Identifying "best" forecasts in EPS for Strategic Intervention

Marion Mittermaier, Clare Bysouth and Ric Crocker

The concept of Strategic Intervention

- ***Allows the Chief Forecaster to select an alternative data source to drive the production process***
  - ➢ Reduced usage of downstream intervention
- Priorities for usage:
  - Days 1-2 (i.e. deterministic)
  - Impact – large impact on customers

1.Major Sc cover errors in winter anticyclones

2.Major heavy rain errors

3.Synoptic errors

Two perspectives

- Site-specific scores (assessment **based on the UK Index** metrics which provide a **comprehensive test of surface weather parameter performance)**.
  - Tells us little about the correct evolution of the forecast as a forecaster sees it. Restricted to land stations. Doesn't account for upstream developments.
- Spatial verification methods can provide a similar intuitive interpretation to that which the forecaster makes.
- Question:
  - Can these methods tell us whether the **broad-scale picture** in a selected member is in fact better?
  - Do the **results agree with the site-specific** verification results?

Interpretation Challenges

- Can we arrive at a ‘best’ forecast for all lead times?
- Can we arrive at a ‘best’ forecast for all parameters?
- Can we arrive at a ‘best’ forecast for all methods?

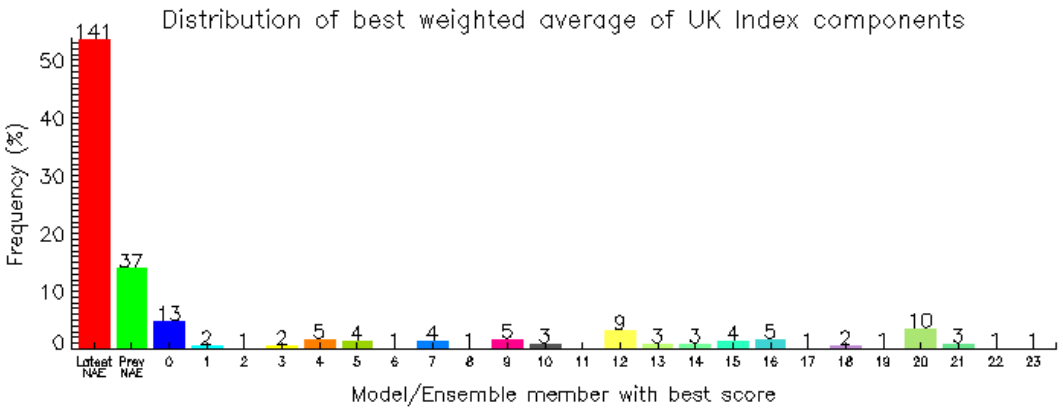
Site specific method

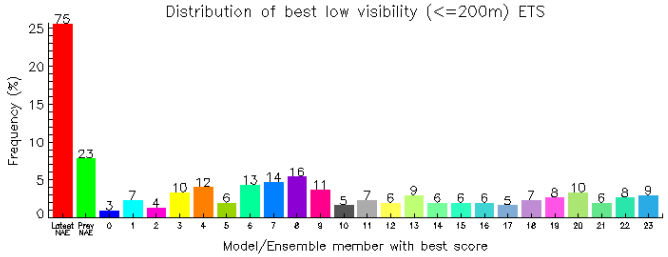
- A single run of the **NAE** is matched up with its **preceding NAE run** and **preceding MOGREPS-R** run.
- The same set of **land-based observations** are used to compile a score for each of the models/EPS members.
- 8 lead times used – T+6,12,18,24,30,36,42,48 from the main run matched with T+12-T+54 from the preceding runs
- **UK Index components** calculated for each lead time and each model/EPS member
  - Temperature, Precipitation, Visibility, Cloud, Wind.
- **This will provide an "envelope" of skill** which will enable us to **rank the forecasts** and determine the position of the selected member relative to all the others.

