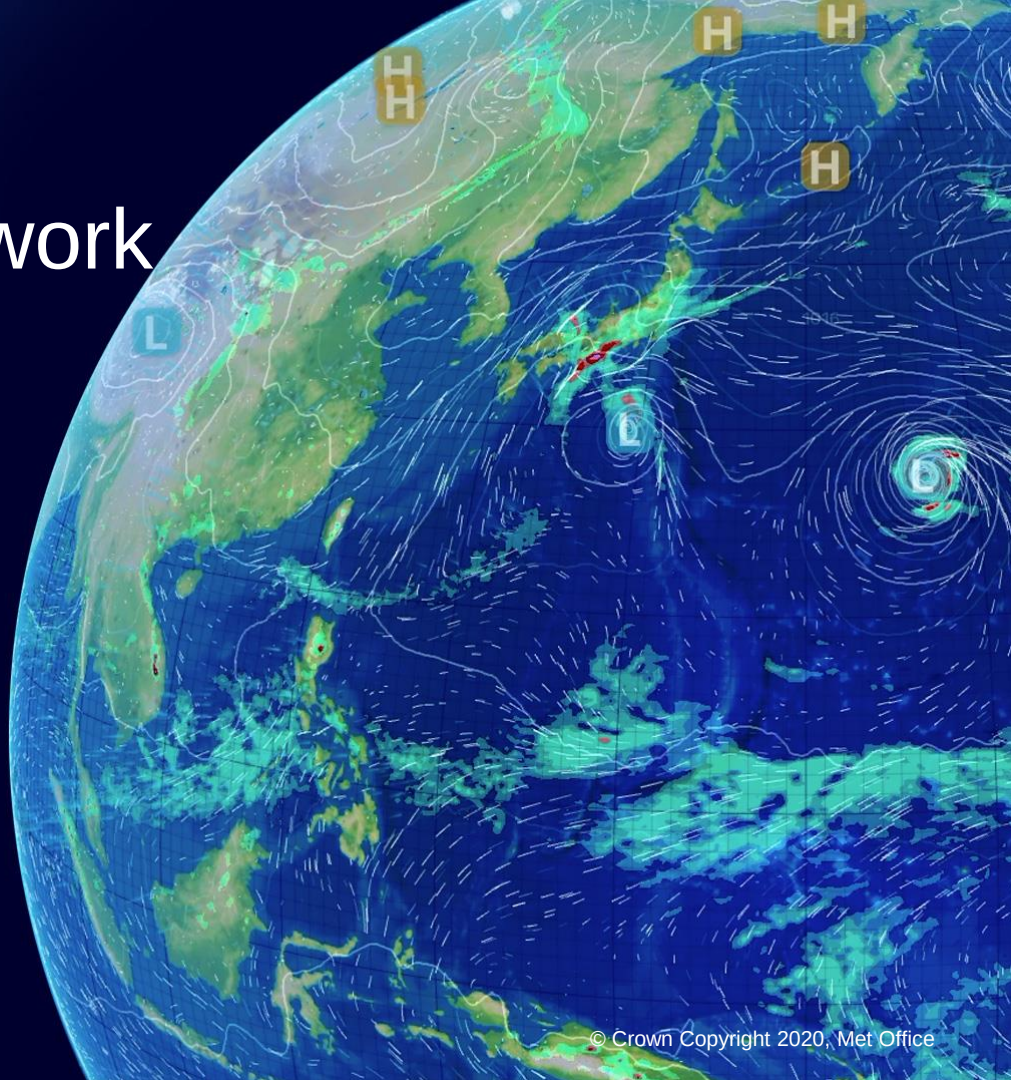
Update on ensemble work at the Met Office

Aurore Porson, Anne McCabe, Nigel Roberts, David Flack, Stuart Webster, David Walters, Steve Willington, Mike Bush, Bruce Macpherson



