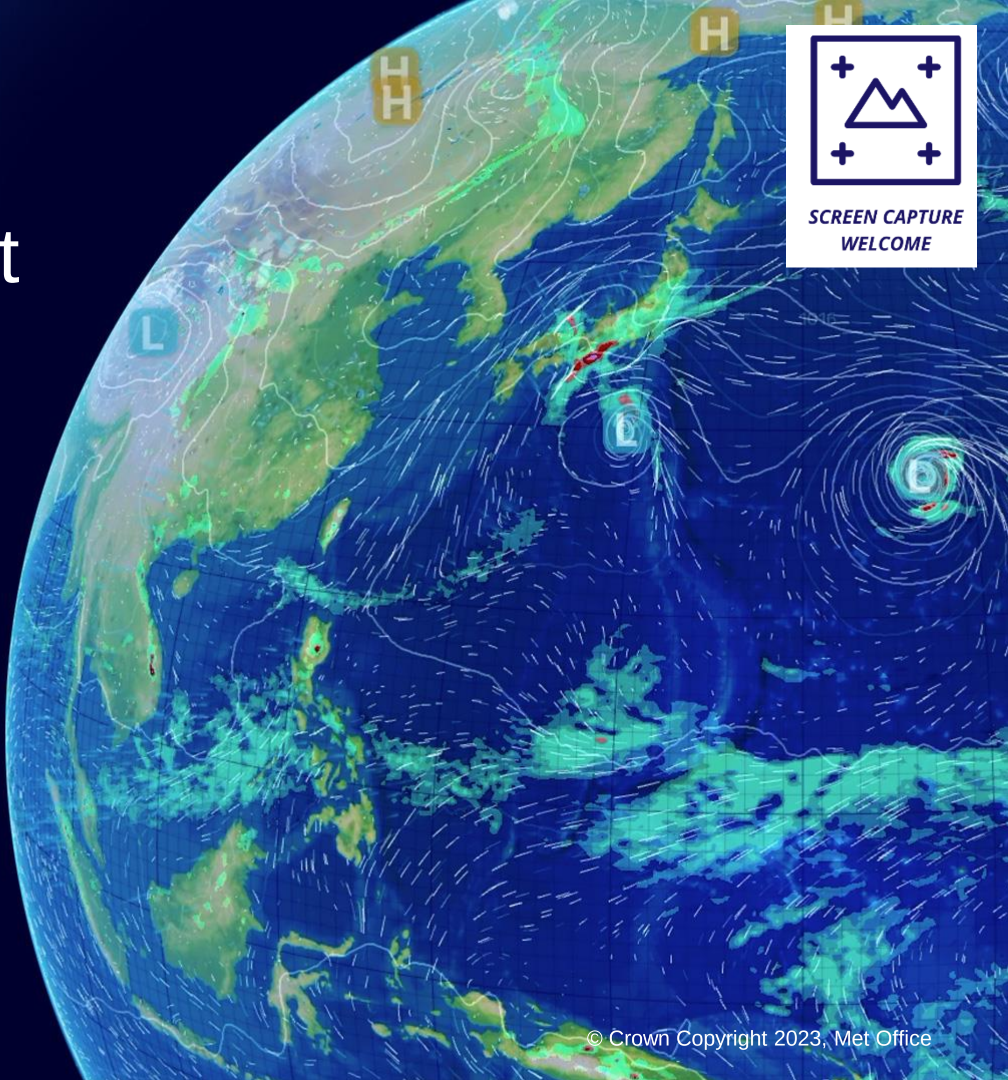
Met Office

Improving the blend of multiple weather forecast sources by Reliability Calibration