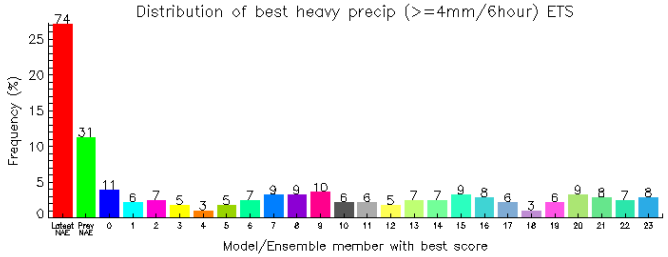
Site specific scores

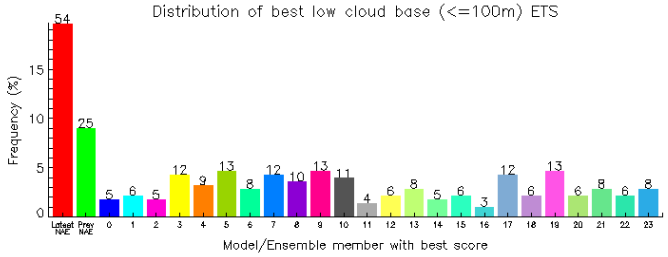
- Scores have been calculated for a 6 month period.
- Allows us to look at individual past cases
- Allows us to answer questions about how the alternatives compare to business as usual **on average**
  - E.g How often is the latest NAE run the best?

How often is the latest NAE run the best?



