Met Office

# MOGREPS ensemble forecasting

Jonathan Flowerdew

International Conference on Ensemble Methods in Geophysical Sciences

Thursday 15 November 2012

© Crown copyright Met Office

# Talk outline

- MOGREPS and initial perturbations
- Ensemble combination and calibration
- Extended-range storm surge ensemble
- Summary

© Crown copyright Met Office

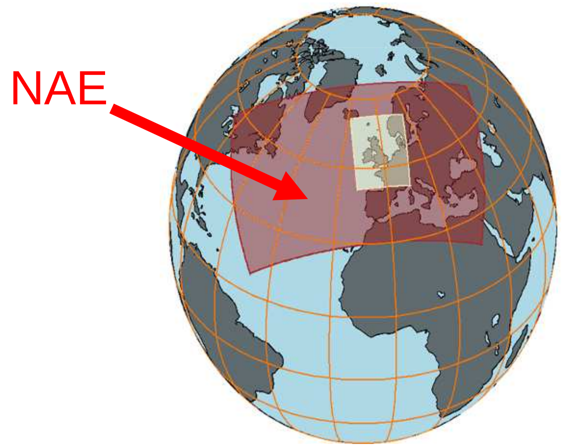
# MOGREPS and initial perturbations

© Crown copyright Met Office

# MOGREPS – The Met Office atmospheric ensemble

9 March 2010 - 27 March 2012

- 24-member ensemble designed for short-range forecasting
    - Regional ensemble over N. Atlantic and Europe (NAE) to T+54 at 06Z and 18Z (18km resolution, 70 levels)
    - Global ensemble to T+72 at 00Z and 12Z (N216 ~60km resolution, 70 levels)
        - Also runs to 15 days at ECMWF for multi-model ensemble research
    - ETKF for initial condition perturbations
    - Stochastic physics

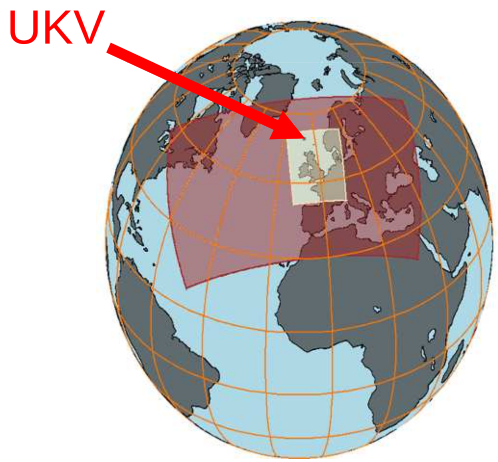



© Crown copyright   Met Office

# MOGREPS – The Met Office atmospheric ensemble

from December 2012

- 44 perturbed members at 00/06/12/18 UTC, of which 11 plus unperturbed control are run to full length each cycle
  - Products use 24 members formed by combining the two most recent cycles
  - UK ensemble to T+36h (2.2km central resolution, 70 levels)
  - Global ensemble to T+72h (N400 ~33km resolution, 70 levels)
  - Medium-range ensemble to T+15d (N216 ~60km resolution, 70 levels)
  - ETKF for initial condition perturbations
  - Stochastic physics (currently global only)