Gavin Evans, Fiona Rust, Benjamin Ayliffe, Ben Hooper

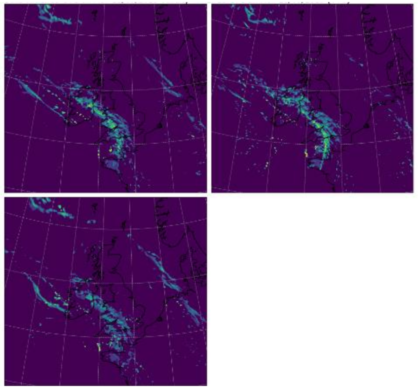
Tuesday 5th September

European Meteorological Society Conference 2023, Bratislava

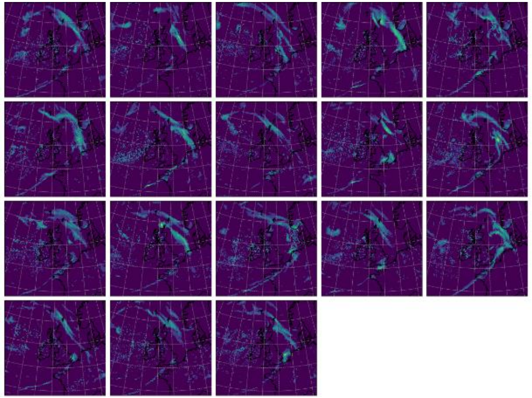


SCREEN CAPTURE WELCOME

Met Office ensemble models



MOGREPS-UK
Out to 5 days
Available every hour
3 members



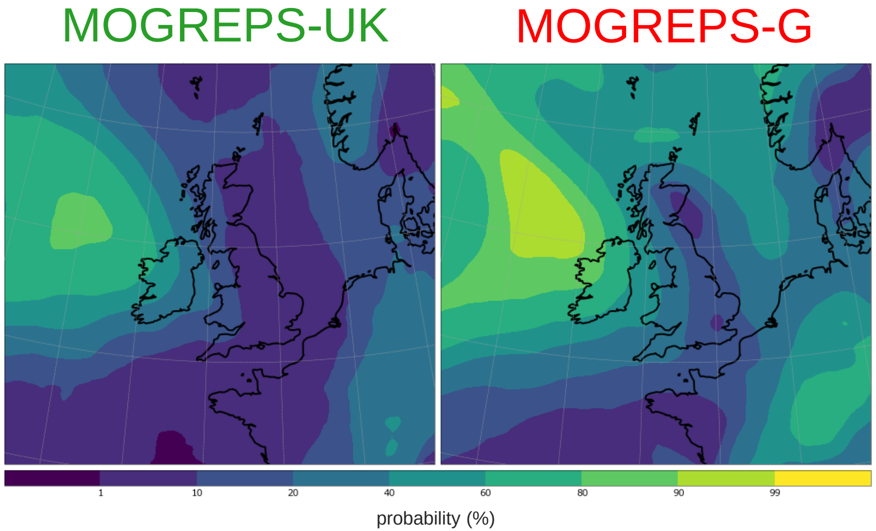
MOGREPS-G
Out to 8 days
Available every 6 hours
18 members

Met Office

Create a seamless probabilistic forecast

High resolution, convective-scale models (MOGREPS-UK) resolve meteorological processes that are parametrized in lower resolution models (MOGREPS-G).

Transitions between models can introduce jumps and inconsistencies.



Met Office

How to improve the blend?

Calibrate MOGREPS-UK and MOGREPS-G using a common truth (high resolution deterministic model analysis - UKV).

Constraints:

- Use a rolling training period to fit within existing IMPROVER infrastructure. A large archive of the inputs to calibration is not available.
- We perform probabilistic blending and therefore it is more difficult to implement a method acting on ensemble members e.g. quantile mapping.
- Want to apply to a range of diagnostics.

Strategy

- Implement reliability calibration (Flowerdew 2014).

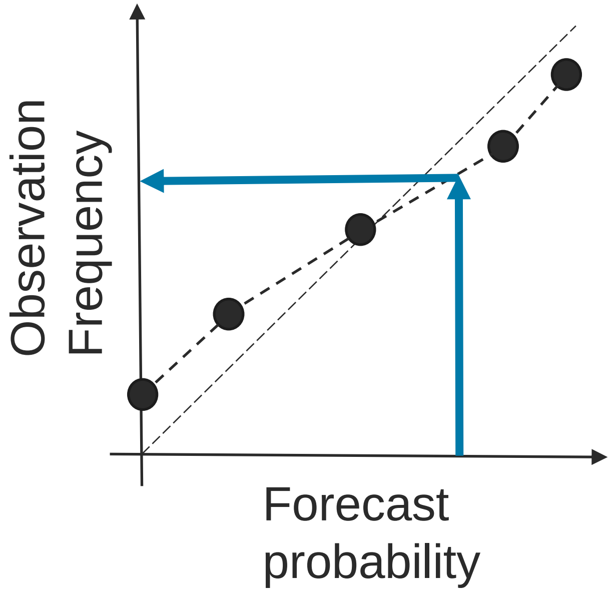
Flowerdew, J. (2014). Calibrating ensemble reliability whilst preserving spatial structure. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *66*(1). https://doi.org/10.3402/tellusa.v66.22662

What is Reliability Calibration?