Content

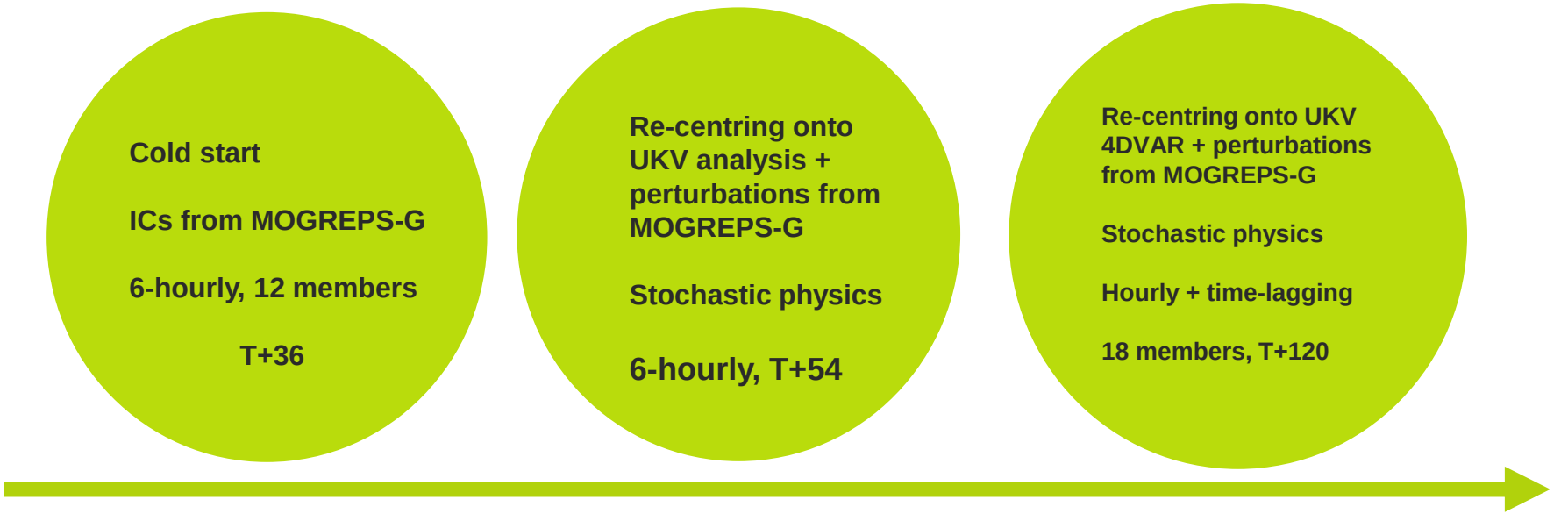
- Operational activities
  - Hourly cycling assessment
  - Parallel suite trials
- Research developments
  - Hazardous Weather Testbed
  - UK mini-testbed
  - Ensemble Toolbox
  - Fog Development
- Conclusions

Met Office

Content

- Operational activities
  - Hourly cycling assessment
  - Parallel suite trials
- Research developments
  - Hazardous Weather Testbed
  - UK mini-testbed
  - Ensemble Toolbox
  - Fog Development
- Conclusions

Towards hourly cycling:



Bowler et al., 2008
Operational in MOGREPS-R, 2012
Operational in MOGREPS-G, 2013

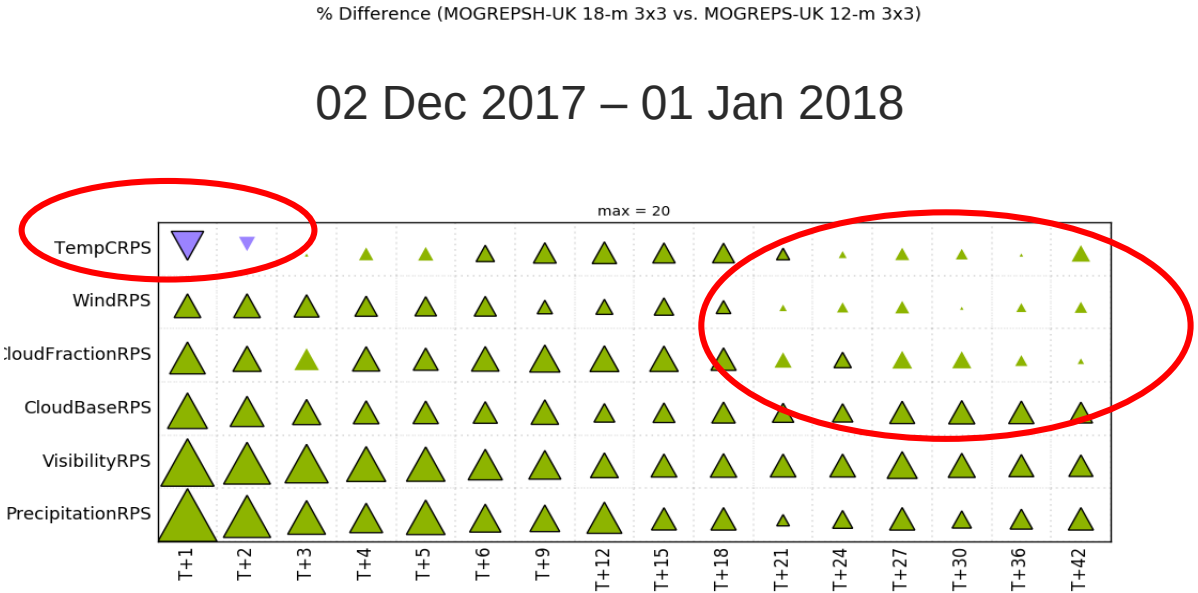
Tennant, 2015
McCabe et al., 2016
Hagelin et al., 2017
Operational in 2016

Operational in March 2019
Porson et al. 2020

Ensemble scorecard: Summary of objective verification

(Aurore Porson, Jo Carr, Susanna Hagelin, Rob Darvell, Rachel North, David Walters, Ken Mylne, Marion Mittermaier, Bruce Macpherson)
Porson et al. 2020, https://doi.org/10.1002/qj.3844

1) Temperature sensitive to differences in the times of the analyses between the 2 configurations.