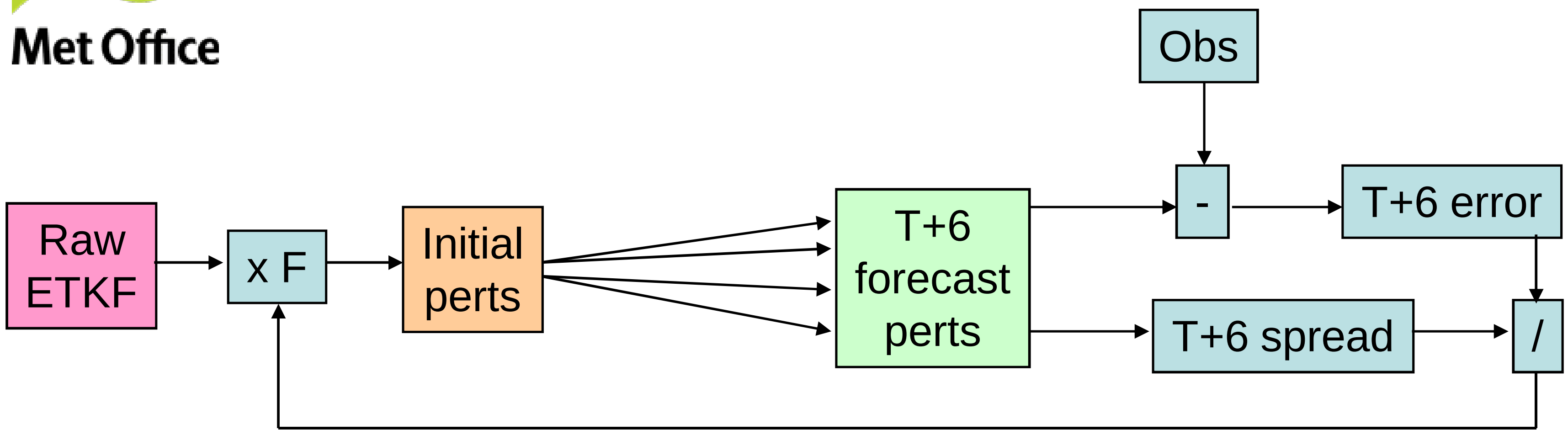

© Crown copyright Met Office

# Pros and Cons of ETKF

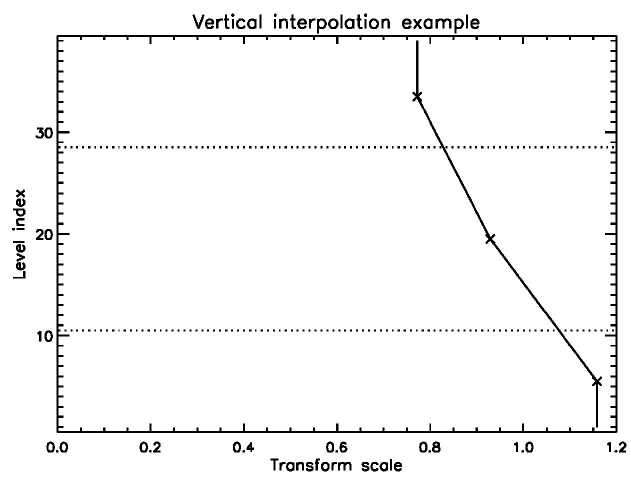
- MOGREPS initial perturbations are generated using an ETKF, which mixes and scales the background perturbations
- Short-range forecast perturbations are now used by hybrid data assimilation
- ETKF pros: cheap, simple,
  developed structures from T+0,
  relation to data assimilation theory,
  mixes balanced background perturbations
- ETKF cons: hard to localise properly, thus needs lots of inflation, hard to iterate like VAR, two separate systems to maintain

© Crown copyright Met Office

# Adaptive inflation



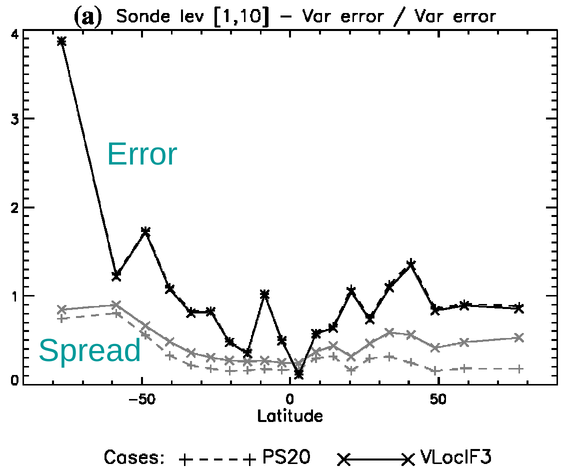

- The ensemble spread is imperfect for several reasons
- The online calibration scales ensemble spread towards the rms error of the ensemble mean, as measured by sonde and ATOVS data
- The calibration varies with run, location, … and now height
- ATOVS observations contribute to vertical band averages in proportion to their integrated Jacobians



