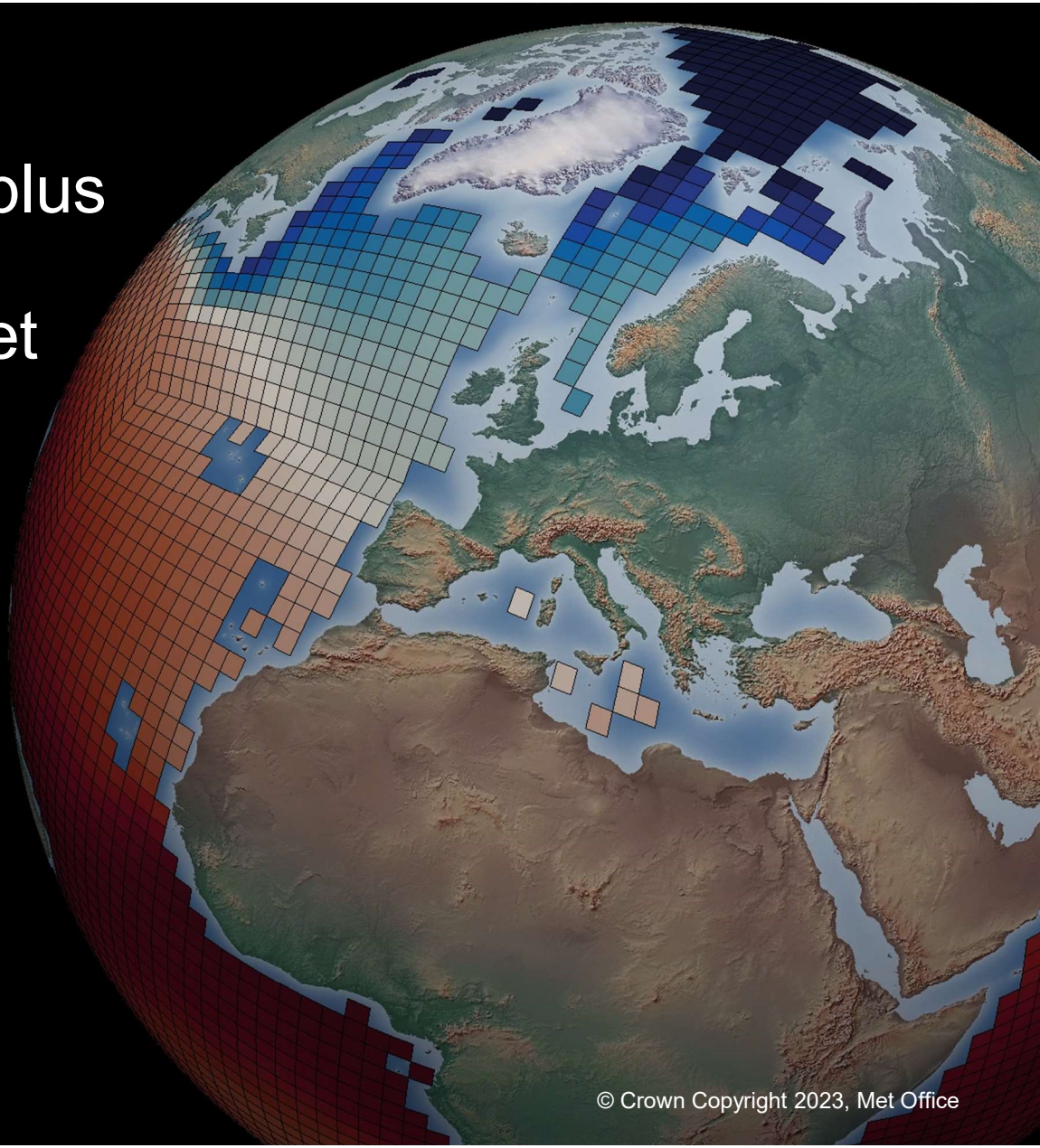
Progress with integrating METplus into the operational and model development process at the Met Office

Marion Mittermaier
Science implementation lead

on behalf of the METplus implementation team

EWGLAM September 2023

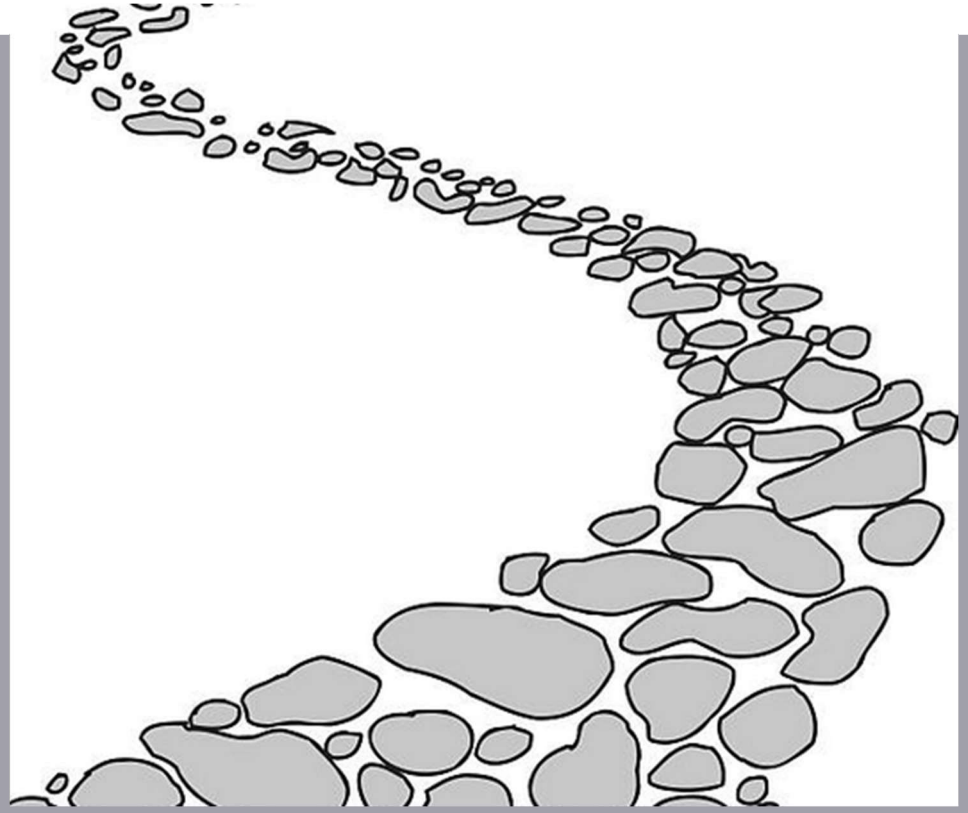


© Crown Copyright 2023, Met Office

Met Office

The long and winding road ….

