added one to each other, in some cases they can have contrasting interaction.

- Sometimes Jo spikes reduced when LSB is used (red line).
  - Lower distance from OBS, LSB leads VAR to give more importance to the background. The analysis increments are lower, and the initial state should be more stable.
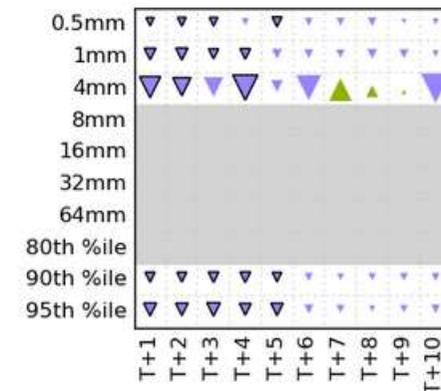
UKV-Var increments	LSB increments
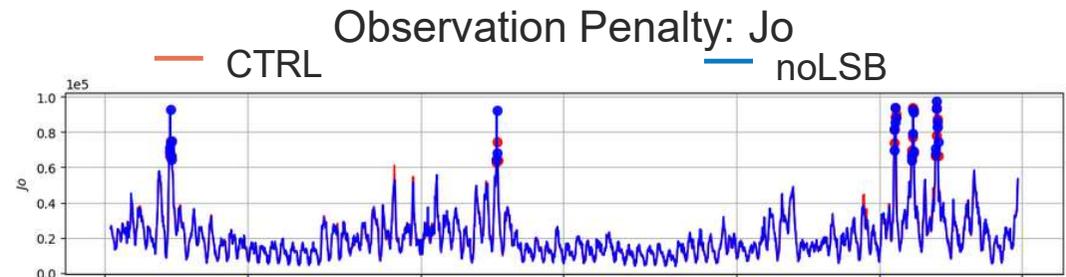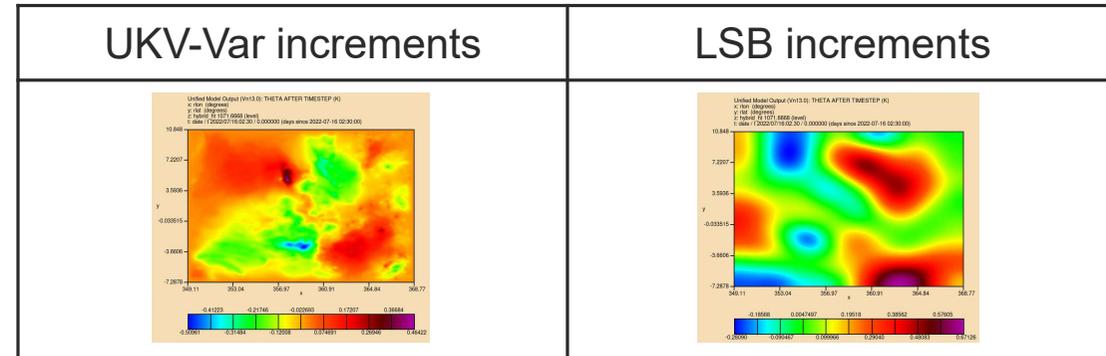
Observation Penalty: Jo
— CTRL   — noLSB

- In the cycle-by-cycle approach, the impact is not cumulative with the trialling and affects equally each forecast generated. Differences can be small and more difficult to interpret as the dump file contains information of the increments added at T-1 in the fully cycling DA suite (the Control).

- The LSB is applied only on four cycles (03Z, 09Z, 15Z, 21Z). To verify the impact of the LSB we need to do the verification on the forecasts starting on these cycles.

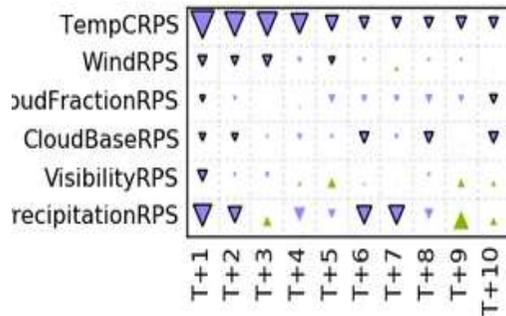winter ALL OFF + LSB vs Control
- One month of data
- HIRA (11 grid lengths)
- Hourly FSS (25 grid lengths)

- From the reproducibility suite analysis, our recommendations for the lead time impact of DA:
  - Can vary strongly between seasons and between fields (more sensitive in summer and for temperature).
  - Does not seem to vary strongly with weather regimes.
  - ALL_OFF+ LSB vs CONTROL max critical lead time: summer = T+3 - T+4; winter = T+5 - T+8.
- From a full cycling DA suite analysis, the differences associated with ALL_OFF can persist well into the end of the run.