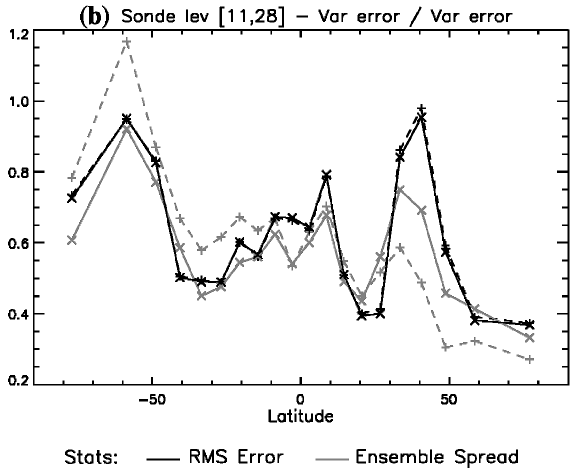

© Crown copyright Met Office

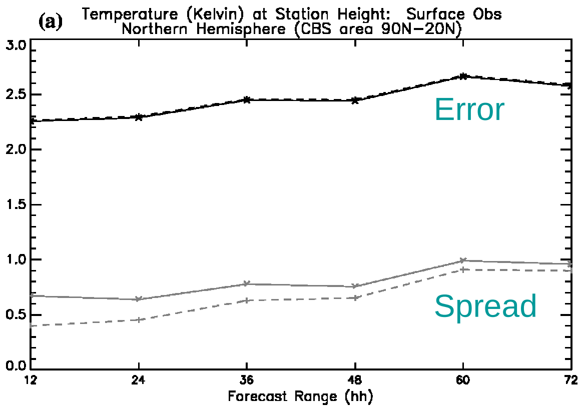
# Vertically-varying inflation: results

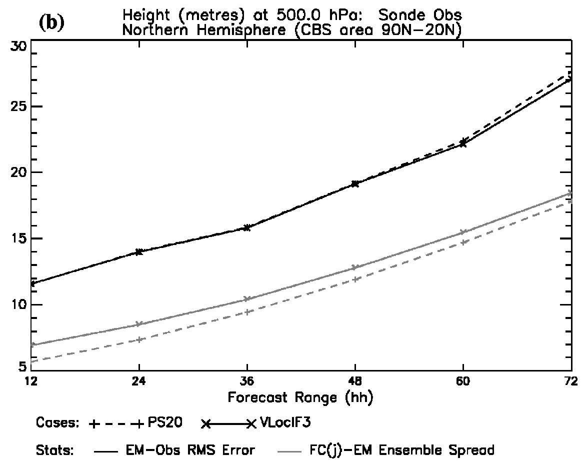
Boundary layer

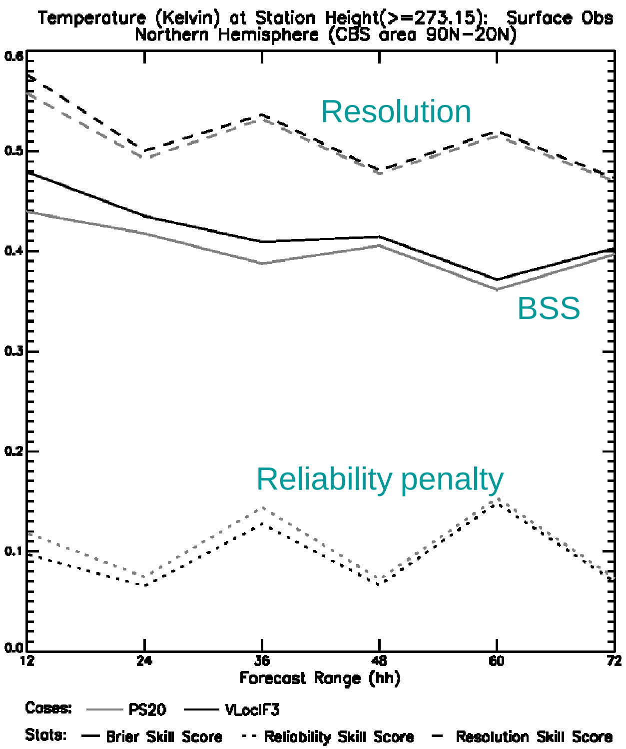



Troposphere




© Crown copyright Met Office

# Ensemble combination and calibration

© Crown copyright Met Office

# Ensemble combination and calibration

- Physically-based ensembles can suffer from systematic errors in climatology, spread and reliability which can vary as a function of location, lead time, diurnal cycle, threshold, …
- Combining ensembles provides more samples, averages systematic errors, and samples uncertainty in model structure
- Statistical calibration adjusts forecasts in an attempt to remove systematic errors detected by comparing past forecasts and observations

© Crown copyright Met Office

# Reliabilty calibration principles

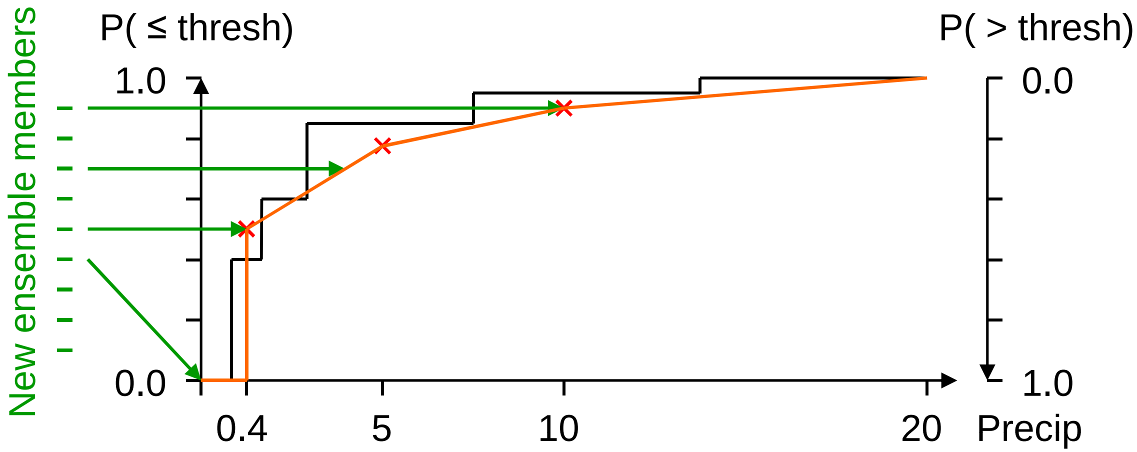
- Probability verification provides the most general measure of ensemble performance
- (Local) reliability is more amenable to improvement by post-processing than (local) resolution
- Forecast realism, derived variables, and downstream models all require the preservation of spatial, temporal and inter-variable structure
- The chosen method makes minimal assumptions about the underlying distributions, and tries to make maximum use of underlying ensemble forecast
- The scheme is as local as it can be whilst achieving a specified noise threshold
- The scheme is trained on a year of recent forecast-observation pairs to be applicable to all systems without requiring reforecasts