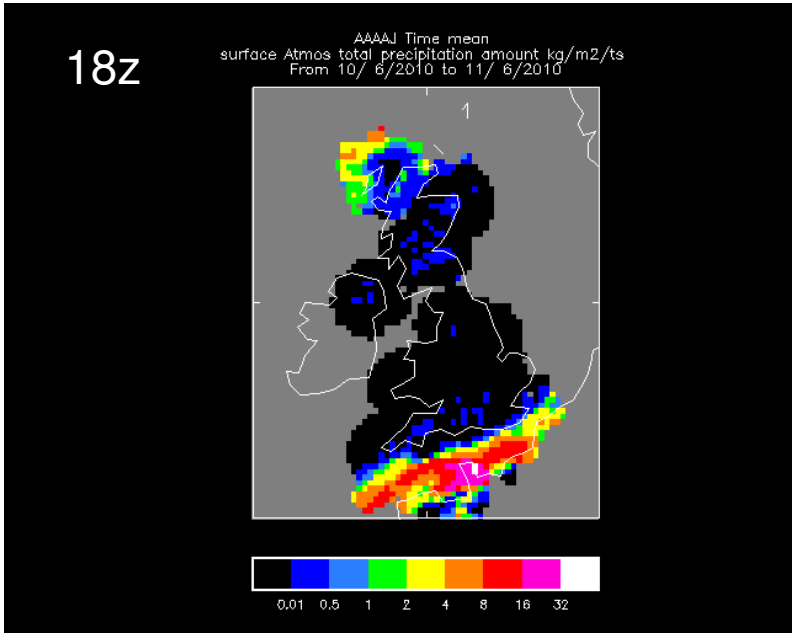


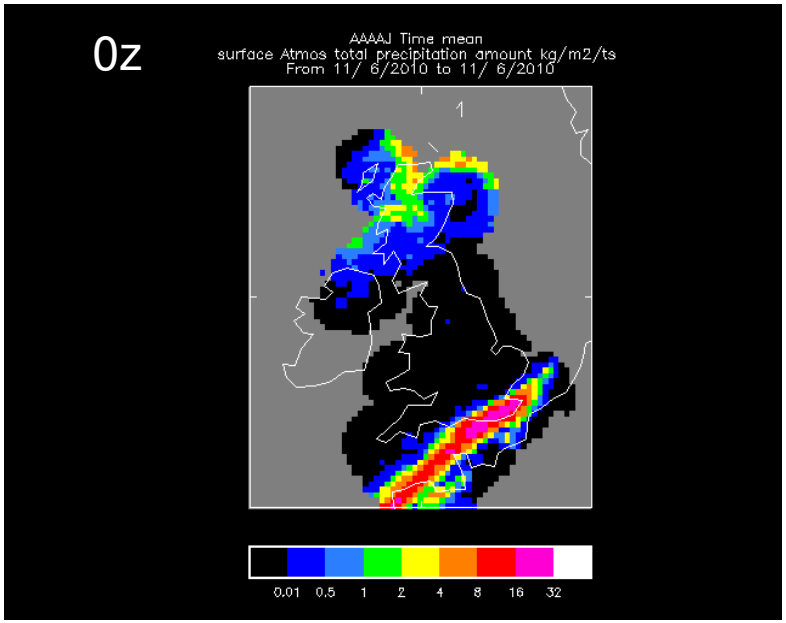




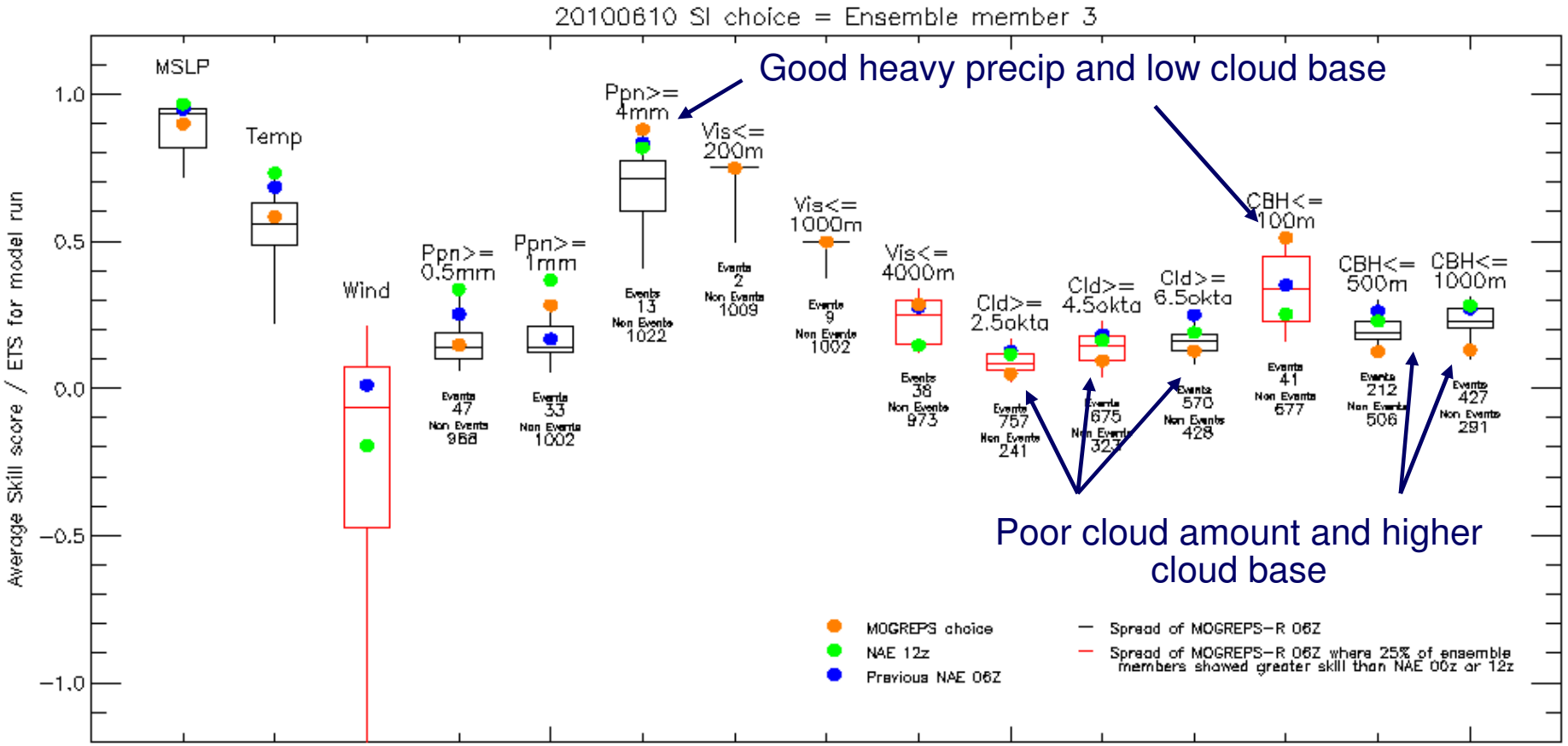
- The latest NAE has the best weighted average for over 50% of the model runs.
- When high-impact events are considered, the latest NAE is best for only one in every four or five runs.
- Not surprising - these conditions are not often widespread and the EPS has 24 chances of capturing them well and "outperforming" the NAE model.