≋Met Office

- In the cycle-by-cycle approach, the impact is not cumulative with the trialling and affects equally each forecast generated. Differences can be small and more difficult to interpret as the dump file contains information of the increments added at T-1 in the fully cycling DA suite (the Control).

- The LSB is applied only on four cycles (03Z, 09Z, 15Z, 21Z). To verify the impact of the LSB we need to do the verification on the forecasts starting on these cycles.

summer ALL OFF + LSB vs Control
- One month of data
- HIRA (11 grid lengths)
- Hourly FSS (25 grid lengths)
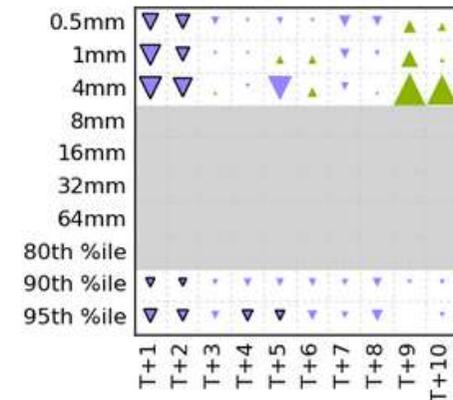
- From the reproducibility suite analysis, our recommendations for the lead time impact of DA:
  - Can vary strongly between seasons and between fields (more sensitive in summer and for temperature).
  - Does not seem to vary strongly with weather regimes.
  - ALL_OFF+ LSB vs CONTROL max critical lead time: summer = T+3 - T+4; winter = T+5 - T+8.
- From a full cycling DA suite analysis, the differences associated with ALL_OFF can persist well into the end of the run.
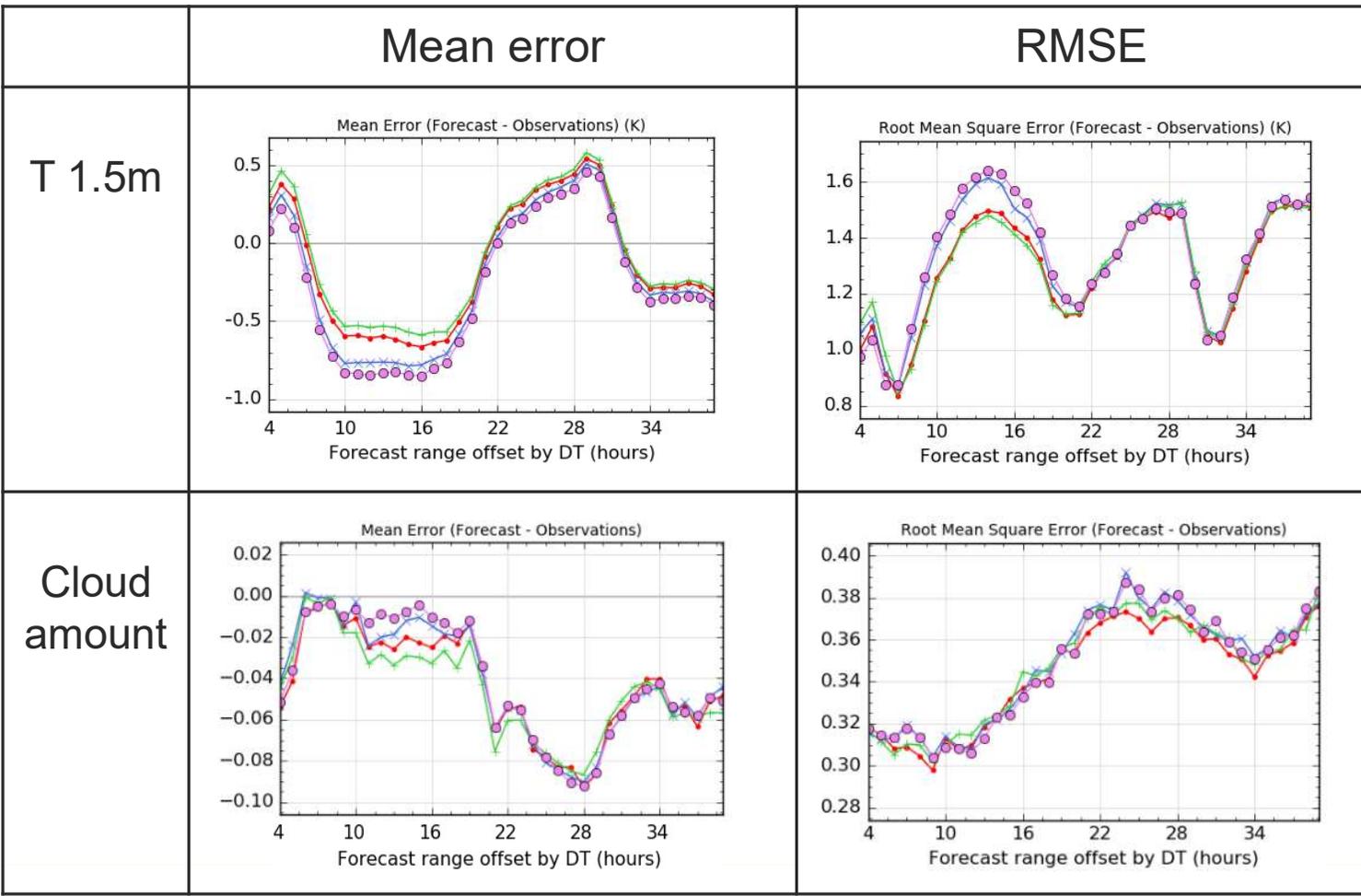
Summer 03Z cycle-by-cycle