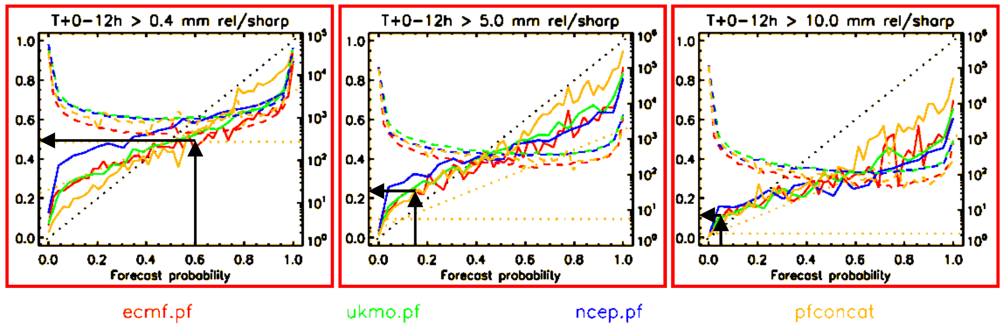
© Crown copyright Met Office

Met Office

# Reliability calibration scheme

1. CDF of raw ensemble
2. Update probability for each threshold
3. This implies a calibrated CDF
4. New members equally divide probability range, as in rank histogram
5. Interpolate members from calibrated CDF in **same order** as raw ensemble (preserves local structures). This step similar to Bremnes (2007)




Standard configuration uses ten thresholds:
(0), 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4, 12.8, 25.6, 51.2, (102.4)

Results are aggregated over space as needed to reach 200 samples

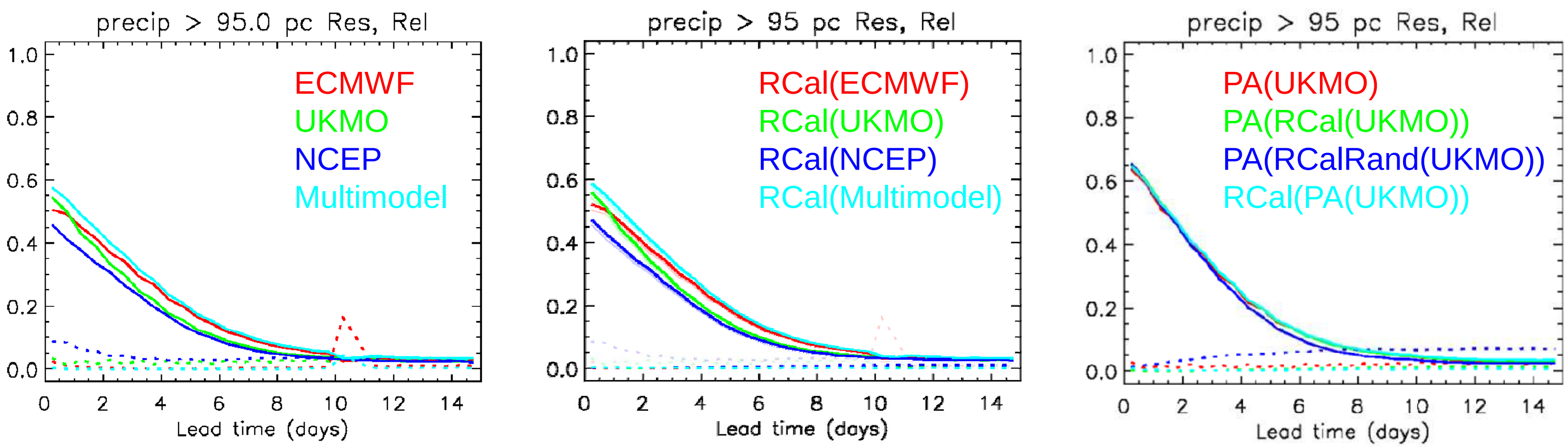
© Crown copyright Met Office

# Verification setup

- ECMWF, UKMO and NCEP perturbed ensemble forecasts from the THORPEX TIGGE archive
- Interpolated to 1˚ grid over Europe
- Twelve-hour precipitation accumulations
- Two-year period (April 2010 to March 2012)
- Verified against Met Office ‘europp’ dataset (composite of raingauge-adjusted radar, satellite-derived precipitation, and high-resolution short-range forecasts)
- Dividing the period into four six-month blocks, the calibration is trained on the two blocks cyclically 'preceding' each forecast
- Results show reliability and resolution components of Brier Skill Score for probability to exceed thresholds defined as percentiles of the 5x5-gridpoint climatology of the same month in the other year

