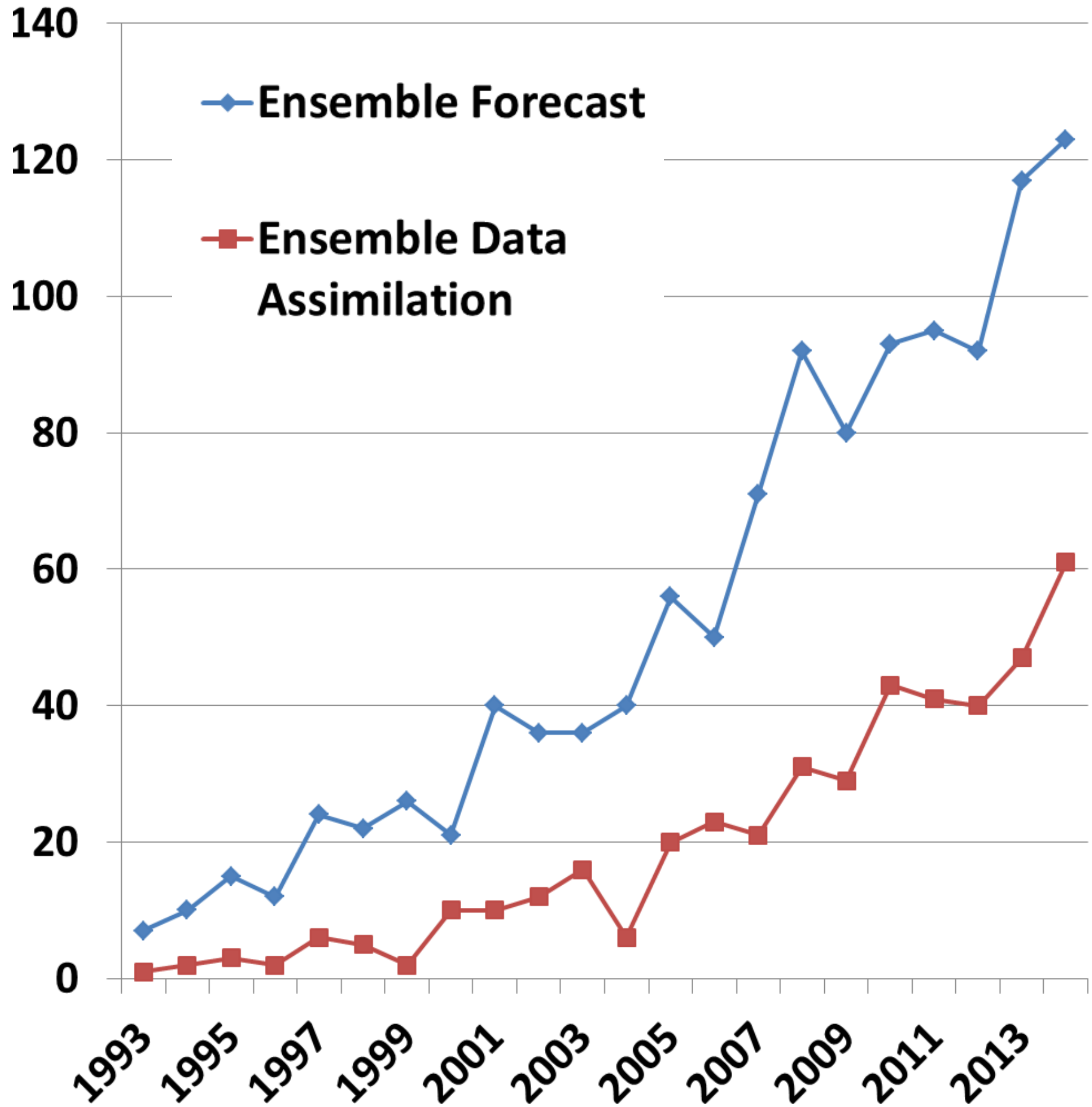
Current Issues and Challenges in Ensemble Forecasting

Carolyn Reynolds (NRL) and Junichi Ishida (JMA)
With contributions from WGNE members