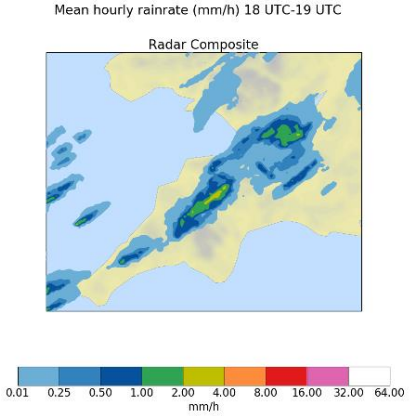
2) Smaller improvements at later forecast times

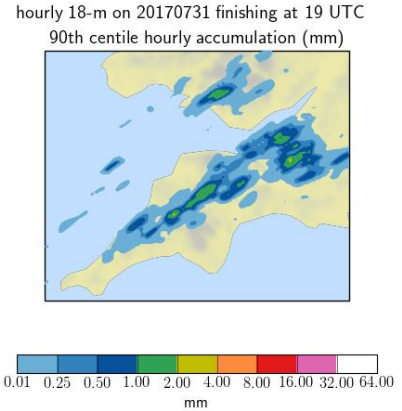
Application to case studies: Examples at short lead times

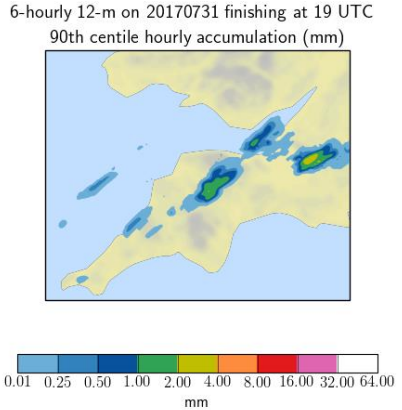
T+4 products now available for the hourly configuration

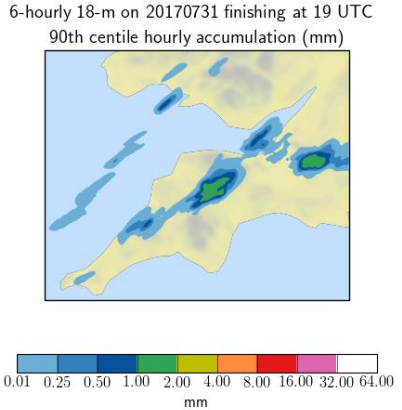
Better spatial structure at short lead times

Better probability products at short lead times with more spread between the members