A brief recap of the setup



Met Office

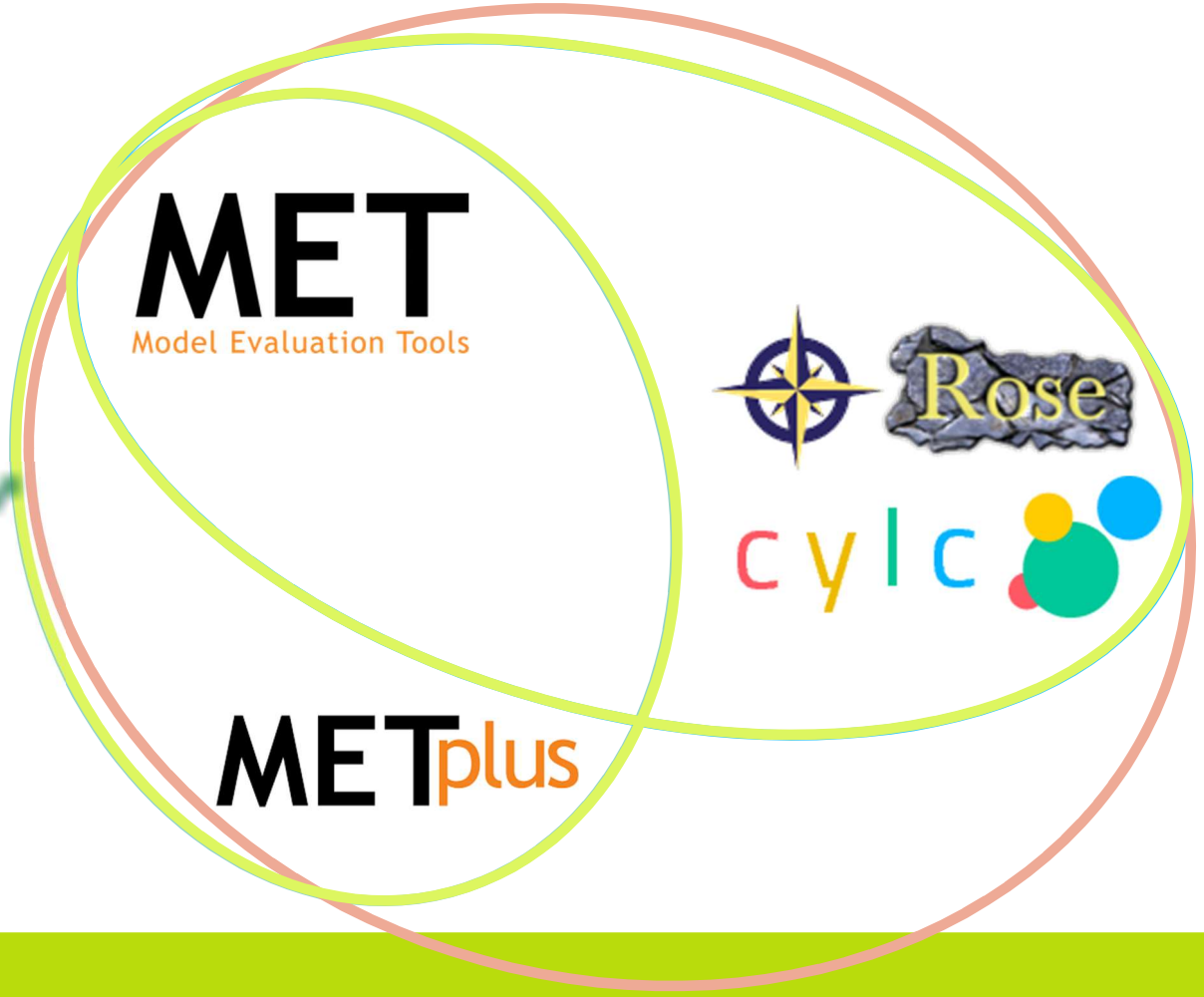
Building a verification workflow

1) METplus wrappers only

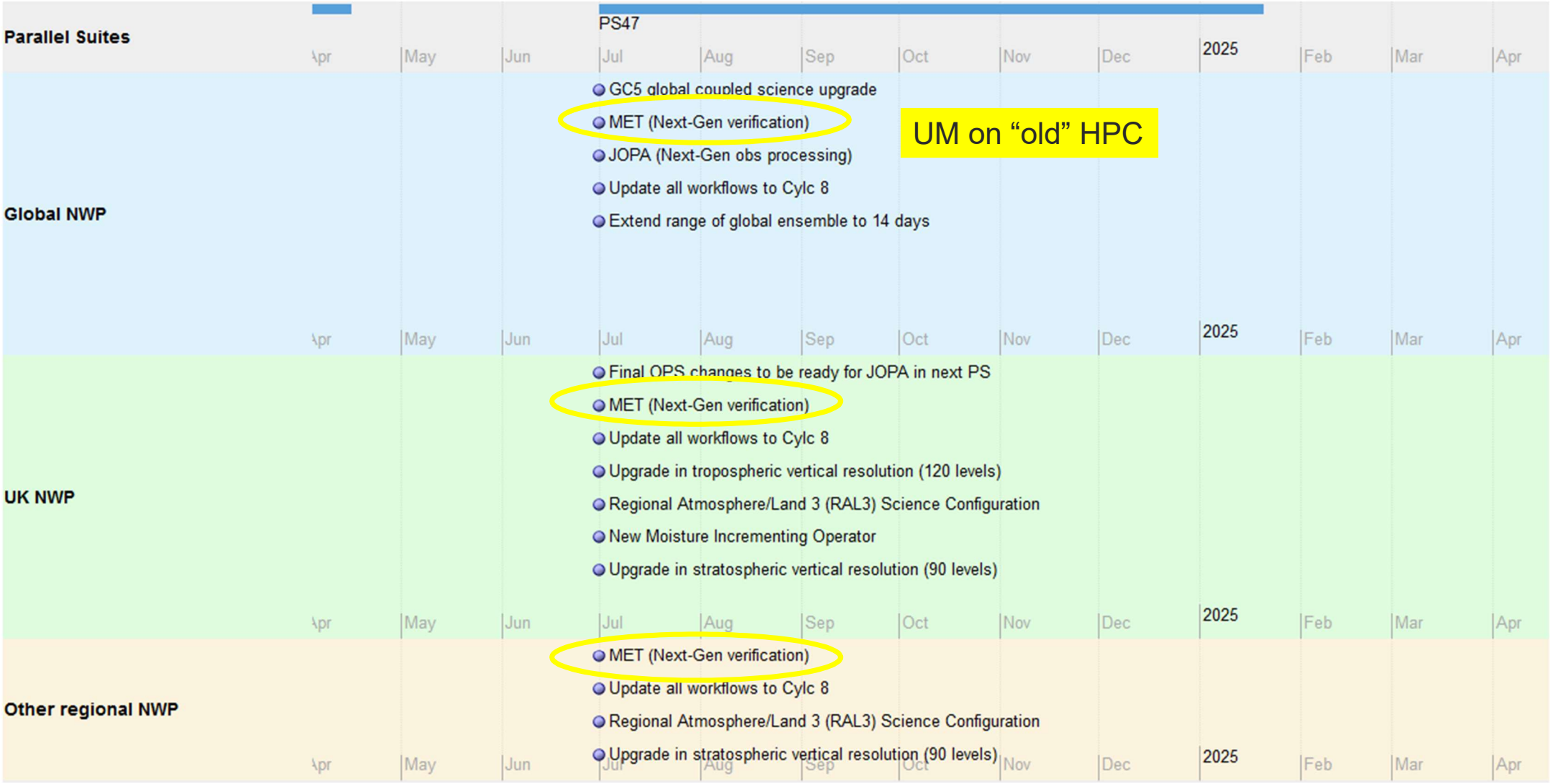
2) Ignore METplus wrapper

3) Incorporate with cylc workflows

4) Use for trials and development



Implementation into operations



Met Office

Since last year we are/have …

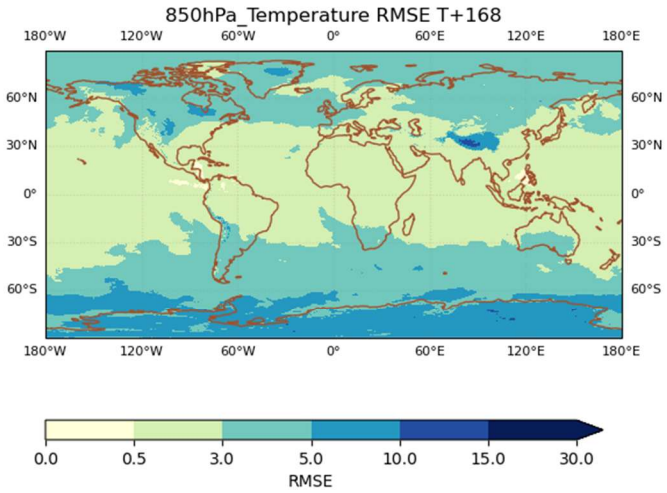
- Tested HiRA for ensemble forecasts
- Testing the ODB as the source of the observations: LNDSYN, SHPSYN, sondes, SATWINDS, AMDAR, AIREPS
- Enhanced decoding of SYNOP messages to extract global precipitation
- Continued preparation for WIGOS identifiers
- Are testing enhancements to current operational capability such as calculating SEEPS against gauges and GPM
- Work has started on UK radar-based verification
- Continuing the process of reconciling the results from the old and new systems
- Approaching “near readiness” for operational parallel running in 2024 (with several caveats!)