30th WGNE
College Park, MD, 23-26 March 2015

- **Recent trends in ensemble-related research**
- **Impact of improved initial conditions**
- **Accounting for model uncertainty**
- **Calibration and post processing**
- **Multi-model ensemble issues and questions**

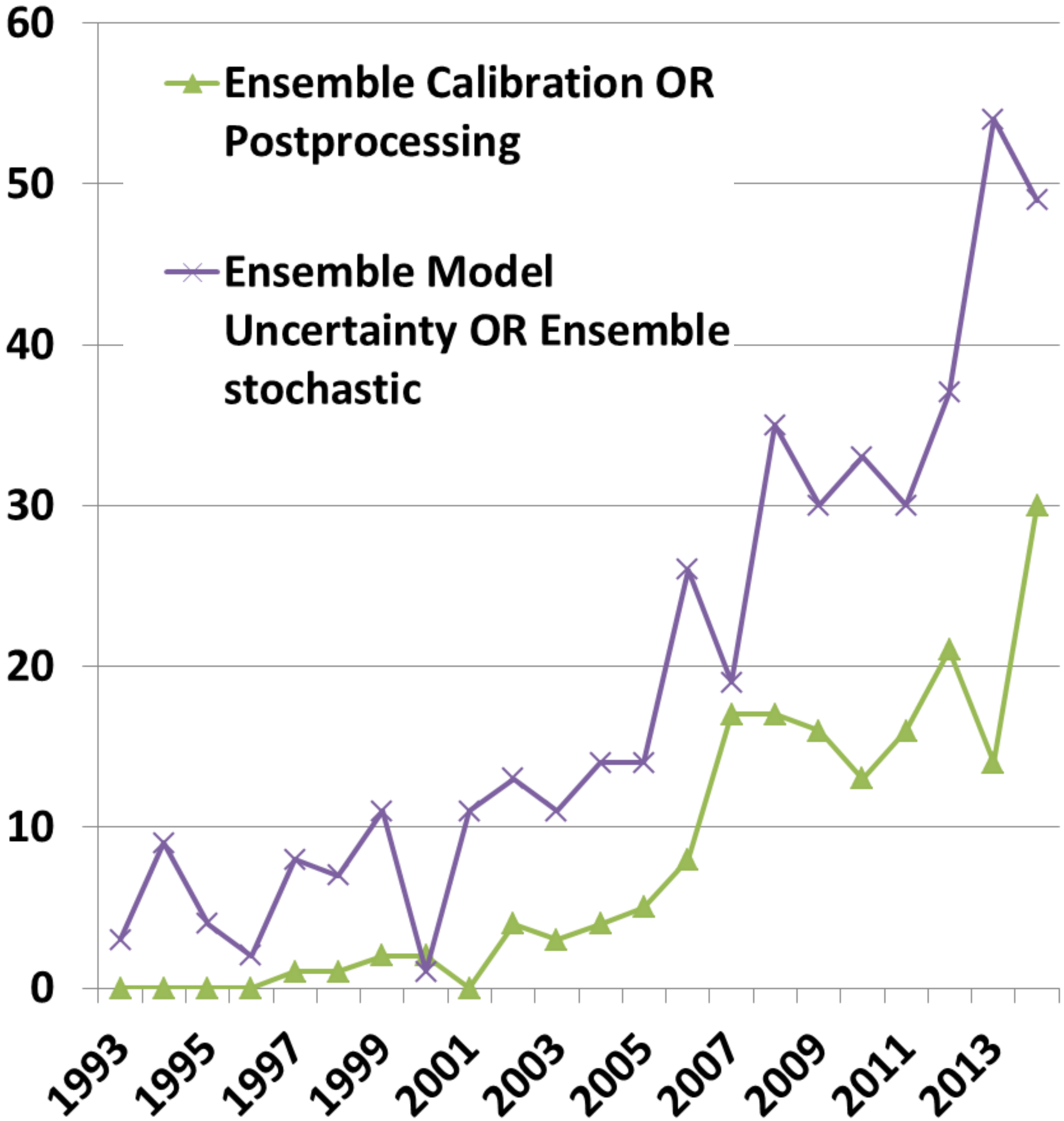
*Number of Article/Year with These Words in the Abstract**



Research in ensemble forecasting and ensemble data assimilation has been climbing steadily since the 1990s.