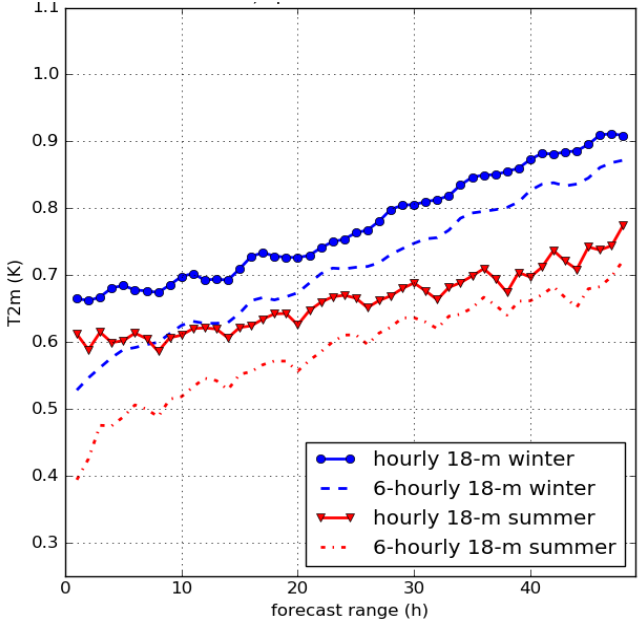


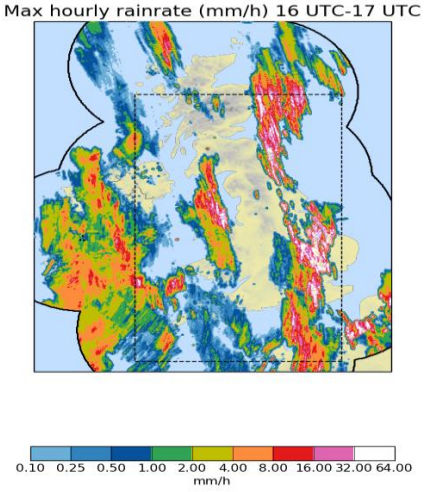


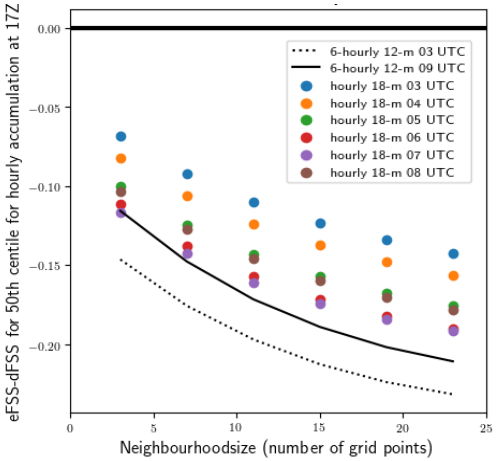


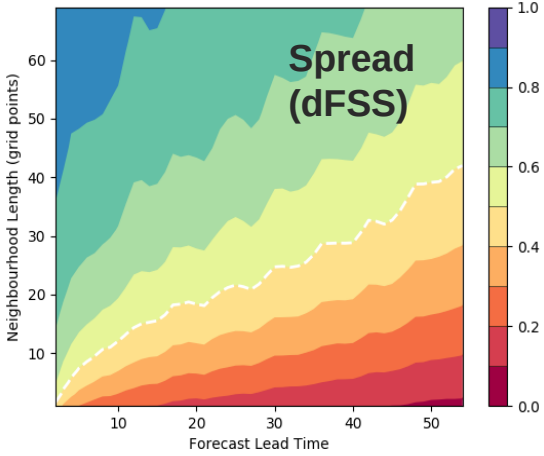
Ensemble spread
(Aurore Porson and Anne McCabe)

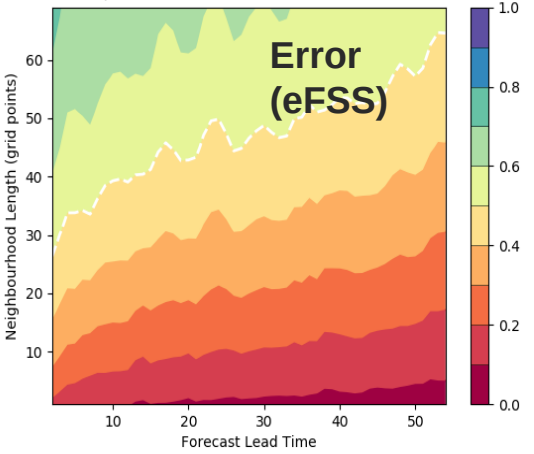
Standard deviation (members to ensemble mean)/
Standard deviation (ensemble mean – obs)









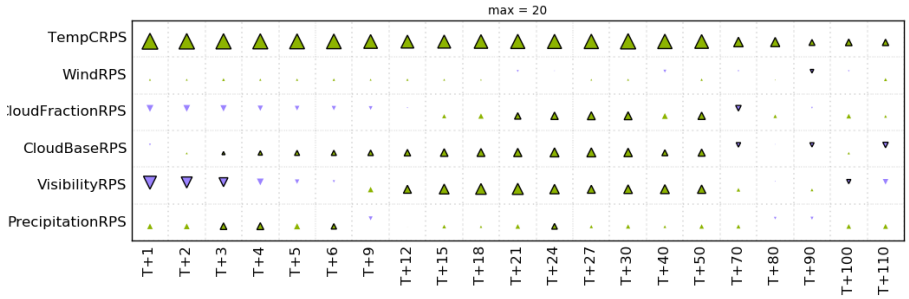


Met Office

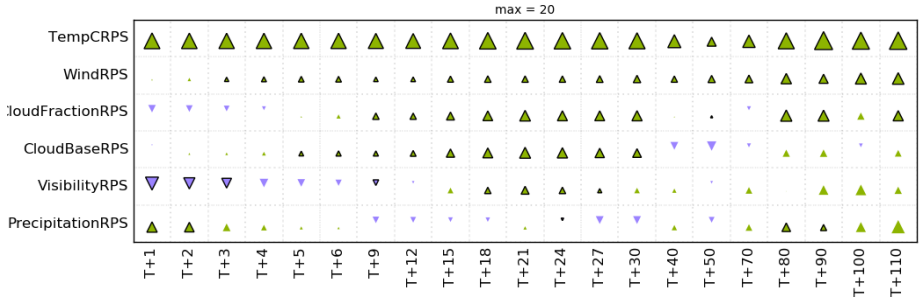
See Mike Bush's talk on Monday 28th September