© Crown copyright Met Office

# Multimodel precipitation results




- The multimodel combination has a consistent advantage over single model ensembles, in both resolution and reliability
- The calibration almost eliminates unreliability, and improves overall resolution

- Calibration at the grid scale (green) performs almost as well as direct calibration at 3x3 scale (cyan)
- Randomised member assignment (blue) is much worse

© Crown copyright Met Office

# Extended-range storm surge ensemble

Thanks to Helen Titley, Ken Mylne, Caroline Jones, Jane Williams

Work funded by the Environment Agency

© Crown copyright Met Office

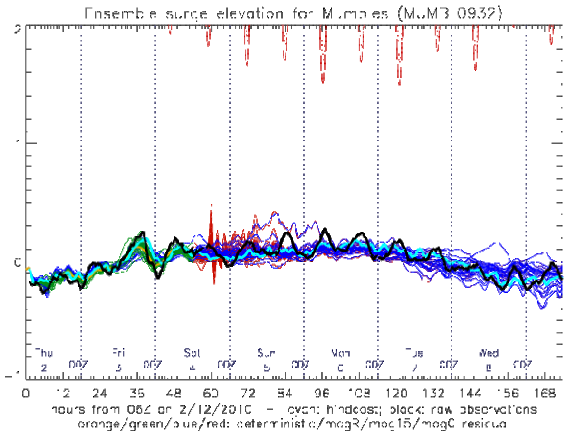
# Coastal storm surges

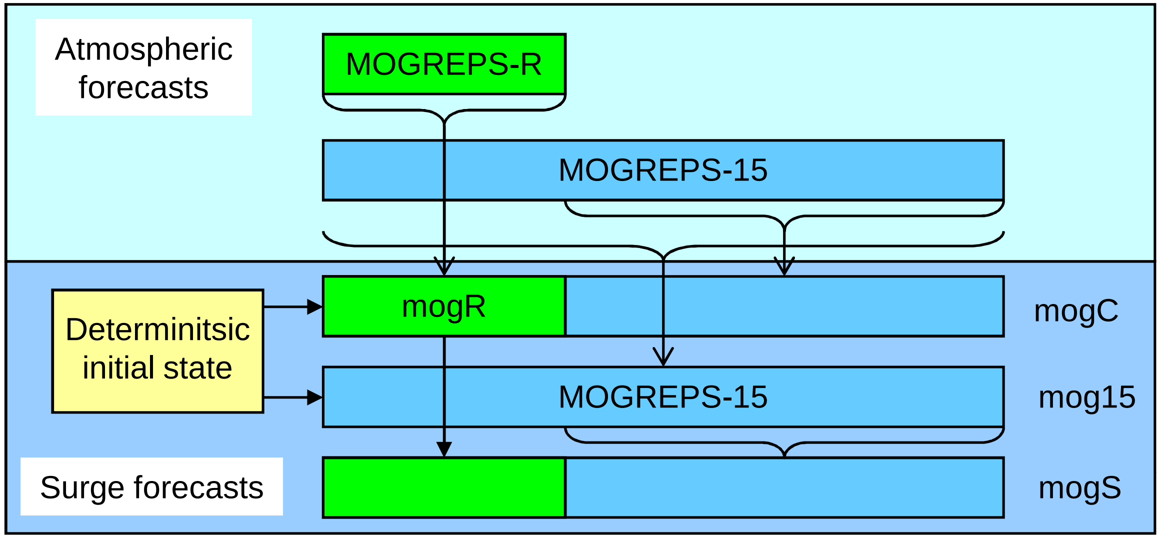
Sea Palling, Norfolk (1 Feb 1953)
This surge was the worst UK natural disaster in recent times, leading to 307 deaths in East Anglia and 1836 in the Netherlands.

- The increase in coastal water level due to atmospheric wind and pressure
- Good forecasts enable protective actions such as warning the public, closing roads, evacuation, errecting temporary defenses, …
- The UK has had a 54-hour operational surge ensemble since December 2009
- Can the lead time be usefully extended based on medium-range atmospheric forecasts?