*AMS journals only

*Number of Article/Year with These Words in the Abstract**



Research in Model Uncertainty grows rapidly in the last three years.

Interest in calibration and post-processing also substantially larger than in the early 2000s.

***AMS journals only**

Integrating DA and Ensembles: Impact of Improved Initial Conditions

Environment Canada Environnement Canada

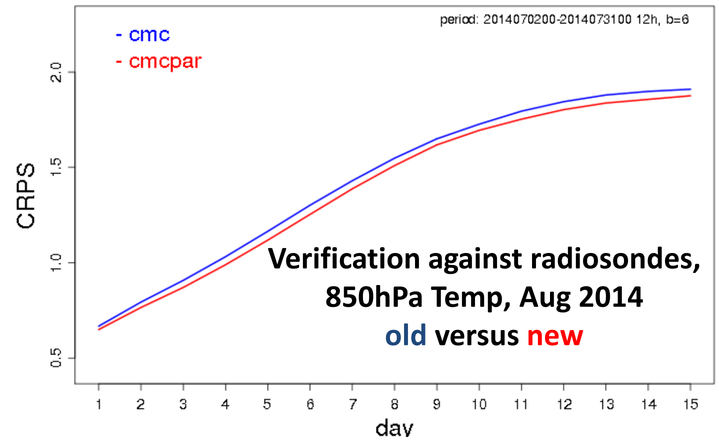
Main changes to the analysis component (EnKF)

- *ensemble size: 192 → 256 members*
- *horizontal resolution: 66 → 50 km*
- *time step: 20 → 15 min*
- *data assimilation:*
  - *RTTOV-10*
  - *4D assimilation of radiosondes*
  - *new bias correction method*
  - *GPS-RO from 1km*
- *further perturbations to the physics (e.g. orographic blocking bulk drag coefficient, thermal roughness length over oceans)*

Main changes to the forecast component

- *horizontal resolution: 66 → 50 km*
- *time step: 20 → 15 min*
- *new method to evolve SST and sea-ice fields*
- *further perturbations to the physics (e.g. orographic blocking bulk drag coefficient, thermal roughness length over oceans)*

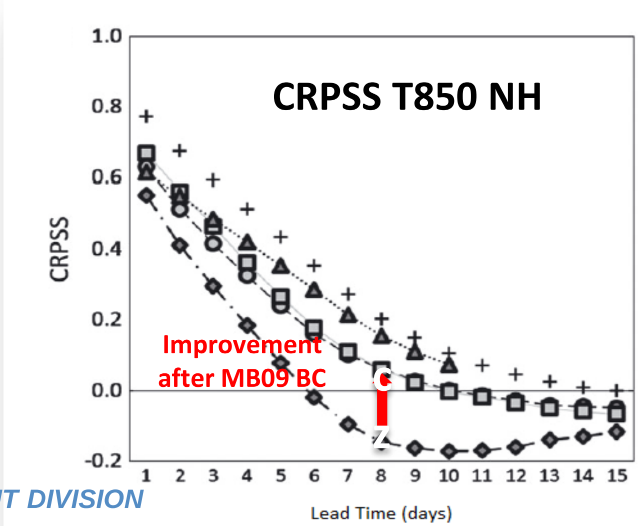
Overall 6-h improvement in forecast skill for atmospheric variables.