WINTER

OS43gl_cons vs control

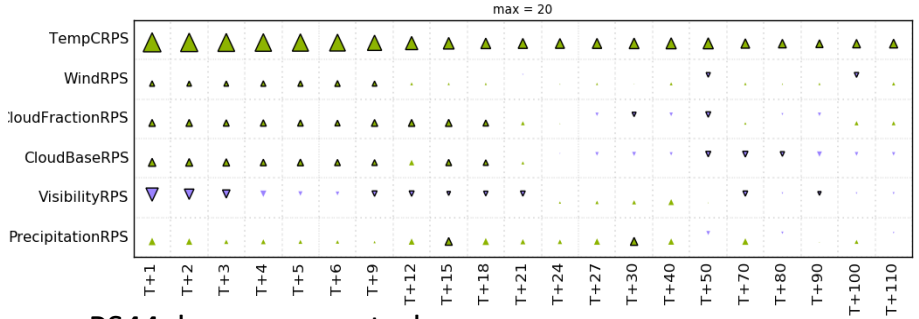


PS44gl_cons vs control

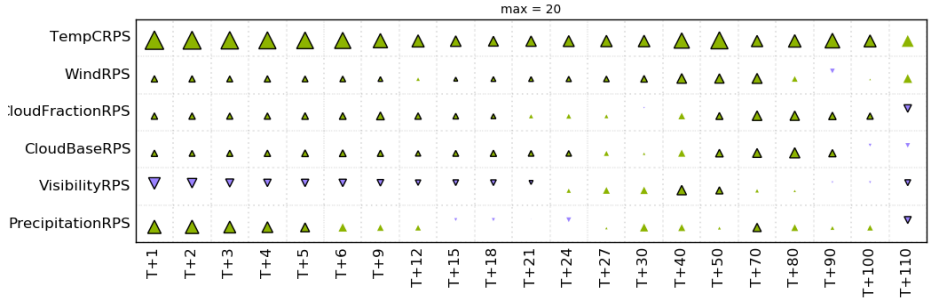


SUMMER

OS43gl_cons vs control



PS44gl_cons vs control



Strong benefit on CRPS

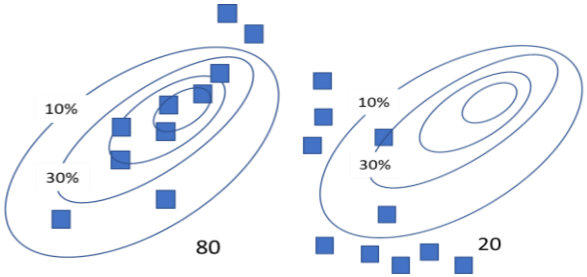
Not shown here, but PS44 driving conditions have more impact on the spread than the physics package itself

Content

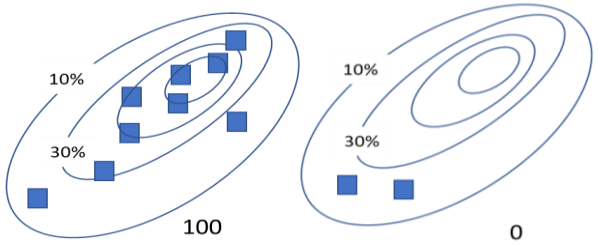
- Operational activities
  - Hourly cycling assessment
  - Parallel suite trials
- Research developments
  - Hazardous Weather Testbed
  - UK mini-testbed
  - Ensemble Toolbox
  - Fog Development
- Conclusions

HWT and New subjective method (Nigel Roberts)

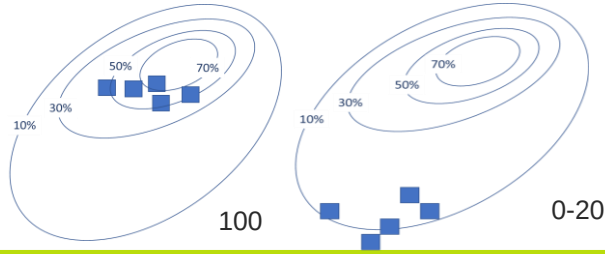
S1 Fraction of observed events captured (hit rate)



S2 High Probability success rate (non false alarm)



S3 High Density Success Rate (non missing)



See Adam Clark and Brett Roberts's talk on Thursday 1st October

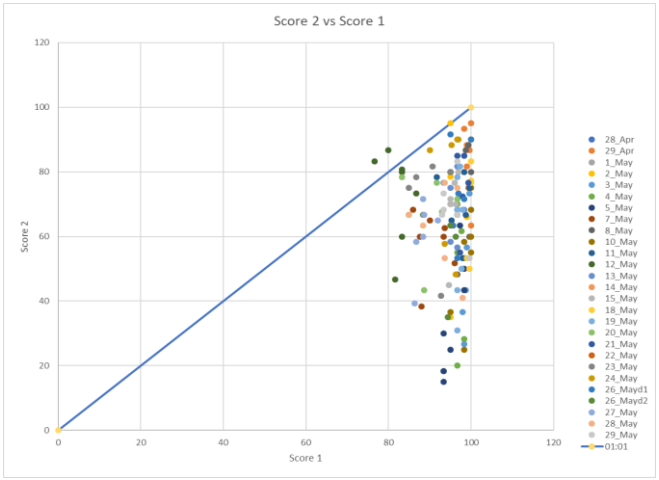
- ***How robust is this method?***
- ***Can it help us to understand what we qualify as "poor spread"?***
- ***Can it help us to understanding the sensitivity to ensemble configurations for severe weather?***

Operational plan HWT 2020
https://hwt.nssl.noaa.gov/sfe/2020/docs/HWT_SFE2020_operations_plan.pdf