Build a reliability diagram for each threshold based on training data.

Calibrate today's probabilistic forecast using the reliability diagram.

Corrects each threshold separately.



Flowerdew, J. (2014). Calibrating ensemble reliability whilst preserving spatial structure. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *66*(1). https://doi.org/10.3402/tellusa.v66.22662

Met Office

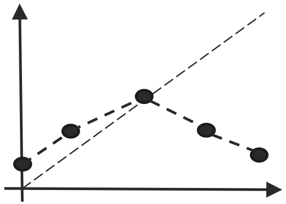
This study

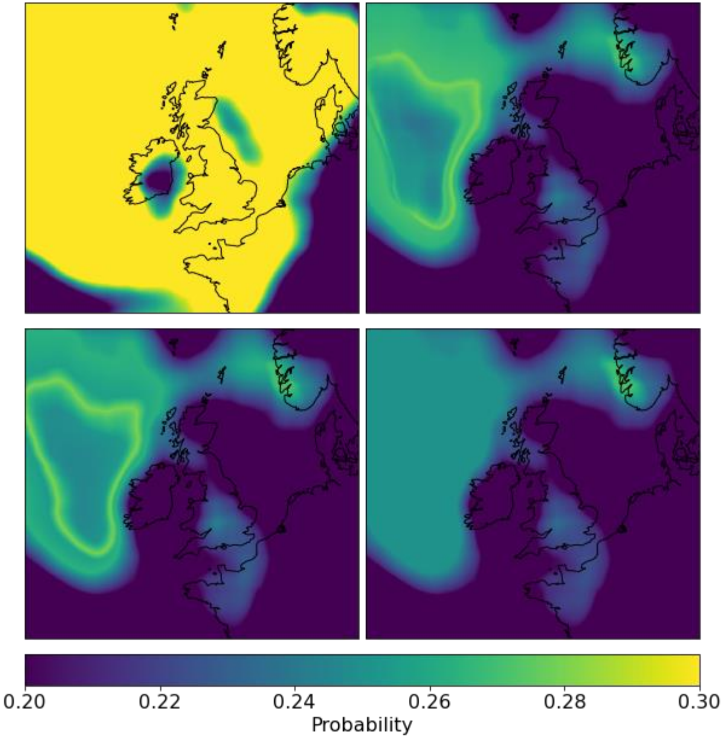
- 30-day training period for total cloud cover and precipitation rate.
- Reliability tables aggregated over all grid points. Separate calibration for each cycletime and leadtime.
- Extend handling of edge cases relative to Flowerdew (2014) following assessing reliability diagrams from training period.

Rust, F. M., Evans, G. R., & Ayliffe, B. A. (2023). Improving the blend of multiple weather forecast sources by Reliability Calibration. Meteorological Applications, 30(4), e2142. https://doi.org/10.1002/met.2142

Met Office

Adaptions to Reliability Calibration



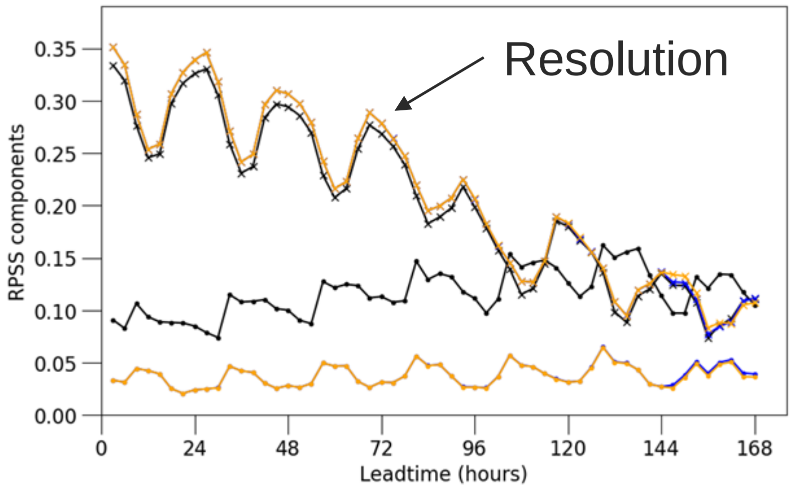


- Original reliability diagram could result in unrealistic lower probabilities in central region.
- Instead we combine bins with:
  - very low sample counts
  - that have non-monotonic observation frequency.
- Enforce monotonicity to avoid edge cases where combining bins is not sufficient.