Met Office

Legend: CONTROL, ALL_OFF, ALL_OFF+LSB, ALL_OFF+VAR+SURF

	Mean error	RMSE
T 1.5m	Mean Error (Forecast - Observations) (K)	Root Mean Square Error (Forecast - Observations) (K)
Cloud amount	Mean Error (Forecast - Observations)	Root Mean Square Error (Forecast - Observations)

- Diurnal impact, run started at 03 Z.

- UKV-Var and Land Surface DA seem to nudge the forecast in the right direction for cloud skills (low bias).

- This seems to have consequences on the temperature bias, where the strong negative bias from Control is accentuated from UKV-Var. Or is the daytime a consequence of the colder initialization?

- This link between Cloud forecast and temperature can lead to wrong interpretation of the DA impact during the forecast (summer).

Winter 03Z cycle-by-cycle

	Mean error	RMSE
T 1.5m	Mean Error (Forecast - Observations) (K)	Root Mean Square Error (Forecast - Observations) (K)
Cloud amount	Mean Error (Forecast - Observations)	Root Mean Square Error (Forecast - Observations)

Legend: CONTROL, ALL_OFF, ALL_OFF+LSB, ALL_OFF+VAR+SURF
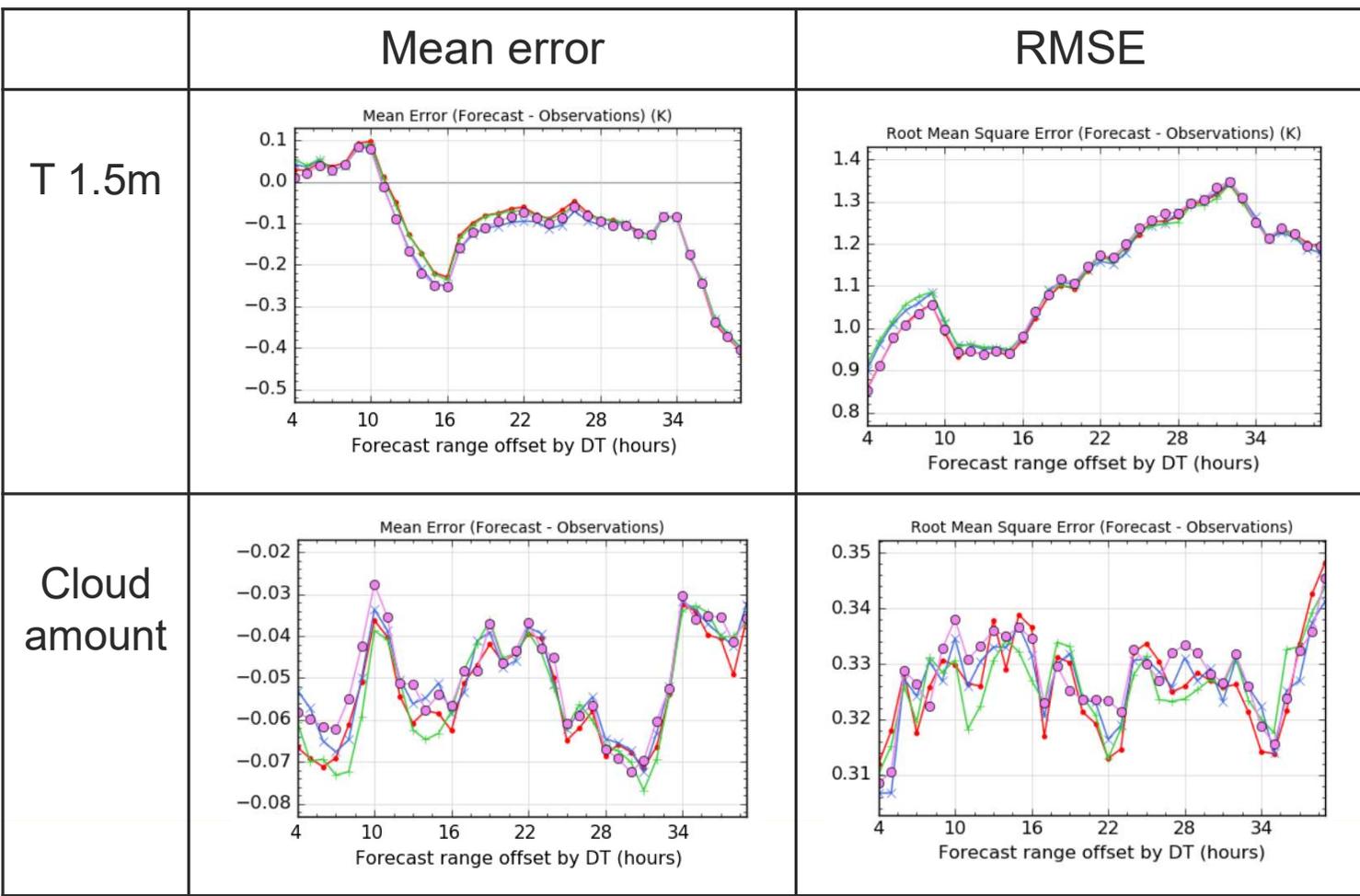
- Diurnal impact, run started at 03 Z.

- Cloud bias: UKV-Var and Land Surface DA tend to reduce the negative bias in the early hours of the forecast.

- For screen temperature the effect in winter is much less visible, probably due to the different meteorological conditions.

CONCLUSIONS

≋ Met Office

- We analysed a winter (December 2021) and a summer (July 2022) month using different sensitivity tests.