** Material kindly provided by Peter Houtekamer and Normand Gagnon*

CPTEC Ensemble Prediction System

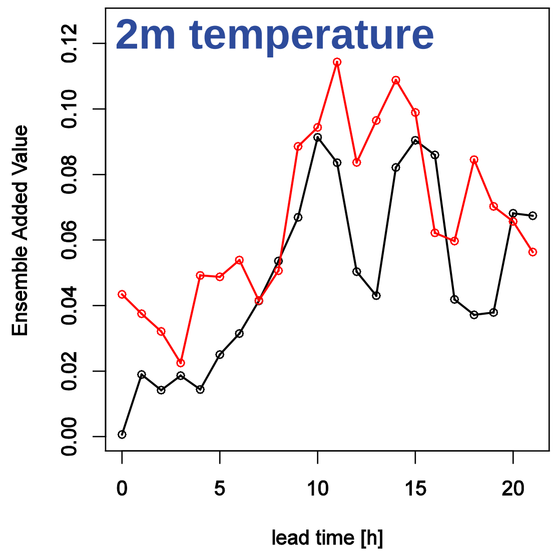
Crosses = NCEP EPS; Triangles = KMA EPS; Diamonds = CPTEC EPS (operational) ; Circles = CPTEC EPS-MB09 (two additional variables, surface pressure and specific humidity, and extended analysis region); Squares = CPTEC EPS-MB09 BC (includes bias correction)

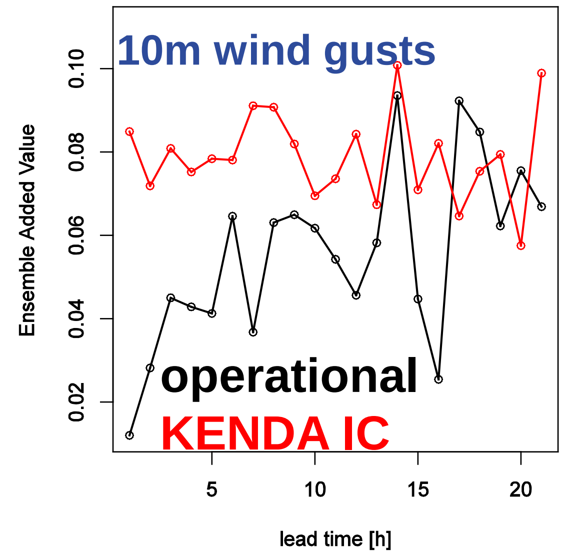


Integrating DA and Ensembles: Impact of Improved Initial Conditions

DWD
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Use of IC from the Kilometer scale ENsemble Data Assimilation (KENDA) based on the LETKF scheme (Hunt et al., 2007)

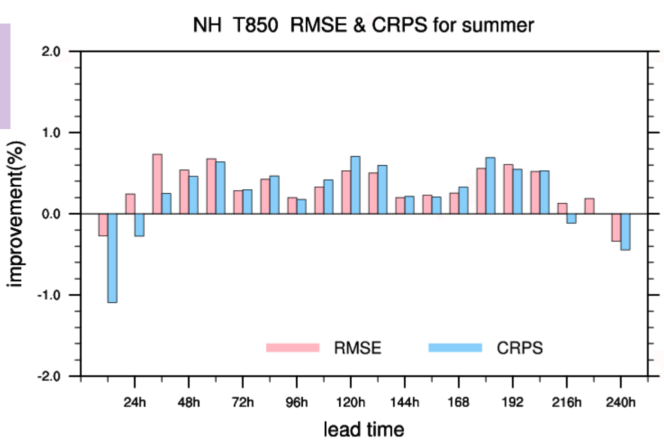


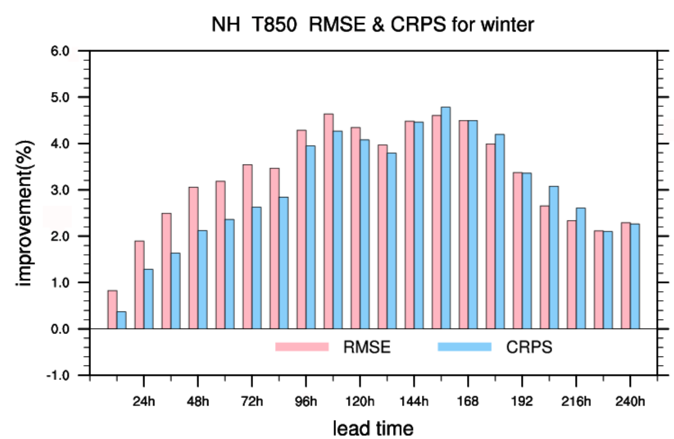


Hybrid Ensemble 4DVAR D.A System (in operation since ‘13 in KMA) benefits global EPS system as well as global deterministic forecast through high quality initial conditions