10/06/10: NAE too slow with rain in south, although amounts not significant. Want to improve spread of heavy rain tomorrow and broken cloud in NW Britain.

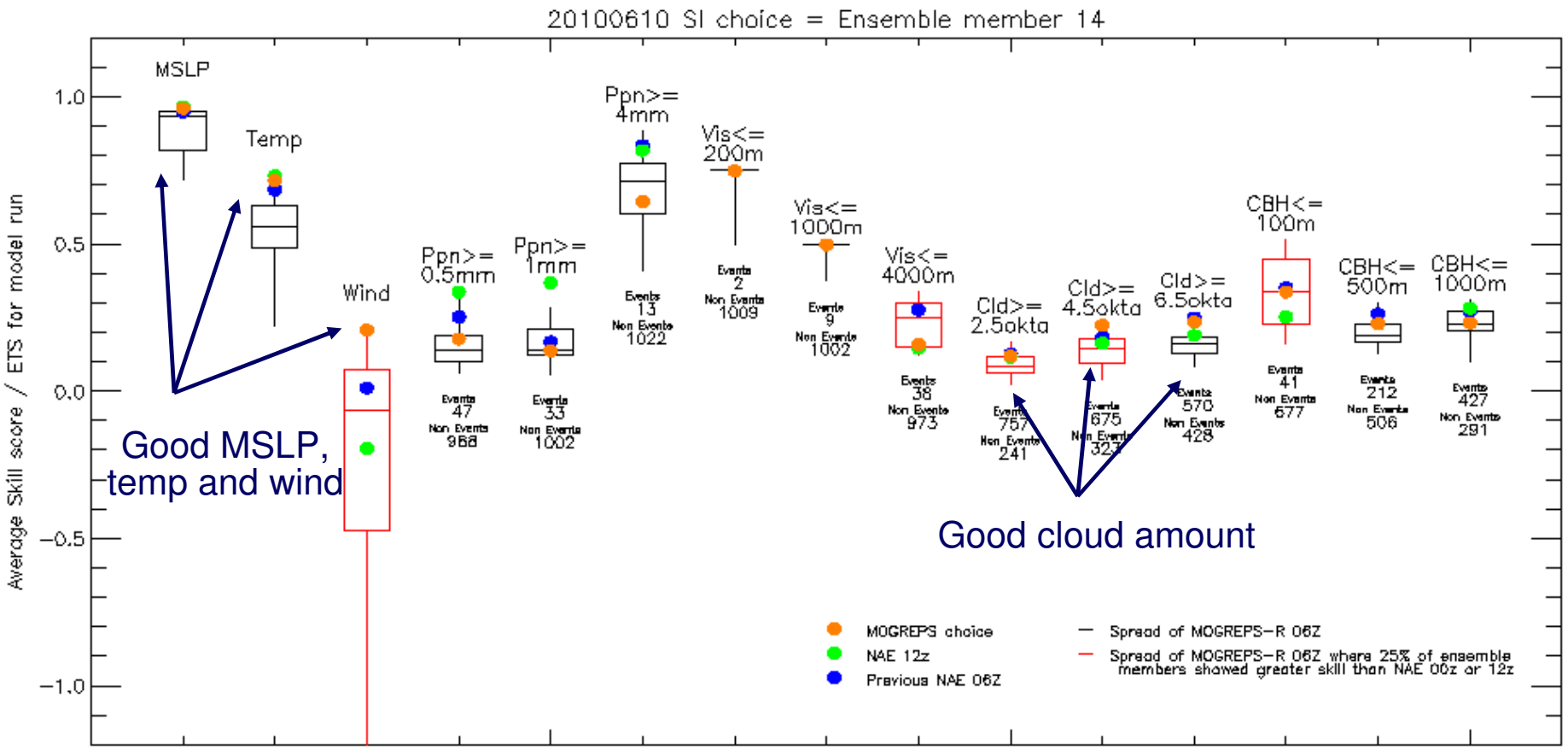




Good choice for Precip



Good choice for Cloud



Spatial Methods

- Spatial verification methods (using model analyses, observations analyses and satellite derived products) can provide a similar intuitive interpretation to that which the forecaster makes. Two tools were used: the fuzzy toolbox (Ebert,2008) and SAL (Wernli et al 2008/2009).