© Crown Copyright 2023, Met Office

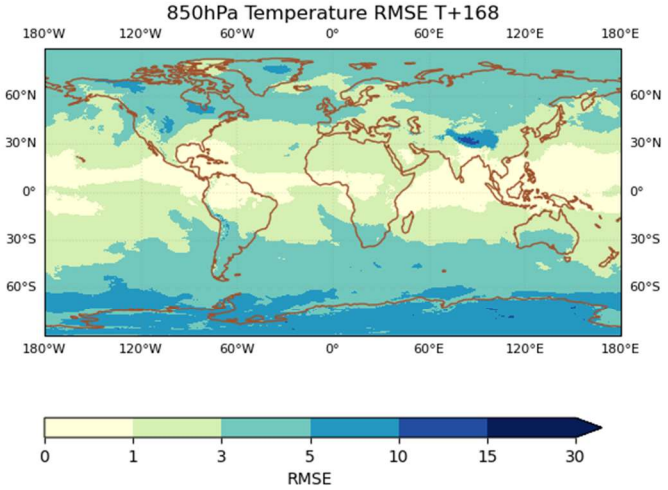
Met Office

Using METplus to evaluate Pangu

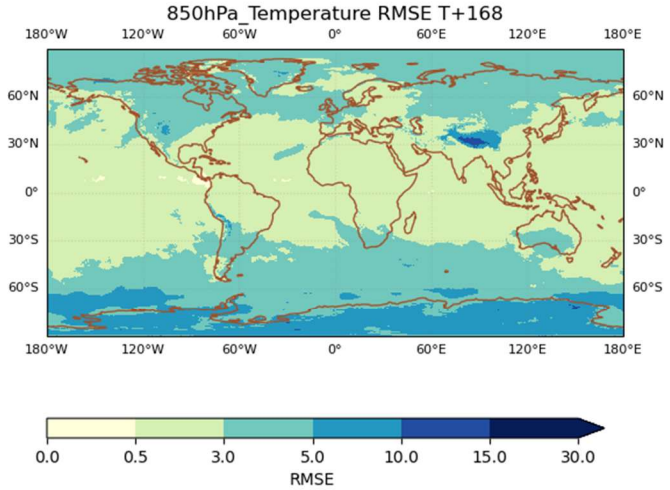
Pangu_GM v GM Analysis



GM v GM Analysis



Pangu_GM v GM



850 hPa Temp RMSE T+168

See Seb Cole's talk in this session

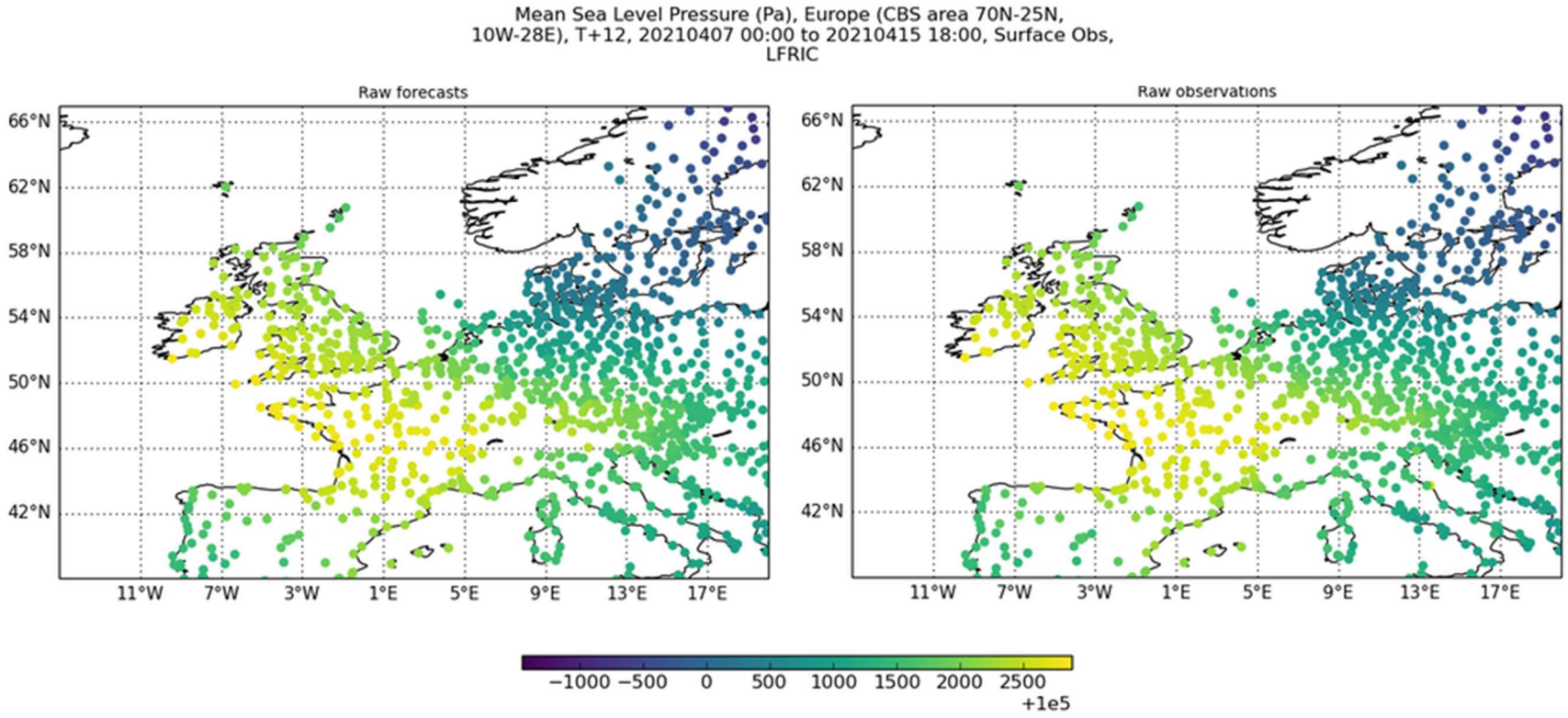
© Crown Copyright 2023, Met Office

Met Office

LFRic specific capability

- Internal regridding from unstructured to regular grids is under development at NCAR DTC (using the atlas library)
- All *gridded* LFRic evaluation at this point is based on externally regridded forecasts.
- *Point-based* verification does now have the capability of extracting the nearest native (cube-sphere) grid point to match to an observation.

LFRic verification against point obs



We now have the ability to extract the nearest unstructured grid point from the grid to compare to obs.

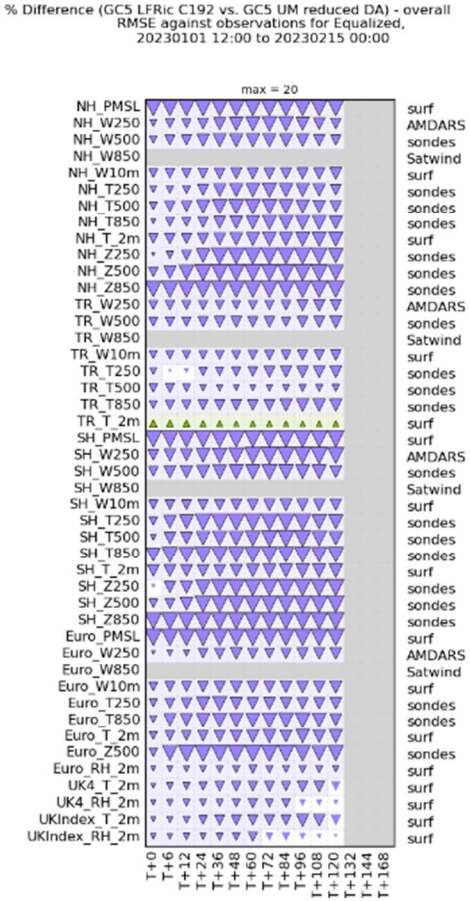
Initial global LFRic DA trial (Feb 2023) scorecard

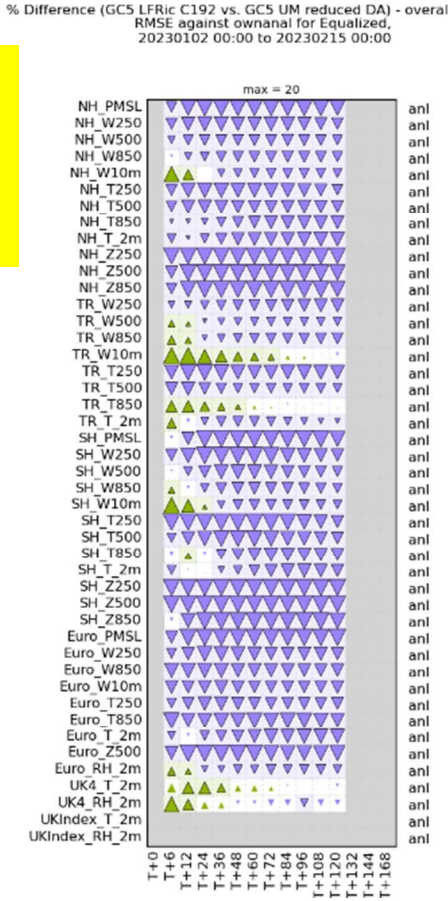
A significant bug was found in the model configuration set up which made the results a lot worse than they should've been