KMA

850hPa T





Growing Interest in Accounting for Model Uncertainty

Strategic Goals for NWP Centres: Minimising RMS error or maximising forecast reliability, T. Palmer, U. Oxford, WWOSC, August 2014

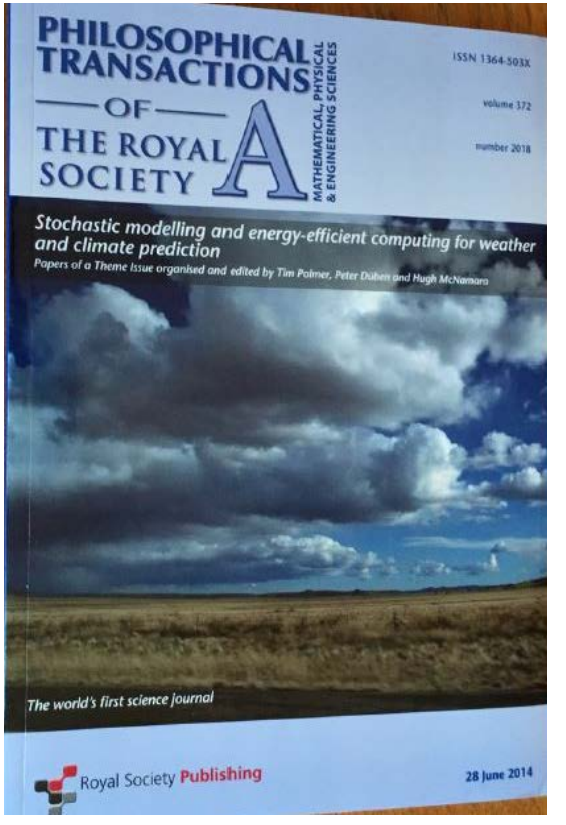
Conclusions

Palmer T.N., 2012: Towards the probabilistic earth-system simulator: a vision for the future of weather and climate prediction. Quart J R Met Soc, 138, 841-861 (Royal Met Soc Presidential Address)

Stochastic parametrisation improves probabilistic scores and can reduce systematic errors. It does not (necessarily) reduce the rms error of deterministic forecasts.

Primary headline metrics should measure the usefulness of weather forecasts for real-world decision making. RMS error/ACC of Z500 does not measure this; CRPSS does.

If RMS Error and Anomaly Correlation Coefficient remain the primary headline metrics to evaluate an NWP Centre's performance, the development of parametrisations with (e.g.stochastic) representations of their own uncertainty will not be given first priority by model development teams.



Recommendations from EUMETNET Joint PHY-EPS Workshop 2013:

- ***Introduce stochasticity only where appropriate (maintain physical meaning).***
- ***Sensitivity studies and process studies, in addition to predictability studies, are necessary to understand impacts.***
- ***Parameter perturbations useful diagnostic to understand spatio-temporal characteristics of uncertainty.***