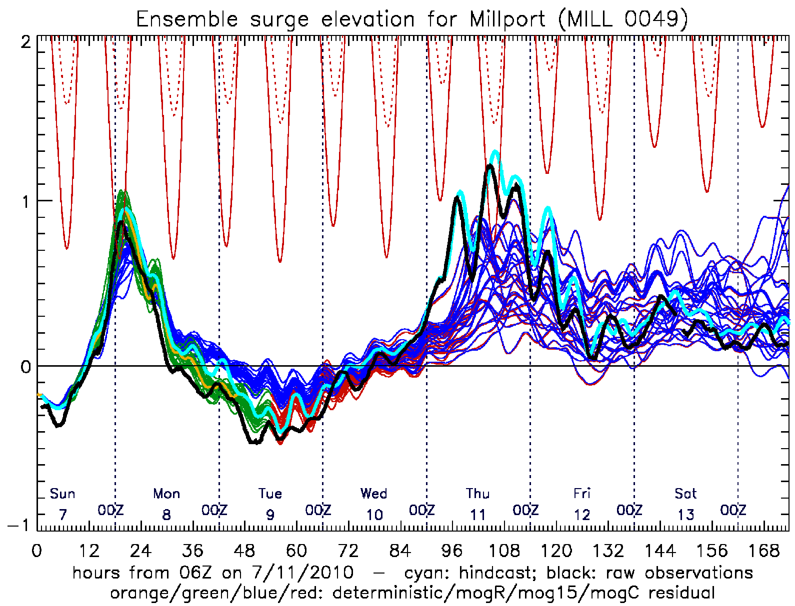
© Crown copyright Met Office

# Extended surge ensemble



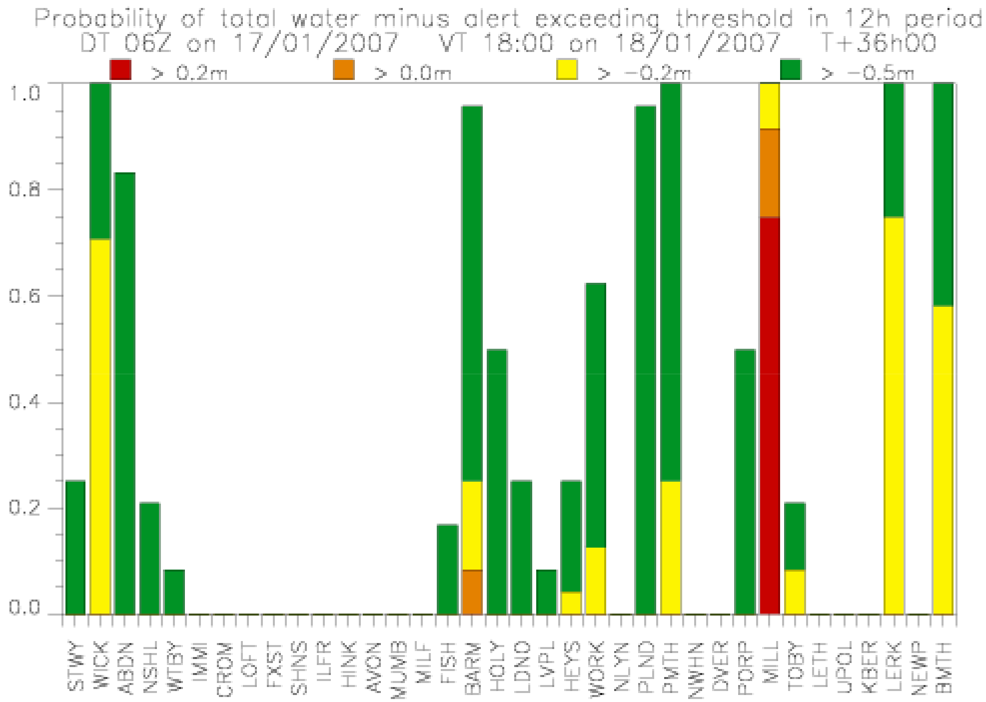



- Green: Surge driven by MOGREPS-R (mogR)
- Blue: Surge driven by MOGREPS-15 (mog15)
- Lower red: Surge driven by MOGREPS-R then MOGREPS-15 (mogC)
- Orange: Deterministic surge
- Black: Surge implied by tide gauge observations
- Cyan: Surge driven by deterministic analyses (hindcast)
- Upper solid red: Alert level minus tide prediction from harmonic analysis of tide gauge data
- Upper dotted red: Alert level minus tide prediction from surge model

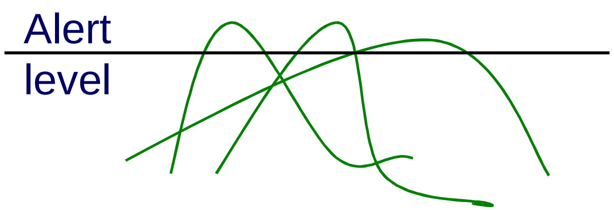



© Crown copyright Met Office

# Port risk charts




- Stacked probability for successively severe events
- Summarises all ports in a single plot, with direct link to cost/loss decision framework
- Accumulated in 12 hour bins starting at 6 hour intervals to match decision-making by tidal cycle