- Question:

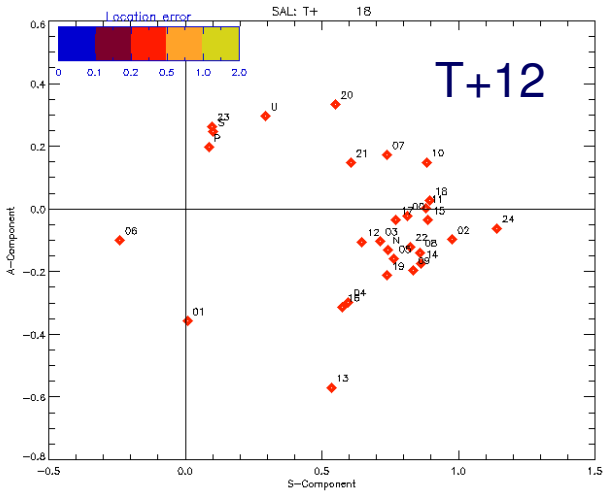
  - Can these methods tell us whether the **broad-scale picture** in a selected member is in fact better?

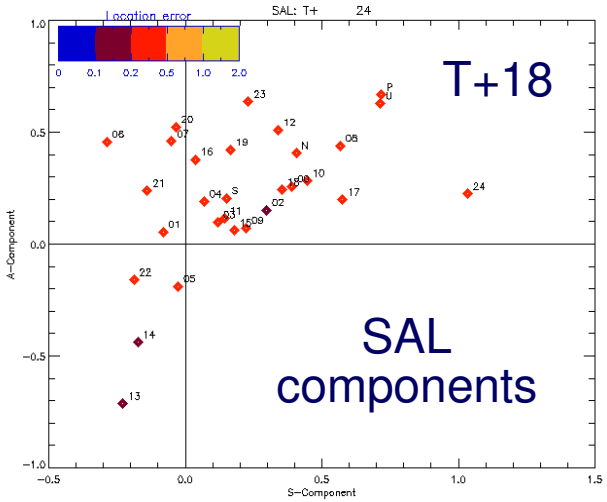
  - Do the **results agree with the site-specific** verification results?

Spatial Methods

- SAL - (Feature-based) Structure, Amplitude and Location error (focussing on higher intensity features) - scores close to zero are better.
- Fuzzy Toolbox - (Neighbourhood-based) A collection of fuzzy verification methods, which verify at different spatial scales and parameter thresholds using user selected scores (e.g. ETS)
- Most spatial methods were developed with focus on precipitation, we are trying to apply them to other parameters as well.