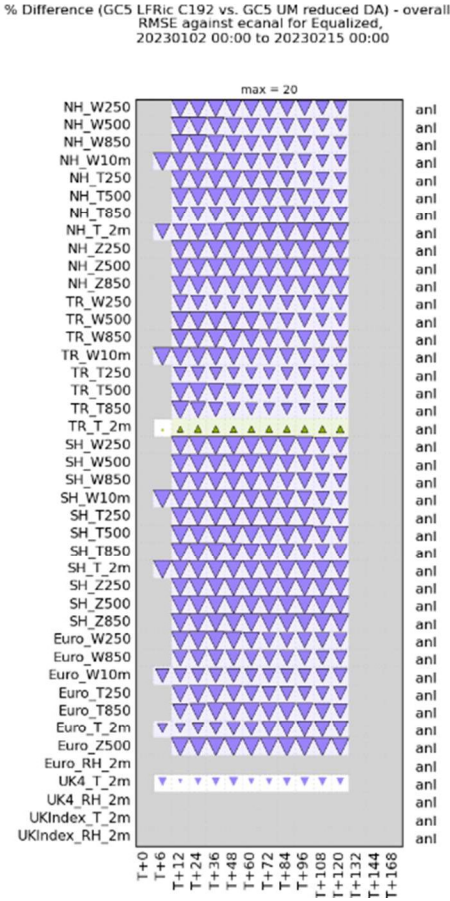
Introduction | Deterministic Summary | Deterministic Verification | Deterministic Assimilation Stats | Documentation

Global Evaluation

Based on regridded LFRic and UM to the 2.5° CBS grid

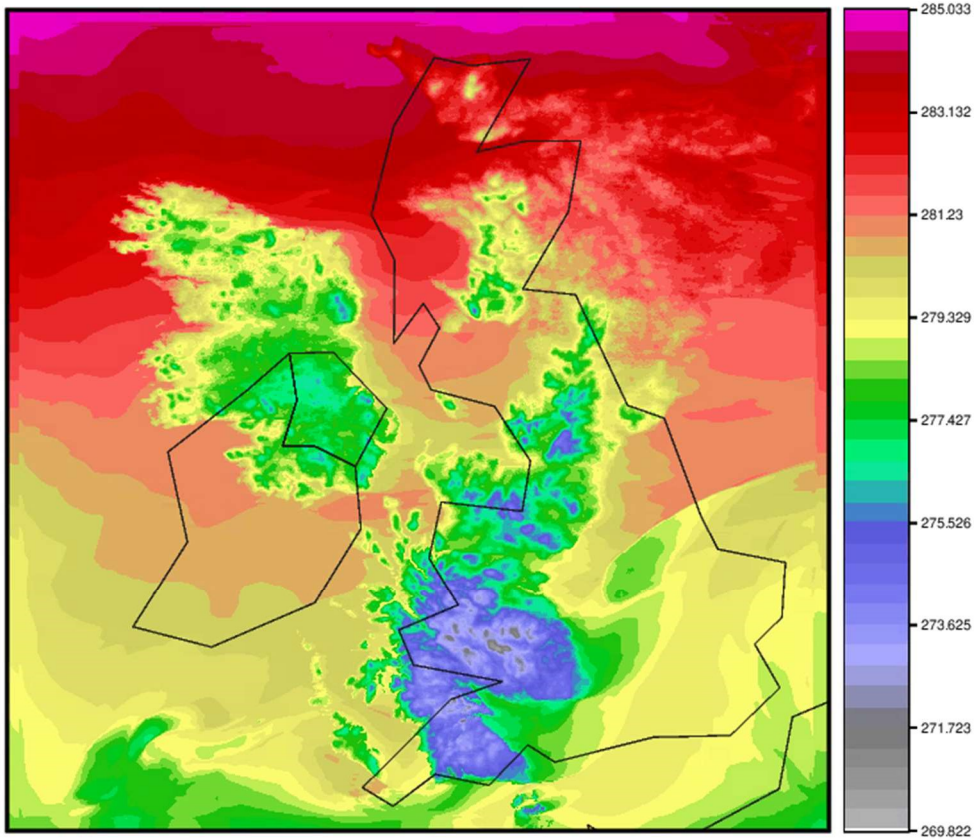




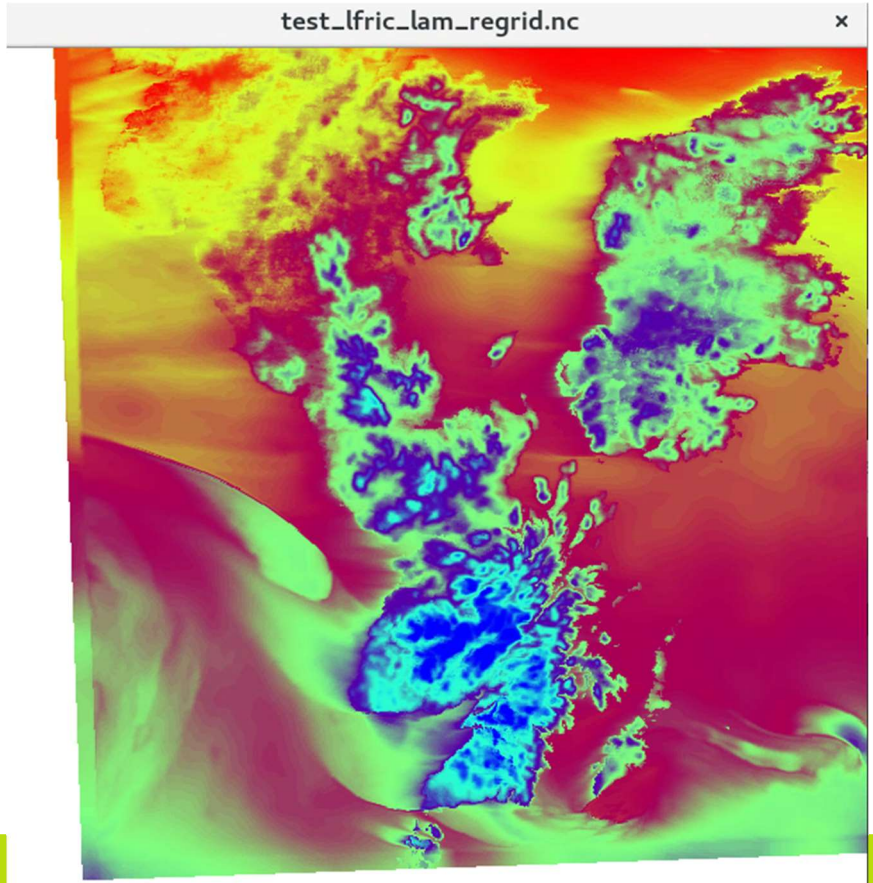


Regional LFRic

Some progress…. MET reads the output from an internal tool called ‘slam’ → but the files are upside down! Regridding also works……..



lfric_struc_0.nc



Met Office

Use case development for R&D

Global LFRic testing has provided the opportunity of developing configuration files for pulling through enhanced capability into the R&D process.

