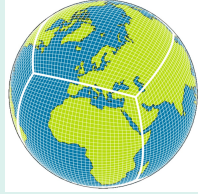


METplus Implementation: Categorical Statistics

Sebastian Cole, Rob Darvell, Phil Gill, Rachel North & METplus Implementation Team

METplus implementation is a part of the Next Generation Modelling System Programme at the Met Office

Rationale - Existing VER verification system has provided operational and trial verification for over 20 years but requires significant re-development to enable verification on the LFRic grid.



Aims – Implementation of the Model Evaluation Tool (MET) to provide trial and operational verification for UM, Ocean and LFRic models.



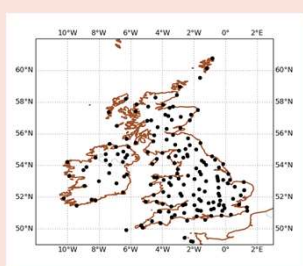
Accuracy - Rigorous testing of MET and comparisons with VER to ensure that verification results are as correct and robust as possible.

Categorical Statistics

Categorical statistics are computed as part of the operational forecast verification at the Met Office. To reproduce the operational output from VER:

- MET's Point-Stat tool was used. Point-Stat provides verification statistics for forecasts at observation points.
- Visibilities from the Unified Model (UM) were verified using VER and MET between 00UTC 01/09/2022 to 18UTC 06/09/2022 at 6 hourly intervals from T+6 to T+36.
- Forecast visibilities were interpolated using the nearest neighbour interpolation method.
- Only stations from the World Meteorological Organisations block 03 station list were considered.

Contingency tables were generated for visibilities >200m, >1000m, >4000m & >5000m. Equitable threat scores and frequency biases were then computed and compared along with contingency table counts.

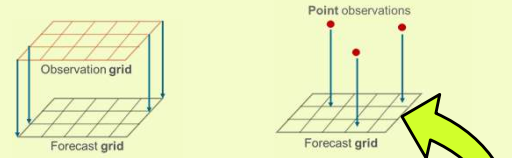


	Observations Category 1 Events	Observations Category 2 Events
Forecast Category 1 Events	A	B
Forecast Category 2 Events	C	D

MET/METplus

MET is a verification package developed by the National Center of Atmospheric Research (NCAR) Developmental Testbed Center (DTC). METplus provides a suite of python wrappers.

- More modular code to allow easier development
- Open source code for easier sharing between organisations
- Wider range of verification tools
- Contract with NCAR for initial developments leading to longer term collaboration

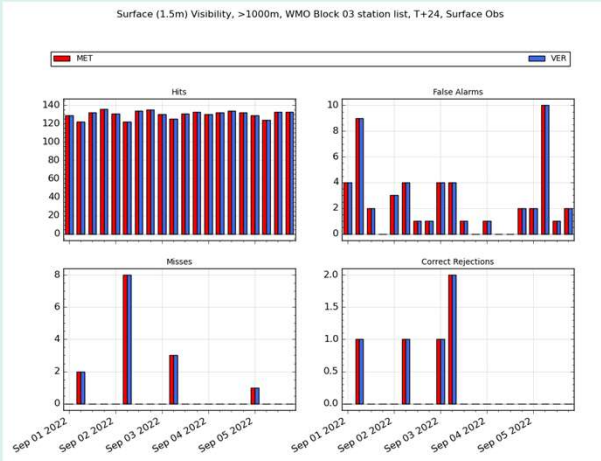


Initial Results

Comparing the VER and MET output, categorical statistics appear to be consistent between the two systems. For visibility, this result was consistent across the 4 thresholds applied. This suggests MET output is directly comparable to VER output and the operational time series of verification scores can continue. More testing is required to ensure this result is not unique to the Block 03 station list.

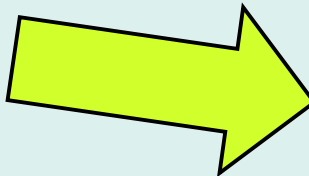
No Influence from Opposing Orders of Processing?

Other operational capabilities tested as part of the METplus implementation have seen the impact of opposing orders of processing between VER and MET. This hasn't been seen to affect the categorical statistics from Point-Stat as no observation point is equidistant from two points on the UM grid in this station list.



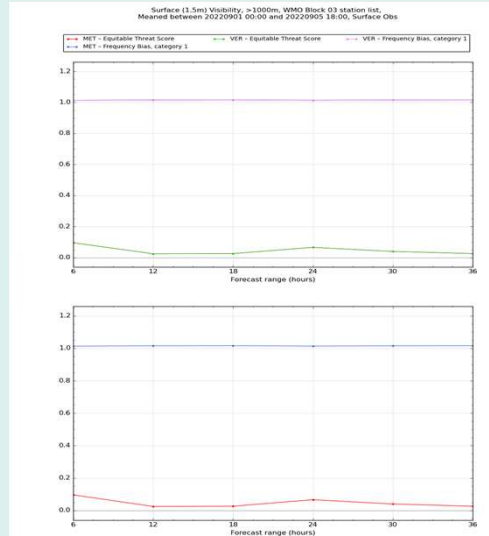
Equitable Threat Score:

$$\frac{(AD-BC)}{((B+C)(A+B+C+D) + (AD-BC))}$$



Frequency Bias, category 1:

$$\frac{(A+B)}{(A+C)}$$



VerPy Updates

VerPy is a Met Office Python library used for handling verification data. It can be used as either a library for loading data into a Python session or as an application for producing verification plots.

The testing and comparison of categorical statistics has highlighted an additional area for development to ensure continuity in the visualisation of statistics.

Next Steps...

- Investigate results over a wider range of station lists and areas.
- Testing of results from the UKV model.
- Explore the impact of different interpolation methods on scores.