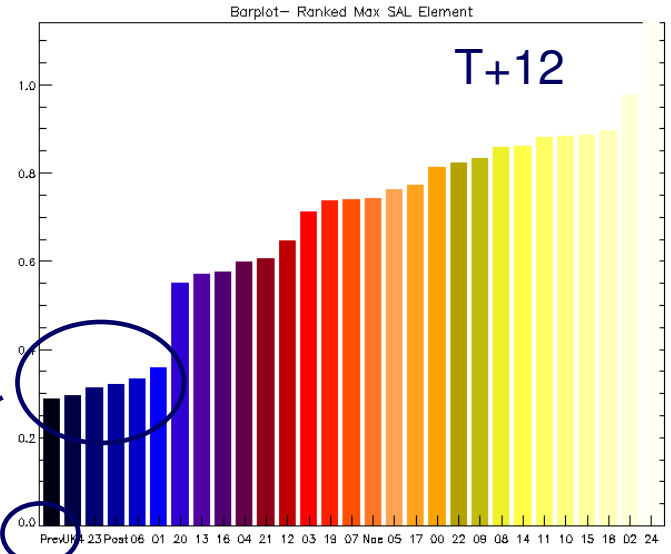
SAL – 10th June

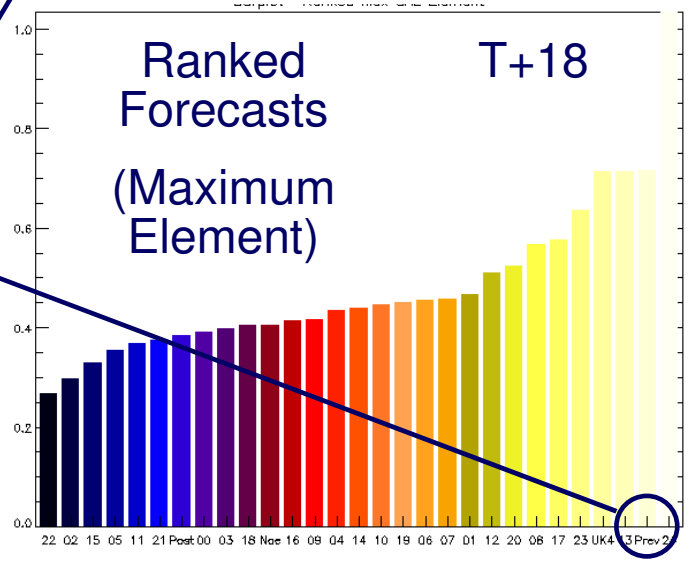




Cluster of possible 'best models'