Simple wrappers have now been developed for:

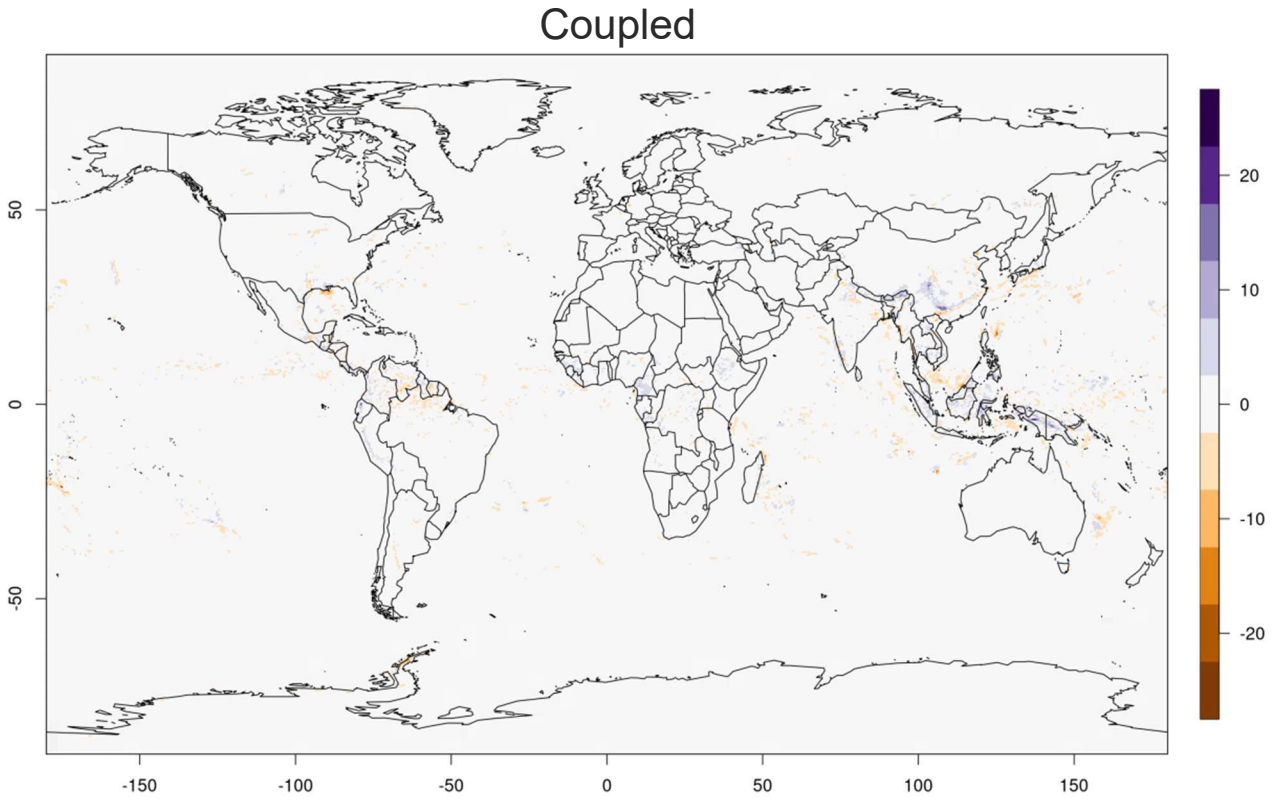
- MODE
- wavelet_stat
- Series_analysis
- GridDiag

This development has also highlighted the need for appropriate visualisation of results before the capability can be made available to users.

Visualisation development activities are planned for this autumn/winter.

Precipitation evaluation of NWP case studies

- 8 coupled and atmosphere-only case studies at N320 (C224) ~ 40 km
- All evaluation was either done LF vs UM (without observations) or against GPM
- Unlike the normal scenario, the **objective** with these comparisons is to **find no (significant) differences** (i.e., the two dynamical cores with nearly identical physics perform very similarly)



Very small differences @ t+24h LF minus UM.

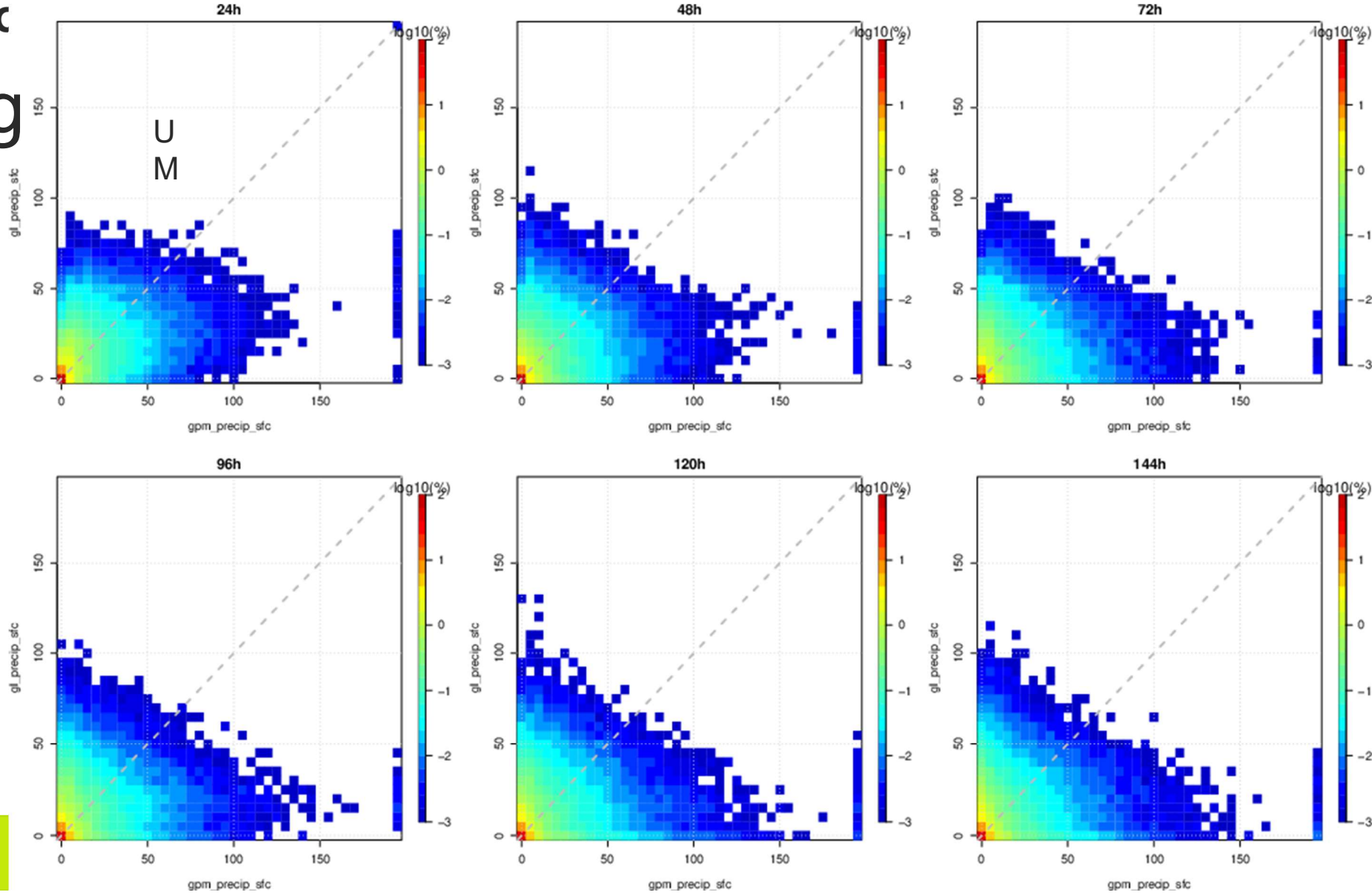
Met Offic

GridDiag

Provides the means of creating the 1-D marginal and 2-D joint (bivariate) distributions between two variables.

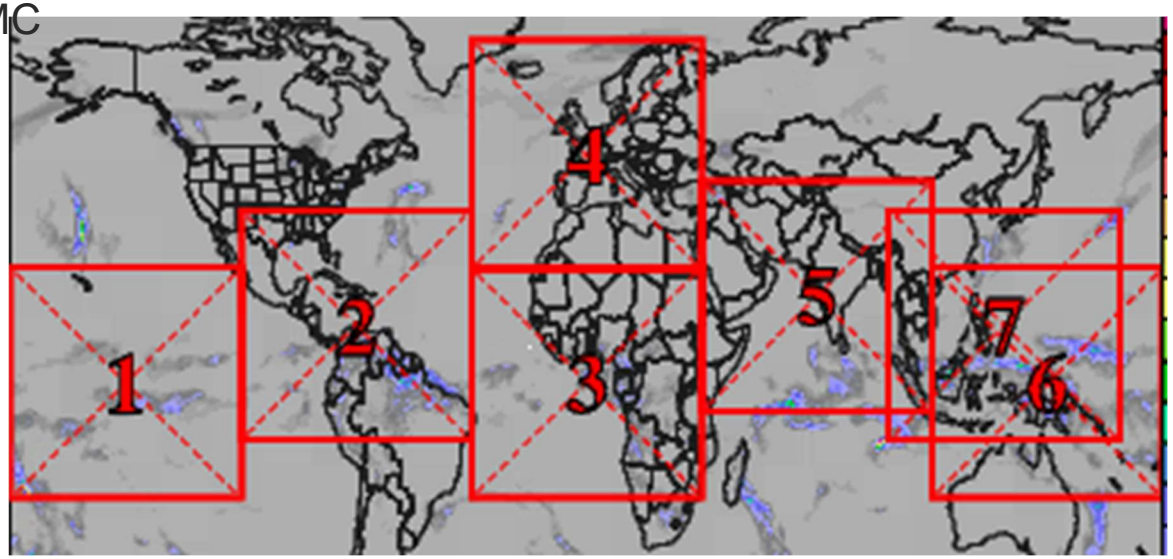
"Perfect" correspondence between two variables would be indicated by only the bins along the diagonal being populated.