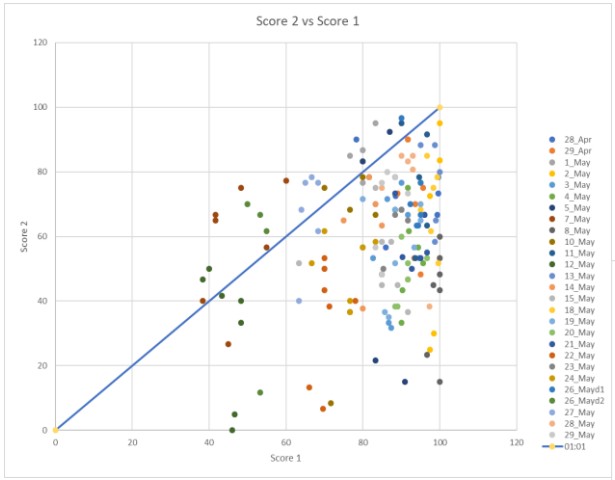
Single–cycle ensembles	Time-lagged ensembles	Multi-model ensemble	Multi-model and time-lagged ensemble	5 different deterministic models + time-lagging 12 hours
UM 00Z	UM TL10, UM TL18	UM+HRRRE	UM+HRRRE TL36	HREF
HRRRE 00Z	HRRRE TL10, HRRRE TL18			

Application of the scores to the ensemble configurations (Met Office, BoM, NOAA/NSSL, NOAA/GSL)

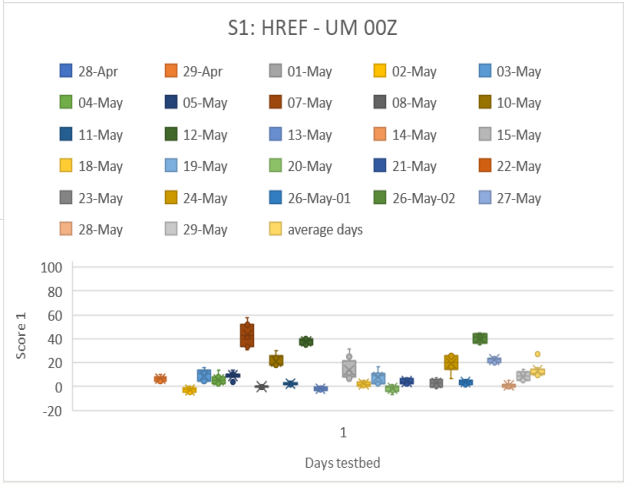
Aurore Porson, Nigel Roberts, Anne McCabe, Marion Mittermaier, Gareth Dow, David Flack, Steve Willington, Harald Richter, Terra Ladwig, David Dowell, Burkely Twiest, Adam Clark, Israel Jirak