Previous NAE run 'best' at T+12…'bad' at T+18

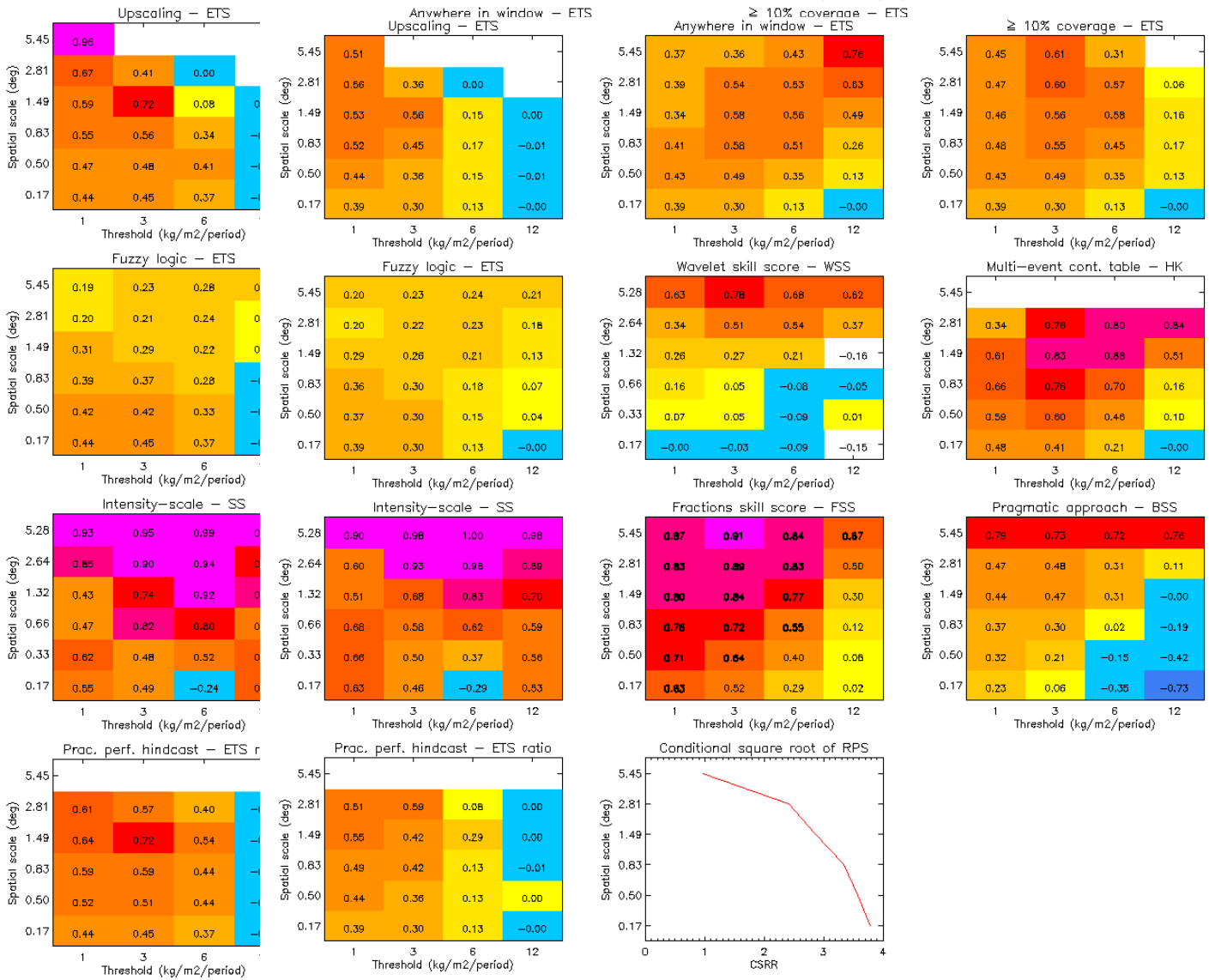




Fuzzy Verification - 10th June

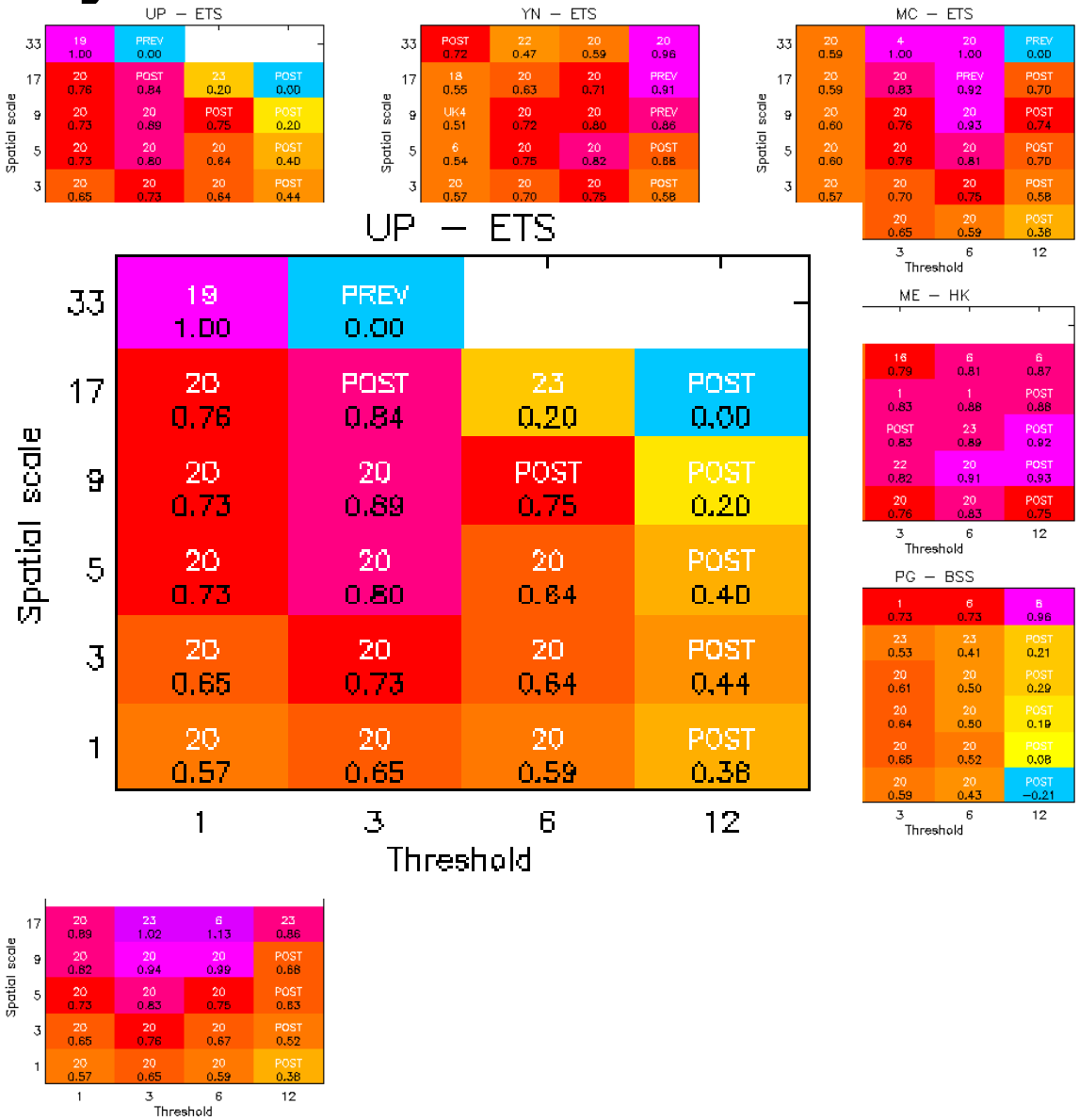
NAE

Ensemble member 1



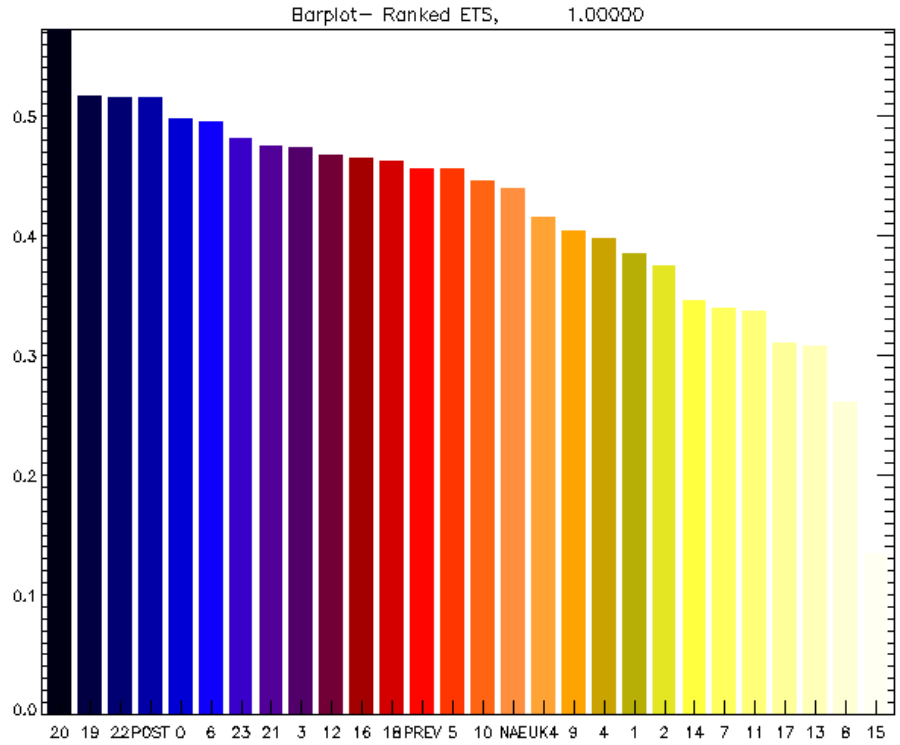
Fuzzy Verification - 10th June

Best Elements, All models



Ranked ETS (Upscaling method)

- Again , the outputs from the fuzzy package can be ranked.
- Pick best ETS over selected spatial scales and thresholds.



Things to ponder

- There is no “best” alternative for all variables at all lead times. – so what is “acceptable” alternative, how do we define this?
- Appropriateness of spatial methods to fields other than precipitation. Initial results show that cloud masks are probably not suitable for many methods. There is no real object.
- Understanding the form of “truth” used and the impact on the choice of method.
- What is the focal point for clouds? For precipitation it is clear, we generally want to get the big events right.
- This analysis framework has revealed many interesting (not understood!) characteristics of our forecast system that are now being investigated in other areas.

Questions?