- The DA impact => using dump retrievals from a fully cycling DA suite.
- Cycle-by-cycle experiments affects equally each forecast generated. Differences can be small and more difficult to interpret as the dump file contains information of the increments added at T-1 in the fully cycling DA suite. The Control retrieves and uses the dump file and all increments.
- Two main sensitivity experiments were made: ALL_OFF and ALL_OFF+LSB. The impact of having LSB on is non-negligible and can affect the decisions regarding lead time impact of the DA forcing.
- From the reproducibility suite analysis, our findings for the lead time impact of UK-DA are:
  - Can vary strongly between seasons and between fields (summer more sensitive and temperature more sensitive).
  - Does not seem to vary strongly with regimes or cases.
  - Small differences were shown to persist during the forecast.
  - On average, based on the Hinton plots, ALL_OFF vs CONTROL max critical lead time: ==summer = T+10-T+15 ; winter = T+18 – T+24.==
  - On average, based on the Hinton plots, ALL_OFF+ LSB vs CONTROL max critical lead time: ==summer = T+3-T+4; winter =T+5-T+8.==
- ==From a full cycling DA suite analysis, the differences associated with ALL_OFF can persist well into the end of the run. Further analysis shows that the strongest impact is up to T+12. Beyond this, this is more case and regime dependent where the impact can last up to T+30.==
- ==The longevity of DA impact is likely to be related to the nature and magnitude of the model error.==

Thank you very much