HREF

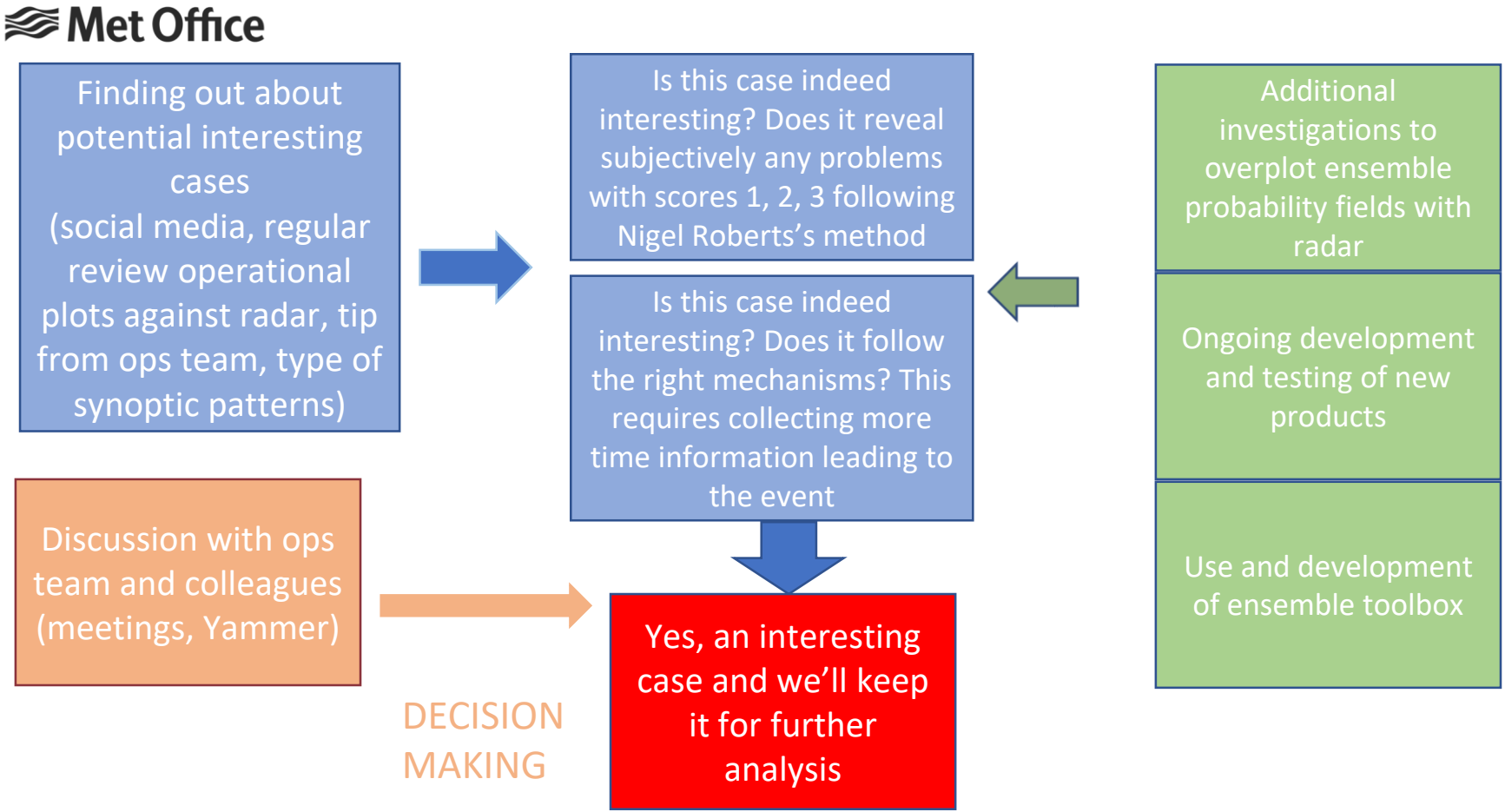


UM

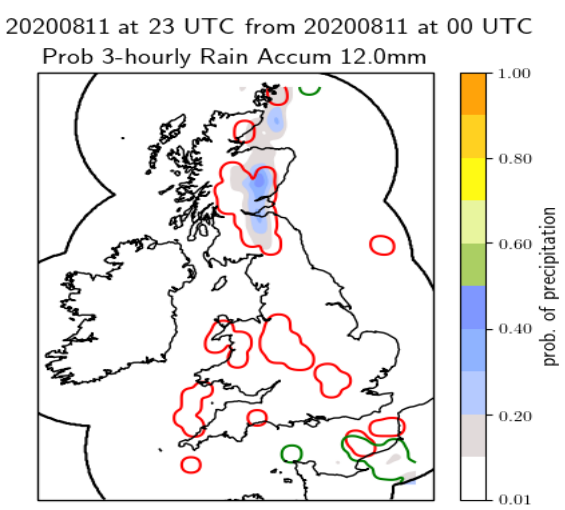


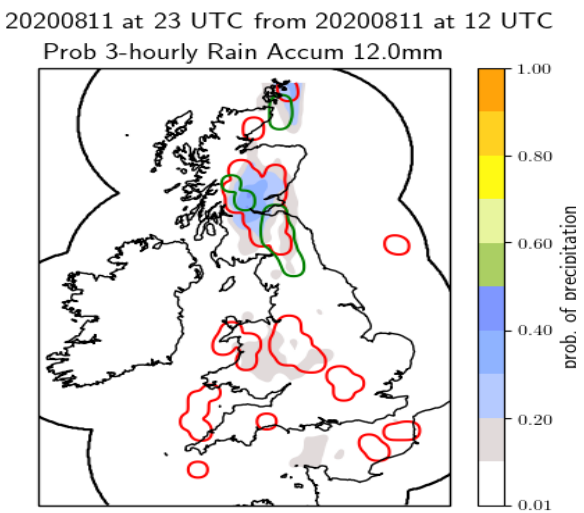
Mini-testbed activity (Aurore Porson, David Flack)

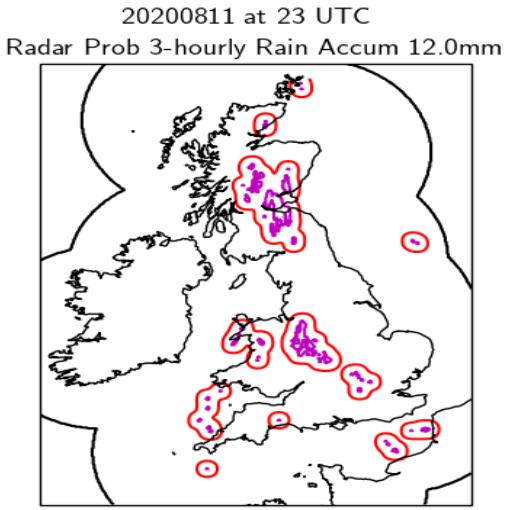
- Select the best cases for multi-model ensemble project
- Understand more fundamentally our operational ensemble - Assess the spread of the ensemble subjectively with the hourly cycling update
- Contribute to raise the profile and value of the ensemble among our scientists
- Provide a good source of data for testing future scientific developments



Direct subjective comparison between radar, ensemble and UKV neighbourhood probability product