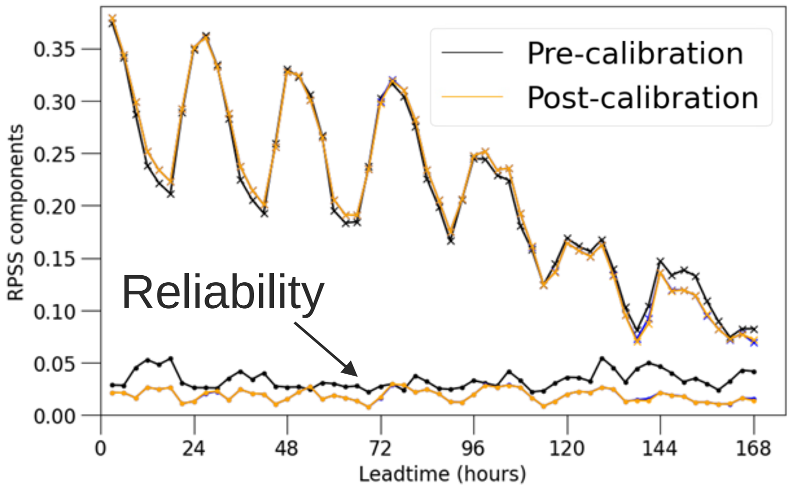
Precipitation rate > 0.03 mm/hr

Ranked Probability Skill Score components – total cloud cover

February, March



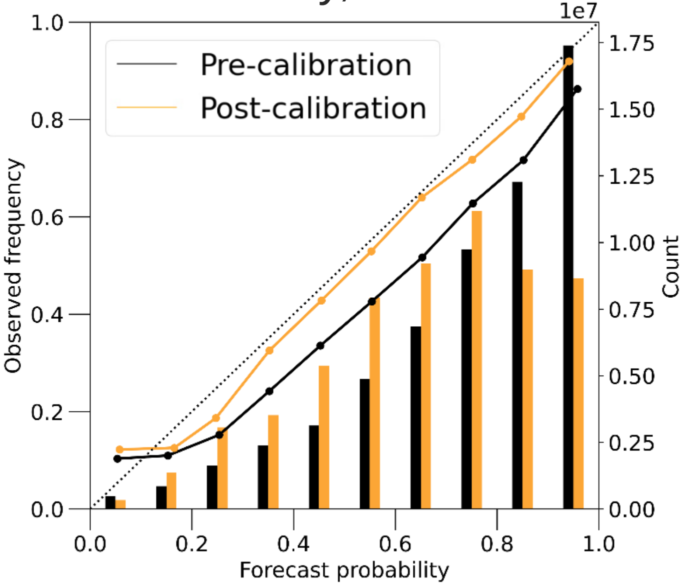
June, July



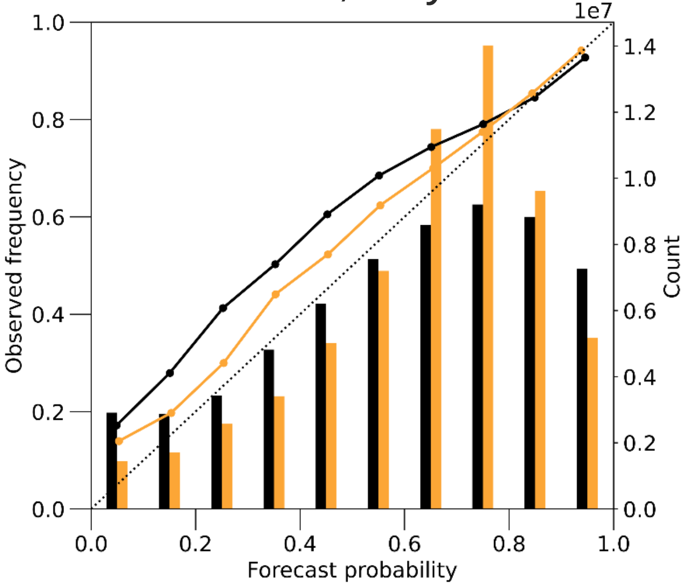
The large reliability penalty term is reduced in February, March.
There is a smaller reduction in June, July.