Tropics coupled



wavelet_stat

1. C+E Pac, 2. TropAtl, 3. TropAfr, 4. Europe, 5. ME+India, 6. WPac+Oz, 7. MC



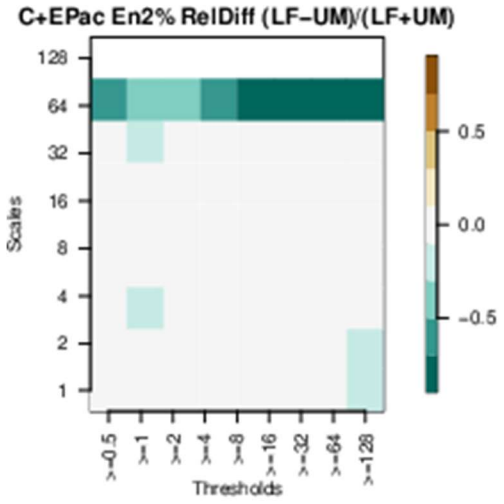
Defined 7 regions of 128 x 128

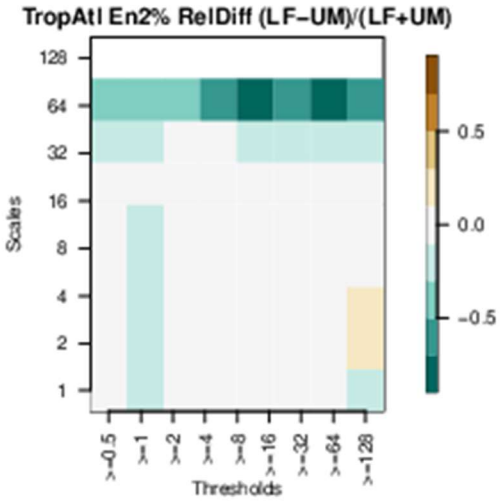
Stay within the latitudes covered by GPM

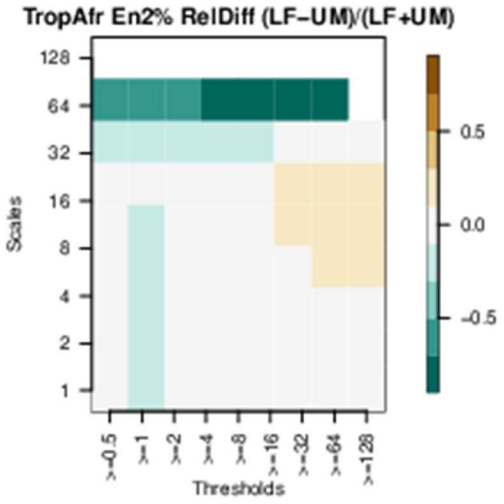
Comparing LF to UM at t+24h only

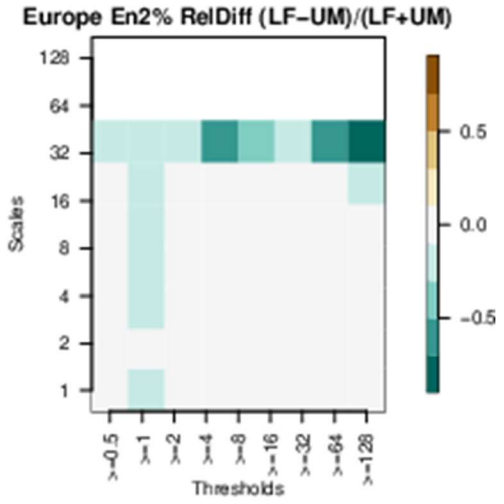
After Casati et al. (2004)

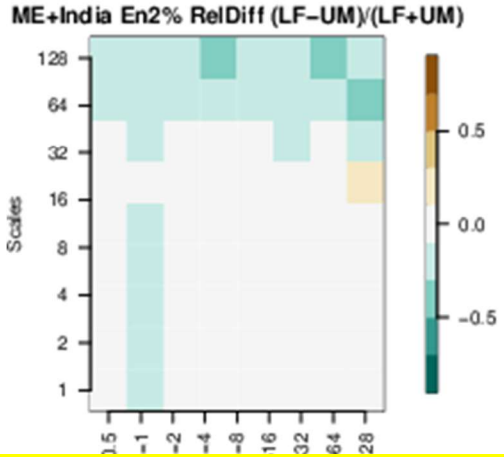
© Crown Copyright 2023, Met Office

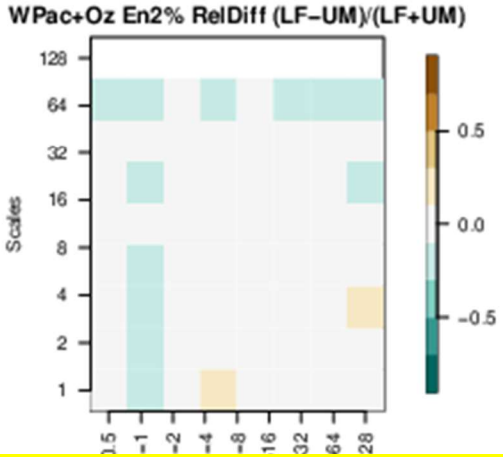


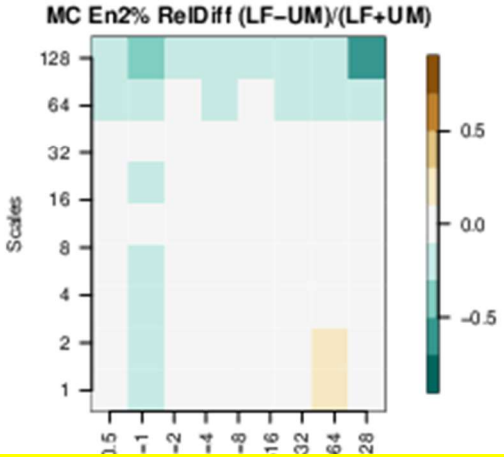








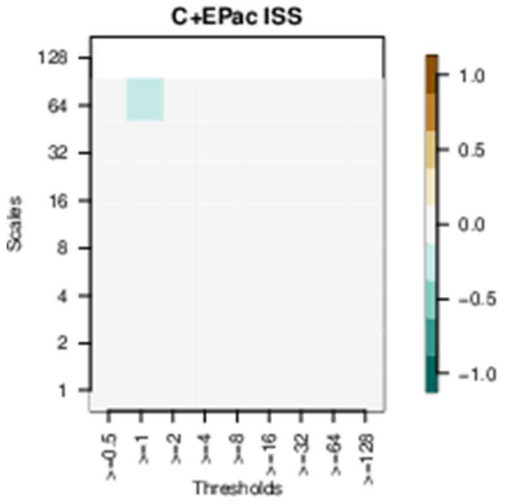


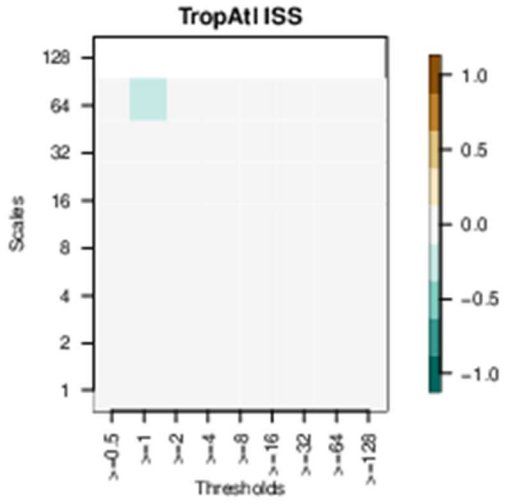


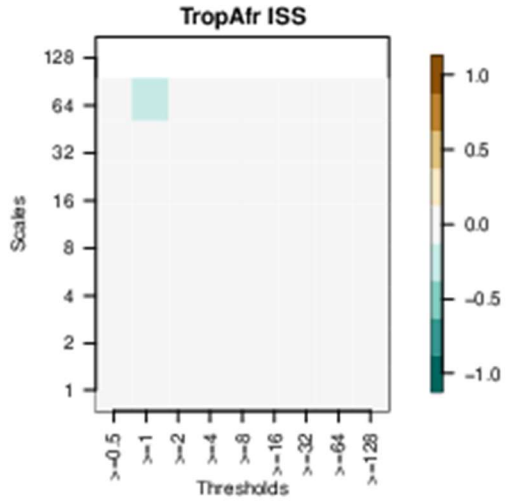
Relative % differences between LF and UM