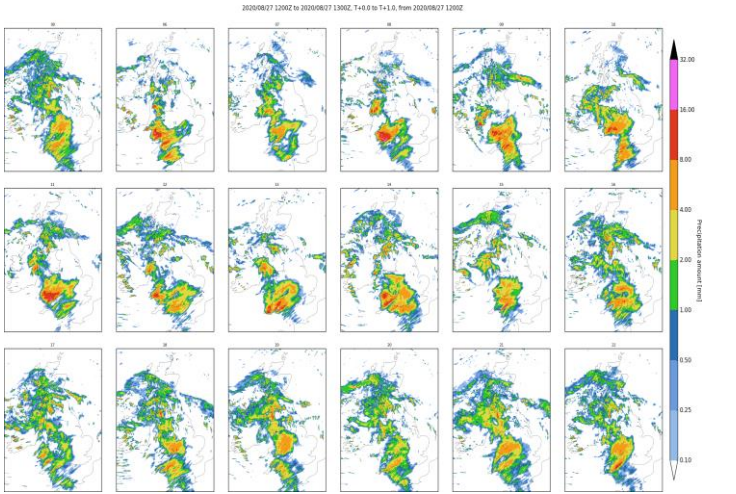




UKV

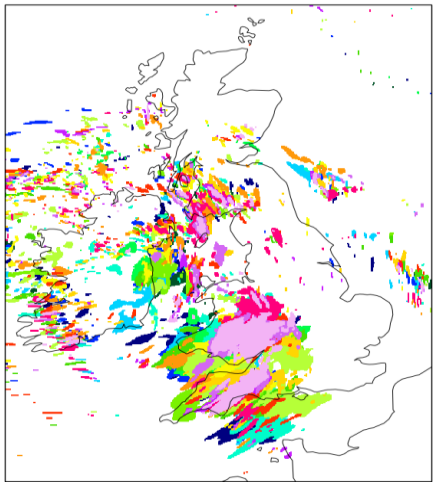
radar

Product development: Ensemble toolbox (Anne McCabe, Stuart Webster, Aurore Porson)

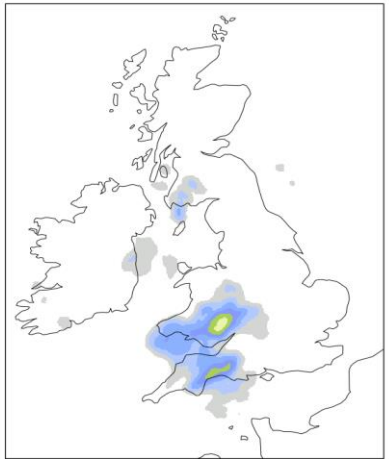


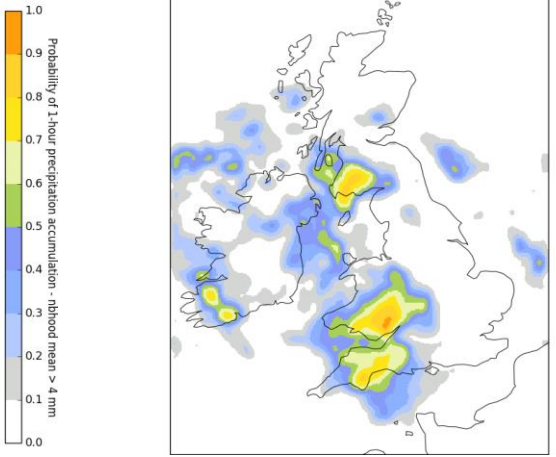
2020/08/27 1200Z to 2020/08/27 1300Z, T+0.0 to T+1.0, from 2020/08/27 1200Z

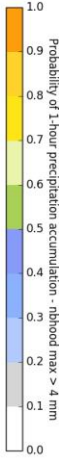
2020/08/27 1200Z to 2020/08/27 1300Z, T+0.0 to T+1.0, from 2020/08/27 1200Z



2020/08/27 1200Z to 2020/08/27 1300Z, T+0.0 to T+1.0, from 2020/08/27 1200Z







Conclusions and Future Work

- Parallel Suite 45 testing (time-varying SSTs)
- Regional Atmosphere RA3 development
- FSS applications to long-term statistics and perturbation analysis
- Collaboration with SRNWP-EPS on multi-model ensemble project
- Continue to test the subjective analysis framework and compare to objective verification in future testbeds
- Fog analysis
- … (being reviewed regularly as important strategic development)

Thank you for your attention!

- Aurore Porson (aurore.porson@metoffice.gov.uk)
- Anne McCabe (anne.mccabe@metoffice.gov.uk)
- Nigel Roberts (nigel.roberts@metoffice.gov.uk)
- David Flack (david.flack1@metoffice.gov.uk)
- Stuart Webster (stuart.webster@metoffice.gov.uk)

Motivation for the hourly cycling and time-lagging

The new hourly time-lagged configuration:

- follows naturally from the UKV (deterministic high-res model) running hourly-cycling 4DVar
- is designed to achieve a larger ensemble size (Hagelin et al. 2017) and longer lead times
- is designed to produce more timely forecasts
- is designed to increase the ensemble spread by:
  - Using multiple cycles of MOGREPS-G
  - Taking into account differences in the high-res DA by staggering the initial conditions

Fog Forecasting with Ensembles (Anne McCabe)

Very sensitive to small changes so well-suited to ensemble forecasting

What is the best way to extract a probability forecast for fog?

What are the characteristics of our current ensemble?

What can we learn from high resolution simulations compared with good quality field observations (e.g. LANFEX and SOFOG)?

How well does the RP scheme represent the model uncertainty? How does it compare to a multi-physics approach?