Verification reliability diagram – total cloud cover

February, March



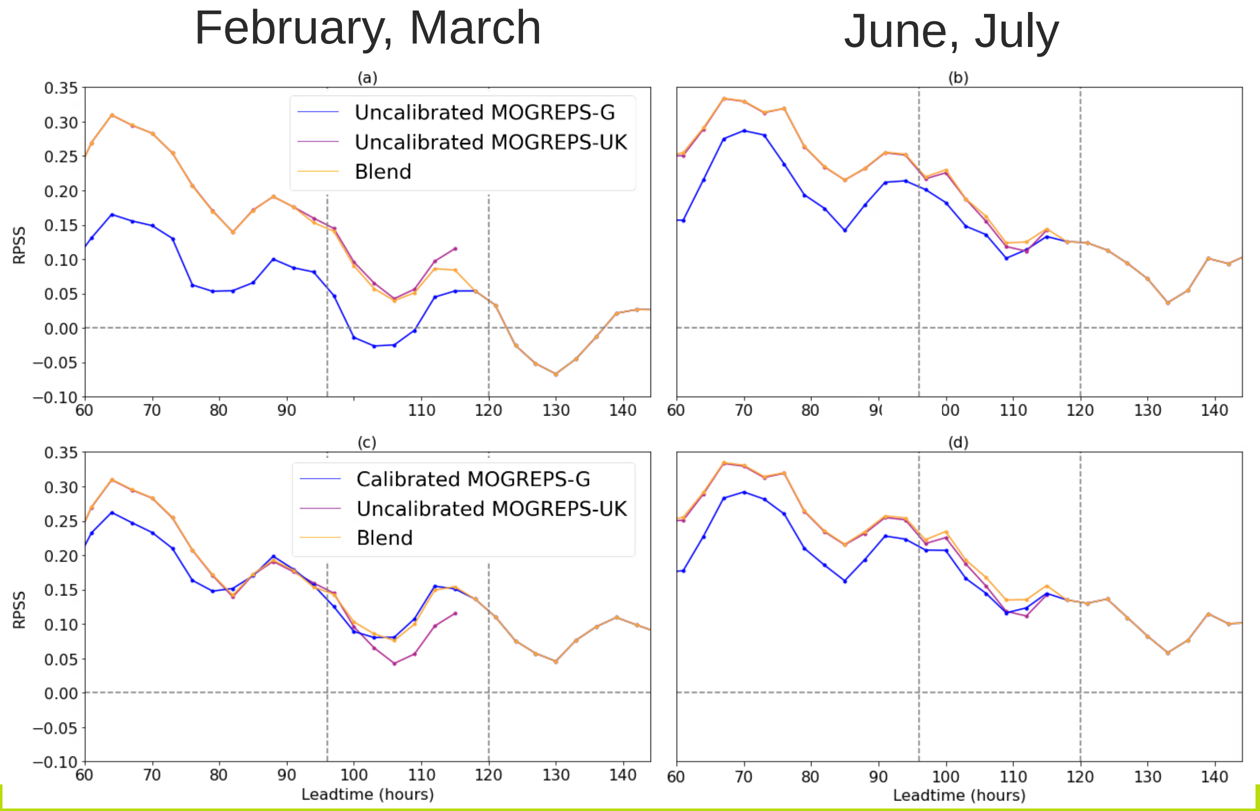
June, July



Reliability improved following calibration

Ranked Probability Skill Score of blended forecasts – total cloud cover

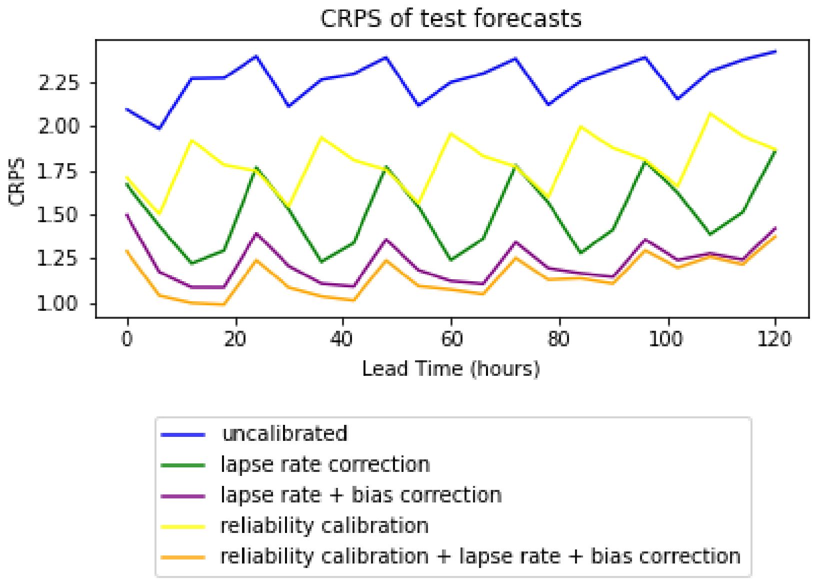
- Improvement in RPSS for MOGREPS-G
- Improvement in RPSS across blending window



Met Office

Applying reliability calibration to temperature

- Part of [Demaeyer et al., 2023]
- Reliability Calibration particularly suited to diagnostics with more complex distributions, so not typically used for temperature.
- Reliability Calibration used in conjunction with other complementary steps.
  - Lapse rate correction
  - Deterministic bias correction
  - Reliability Calibration
- Competitive with other approaches



Summary

- Reliability calibration has been applied to total cloud cover, precipitation rate and temperature forecasts.
- The configuration from Flowerdew, 2014 has been amended to operational implementation and to avoid artefacts. Details in Rust et al., 2023.