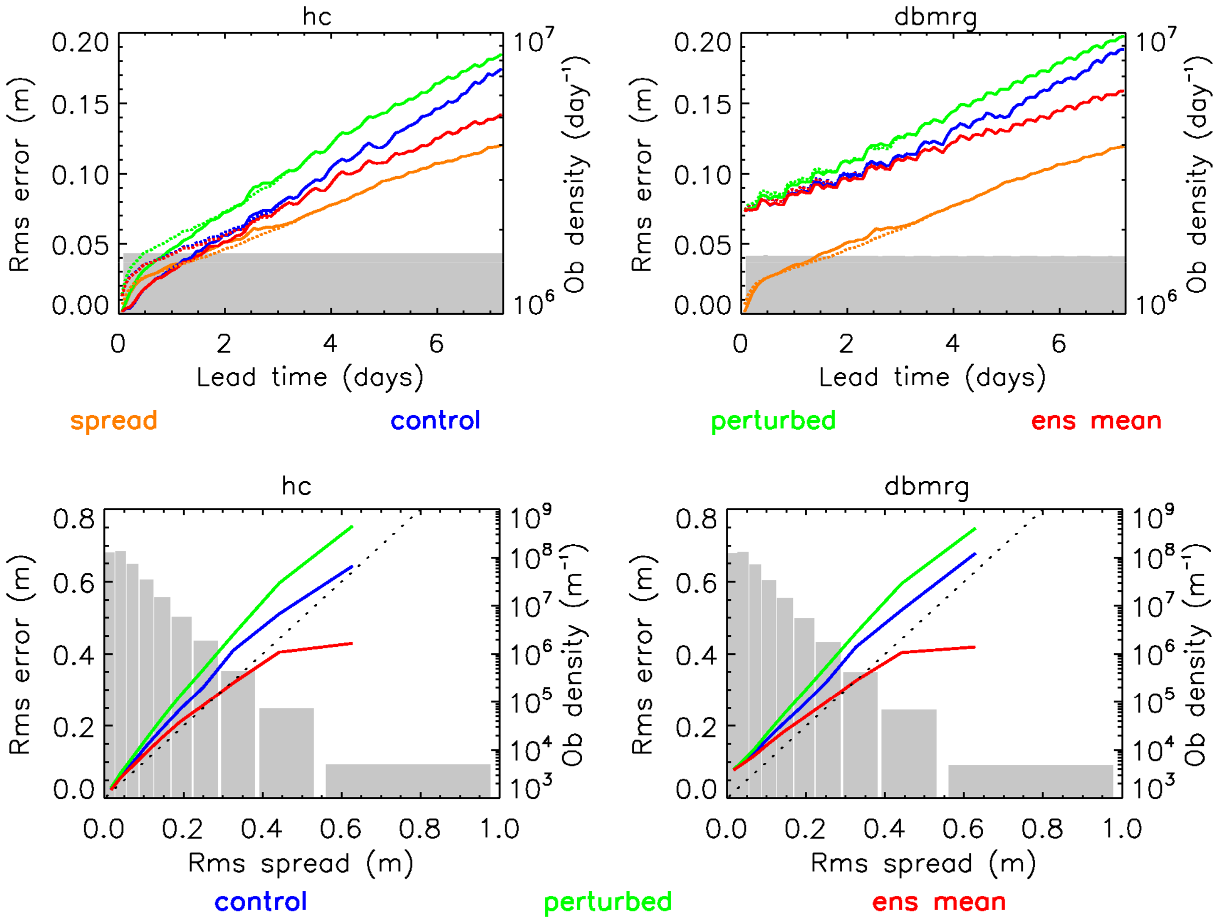

Time binning example: at any one time, only one member is above threshold, but three members exceed it at some point within the time window