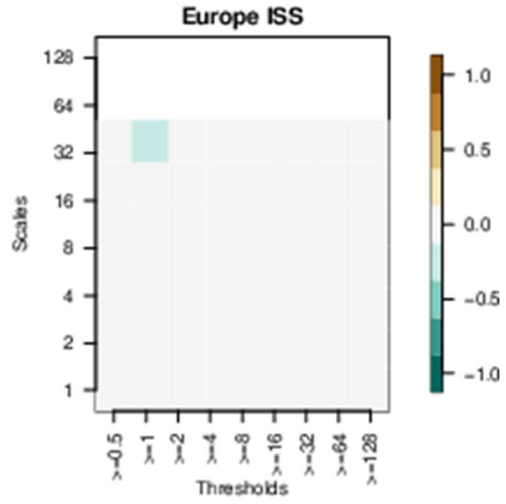
- This highlights more energy at larger scales for all accums

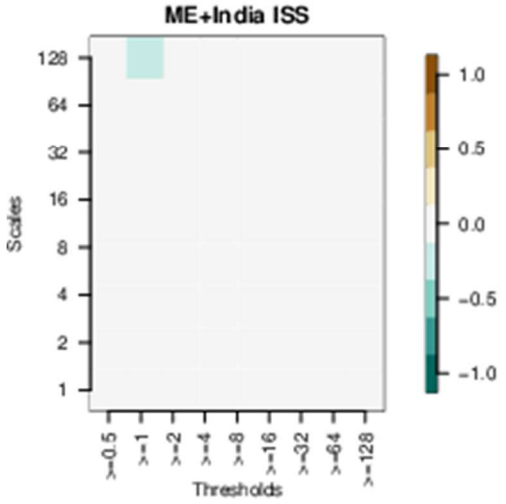
The En2 bias can be assessed via the En2 relative difference, equal to the difference between forecast and observed squared energies normalized by their sum. The En2 relative difference is sensitive to the ratio of the forecast and observed squared energies. As for the MSE, the sum of the energy of the scale components is equal to the energy of the original binary field. Therefore, the percentage that the En2 for each scale contributes to the total En2 may be computed. Typically, low thresholds exhibit most of the energy % on large scales (and less % on the small scales), since low thresholds are associated with large scale features, such as fronts. For higher thresholds, the energy % is usually larger on small scales, since intense events are associated with small scales features, such as convective cells or showers. The comparison of the En2% RelDiffs provides feedback on how the events are distributed across the scales and enables the comparison of forecast and observation scale structure.