- The reliability calibration is producing more reliable forecast for thresholds that were previously over-forecast.
- The impact of reliability calibration is clear on a case study of the blended forecast.
- Code available in https://github.com/metoppv/improver.

Met Office

Future work

Improvements to usage:

- Training dataset length.
- Choice of truth – currently using UKV analysis.
- Apply to other variables.
- Calibrate convective and dynamic precipitation rate separately.
- Compare with other techniques.

Improvements to algorithm:

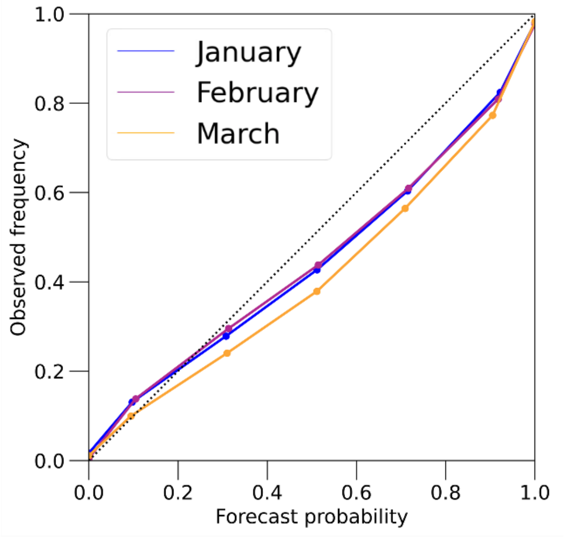
- Local correction i.e. use a reliability diagram per grid point.
- Impact on extremes – do we need explicit extreme handling?
- Transitions from dry training dataset to wet forecast. How will this be handled?

Thank you

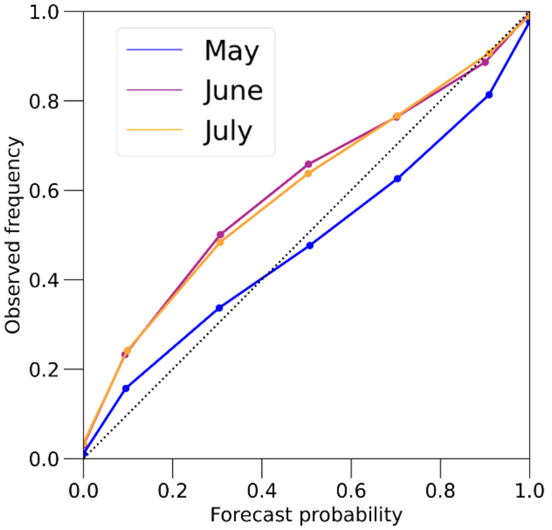
Email: gavin.evans@metoffice.gov.uk

Reliability diagram from training dataset – total cloud cover

January, February, March



May, June, July



Clear over-forecasting signal in Jan/Feb/Mar.

Over-forecasting in May transitions to under-forecasting in June and July.