© Crown copyright Met Office

# Spread and error statistics

Verification against hindcasts

Verification against observations

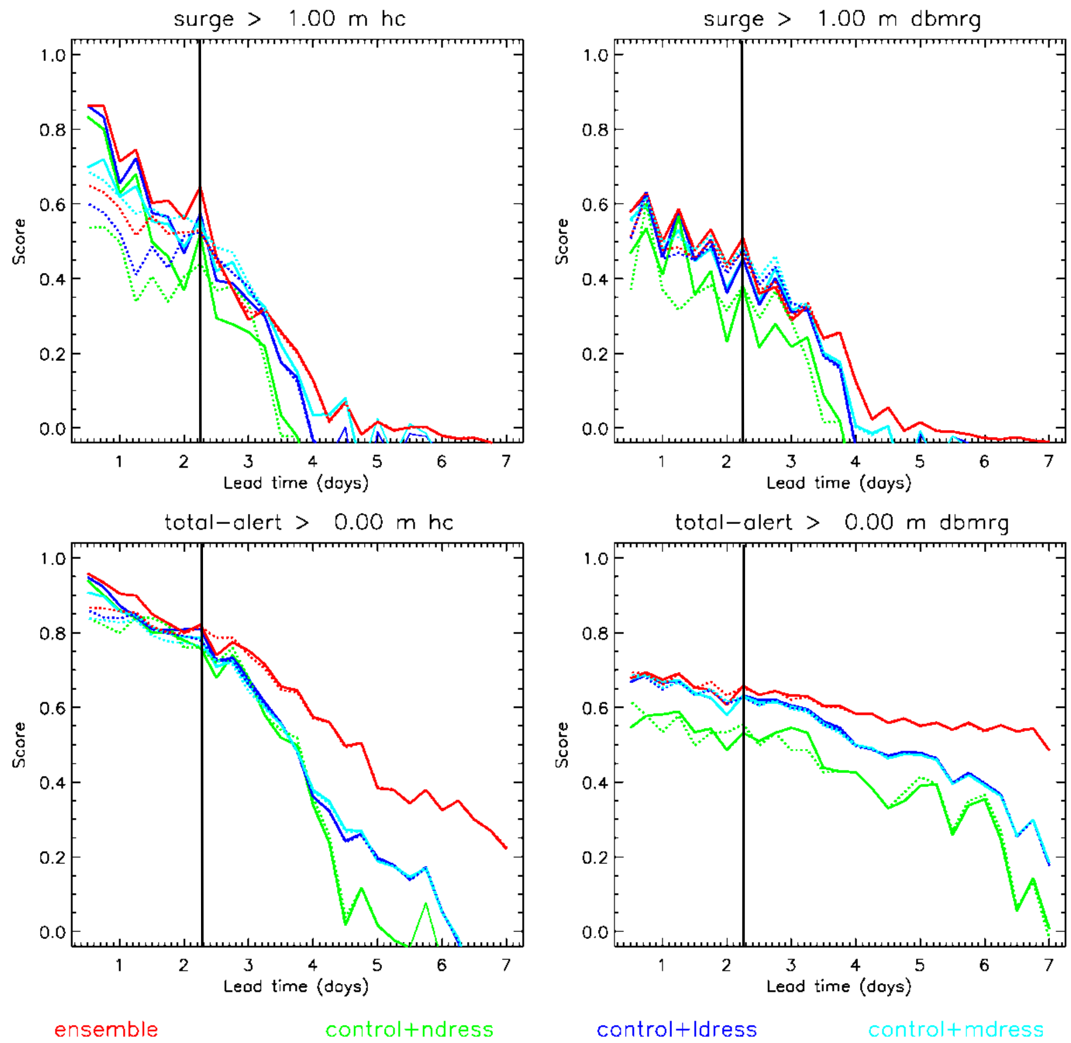



- Solid: mogC; dotted: mog15
- The higher-resolution forcing gives lower error and higher spread
- There is a small bump in mogC error and spread after the T+54h transition
- Non-meteorological errors domainte the error against observations in calm situations
- The ensemble is good at predicting the much larger errors associated with major surge events

© Crown copyright Met Office

# Probability verification

- Ensemble (red) is generally superior to dressed control (blue/cyan) and much superior to pure control (green)
- There is some skill across the full lead time range
- The higher-resolution forcing available for the first 54 hours is generally beneficial, but mog15 is doing reasonably well by the time of transition
- There are several hints of inferior mogC performance following the transition




© Crown copyright Met Office

# Overall summary

- The next operational suite will nest the 2.2km UK ensemble directly within N400 global
- The revised online spread calibration scheme improves the vertical distribution of spread
- The reliability-based calibration method improves probabilistic performance whilst preserving spatial structure
- Combining ensembles produces slightly better forecasts, before and after calibration
- The forecast range of the storm surge ensemble has been successfully extended using input from MOGREPS-15

© Crown copyright Met Office

# References


- Flowerdew J, Bowler NE. 2012. Online calibration of the vertical distribution of ensemble spread. Accepted by Q. J. R. Meteorol. Soc.
- Bowler NE, Flowerdew J, Pring S. 2012. Tests of different flavours of EnKF on a simple model. Accepted by Q. J. R. Meteorol. Soc.
- Flowerdew J, Mylne K, Jones C, Titley H. 2012. Extending the forecast range of the UK storm surge ensemble. Q. J. R. Meteorol. Soc. Early View.
- Flowerdew J, Bowler NE. 2011. Improving the use of observations to calibrate ensemble spread. Q. J. R. Meteorol. Soc. 137: 467-482.
- Flowerdew J. 2012. Calibration and combination of medium-range ensemble precipitation forecasts. Met Office Forecasting Research Technical Report 567.


© Crown copyright Met Office

# Questions & answers

© Crown copyright Met Office