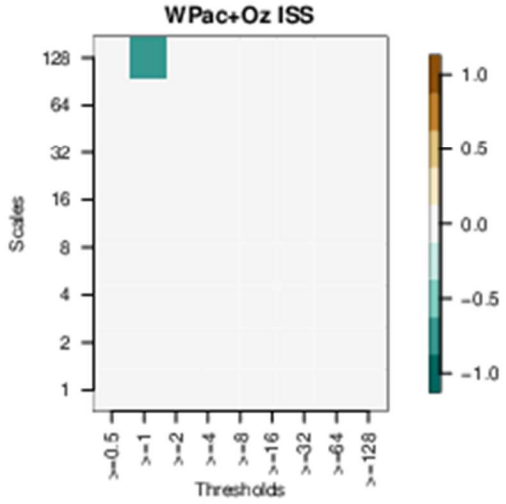


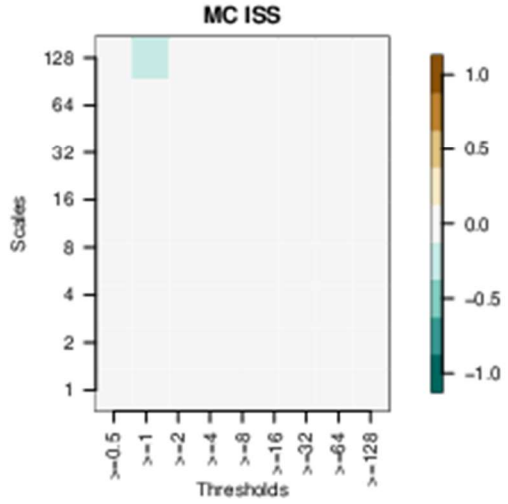








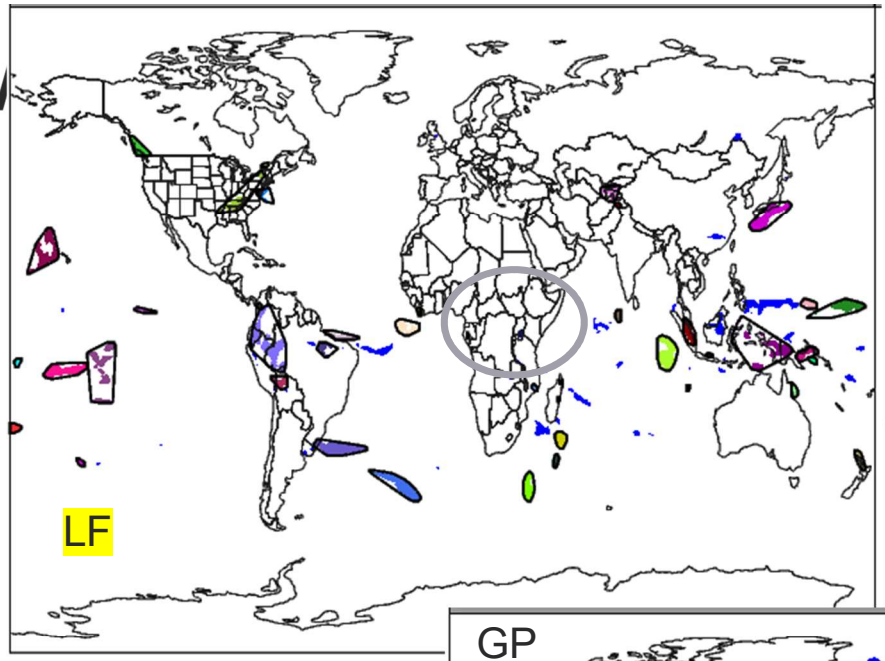


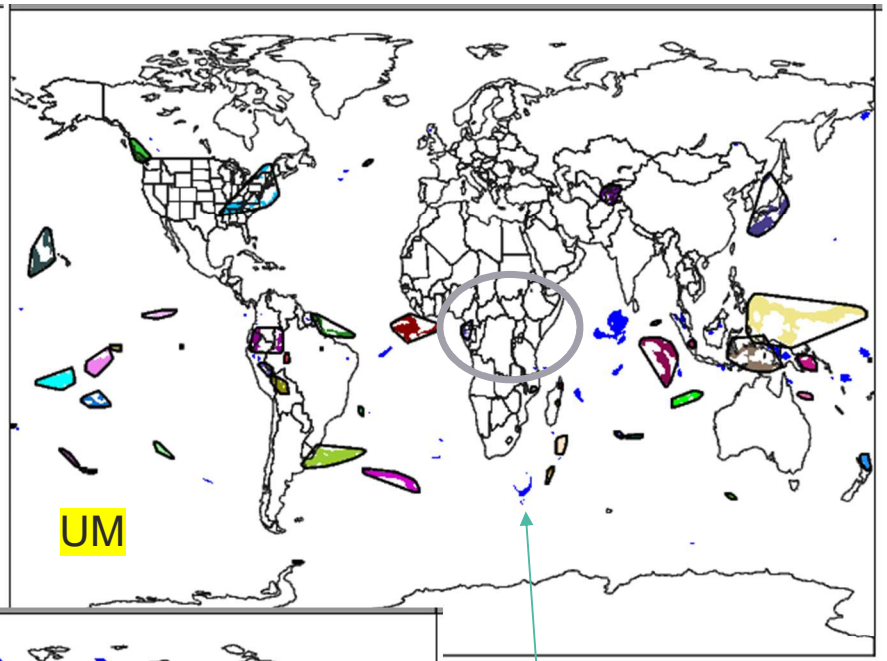


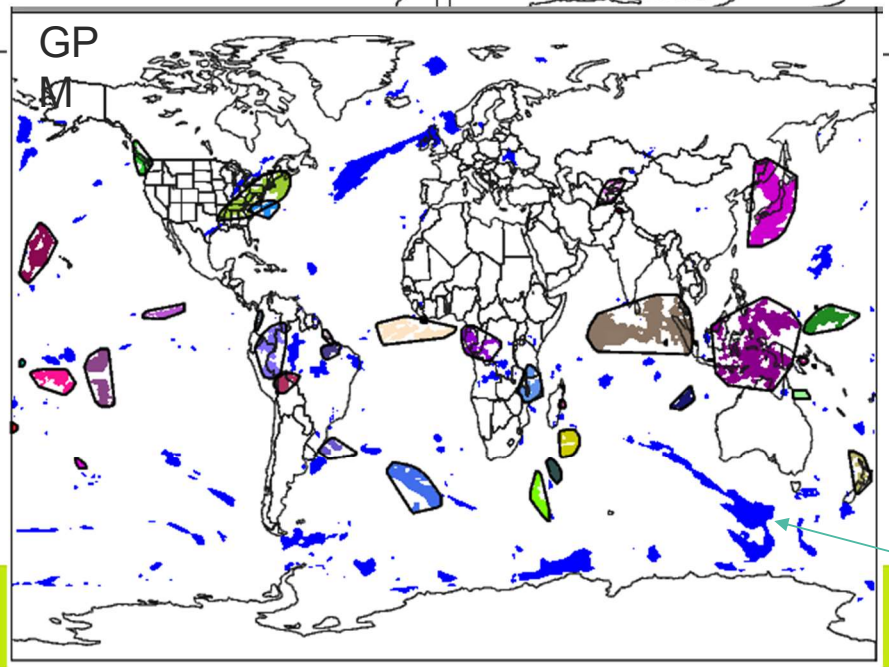
Intensity scale score
LF vs UM

- Largely neutral (good)
- Highlighting some differences at scale 128 ~= 5000 km

The skill score is based on the MSE for each threshold and each scale. It is computed relative to random chance. The Intensity-Scale (IS) skill score evaluates the forecast skill as a function of the precipitation intensity and of the spatial scale of the error. Positive values of the IS skill score are associated with a skilful forecast, whereas negative values are associated with no skill. Usually large scales exhibit positive skill (large scale events, such as fronts, are well predicted), whereas small scales exhibit negative skill (small scale







Graphical output produced when tool is run

2D objects written to netcdf for further manipulation

MODE

- 25 mm/24h objects
- 20210324 120h fcst

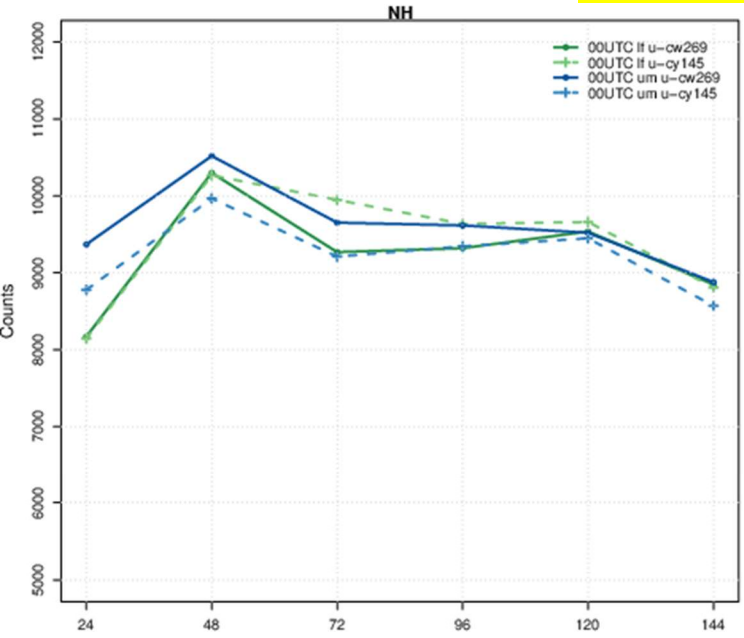
Note colour matches only valid for LFRic
Matches for UM different (though base field the same, i.e., showing only once here)

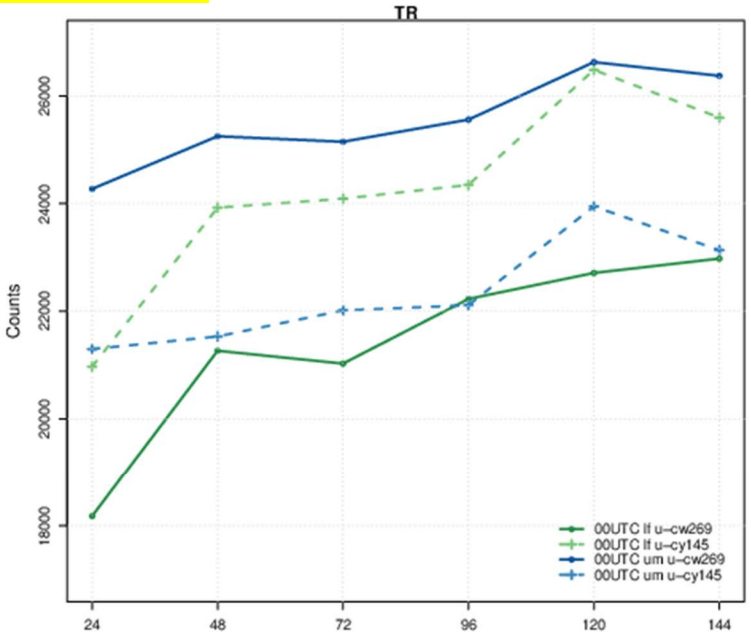
Blue “blobs” with no outlines are

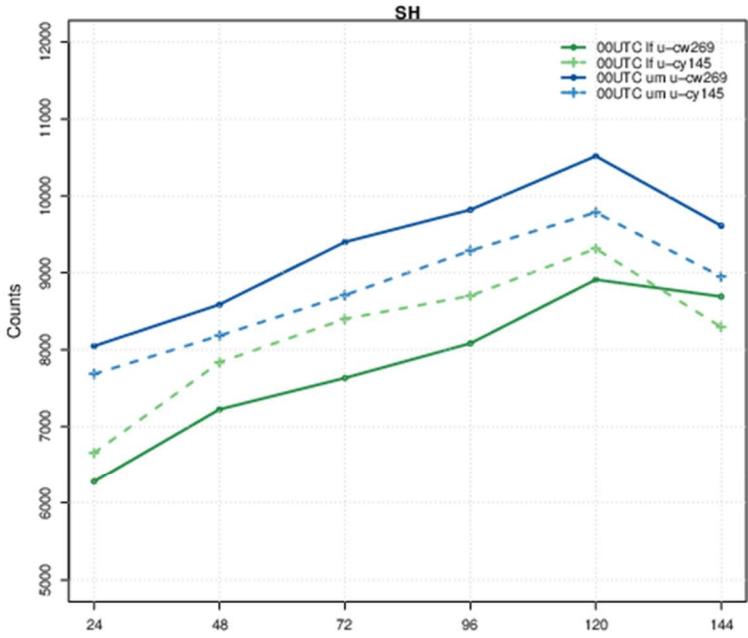
Bulk object attributes by region: grid counts

Objects >= 25 mm/24h

Dashed = coupled







- Generally **LF** has smaller (and fewer, not shown) objects (in terms of total grid counts) in NH and SH; **coupled LF higher**
- Largest differences in TR; increase for coupled

Met Office

To summarise

- We're clos*er* to putting together an operational workflow ready for implementation in PS47 with the UM (mid-2024?)
- The workflow components will continue to evolve.
- LFRic capability for R&D is increasing and evaluation of LFRic has begun. Much more rigorous evaluation will commence when full global DA trials can be run from early 2024.
- Regional LFRic development is offset by 6-12 months from the global timelines but will benefit from all the capability that is being developed for the global model.

Met Office
Questions?
www.metoffice.gov.uk
© Crown Copyright 2022, Met Office