Parameterization of Moist Processes for Next-Generation Weather Prediction

NOAA Center for Weather & Climate Prediction, College Park, Maryland

January 27-29, 2015

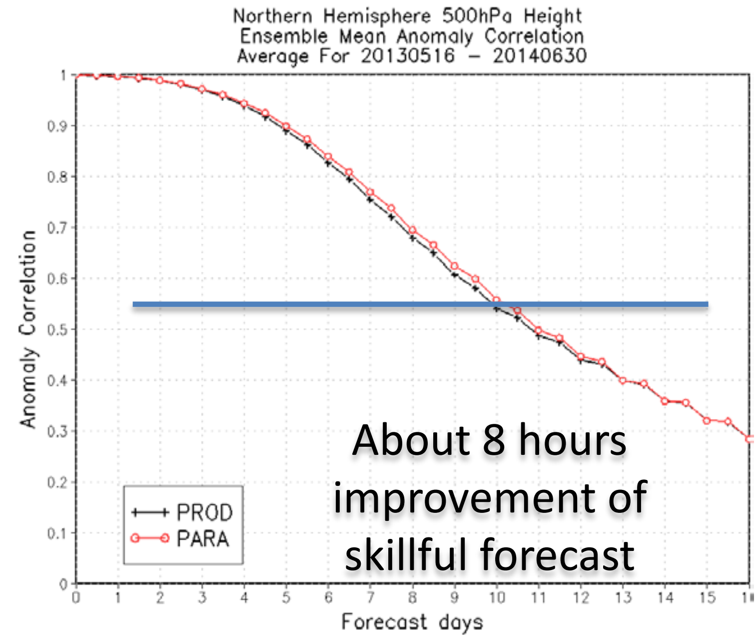
Probability distributions are useful in two distinct contexts: 1) for representing variability at scales below or approaching the model resolution, and 2) to describe uncertainty and improve spread-skill relationships in probabilistic ensemble forecasts. ***It is natural to expect that model uncertainty could be estimated directly by parameterizations and expressed by, for example, drawing the parameterization tendency from a distribution of expected outcomes.***

However, the parameterization community is not yet ready to provide estimates of state-dependent parameterization error to replace current ad-hoc estimates of model error to increase ensemble spread. Data assimilation, sensitivity assessment, and parameter estimation are the most useful current approaches for developing understanding of the response of model output to changes in parameters, how this response maps onto the resolved scales, and how the local and grid scale response changes with environment, flow, etc. Nonetheless, ***ad hoc perturbations to physical tendencies remain the most effective solution for maintaining the dispersion of ensembles through the duration of a forecast***.

Accounting for Model Uncertainty

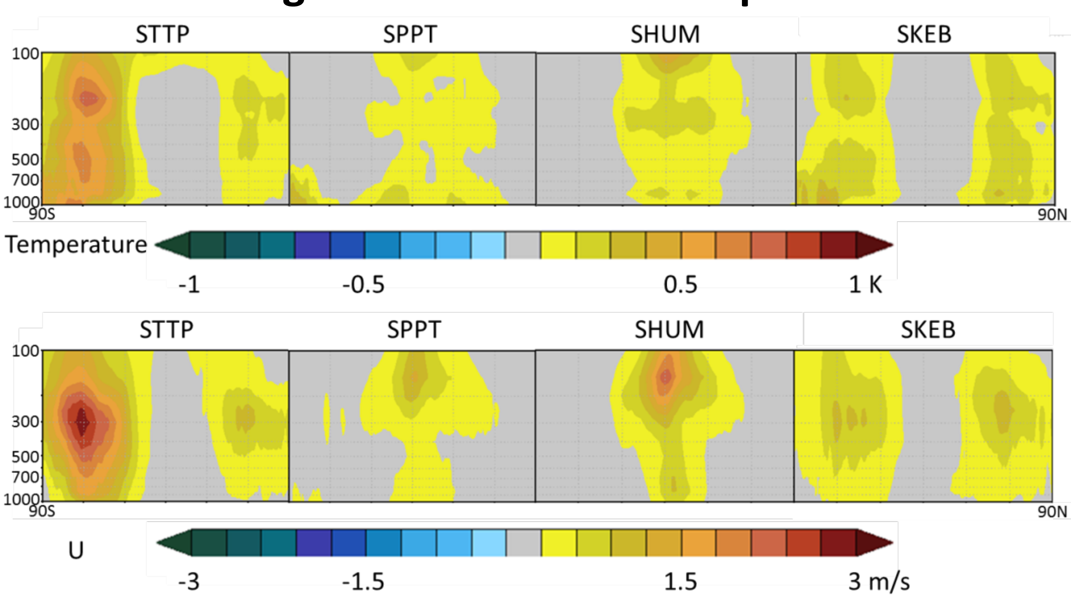
Next GEFS (V11.0.0) configuration
Yuejian Zhu (EMC/NCEP/NWS/NOAA)

- Model: GFS SL (V10) from GFS Euler model (V9.0.1)
- Increased horizontal and vertical resolution
- Initial conditions: EnKF (from BV-ETR)
- Plans: test SKEB, SPPT, SHUM, Stochastic perturbed land surface (current, STTP)

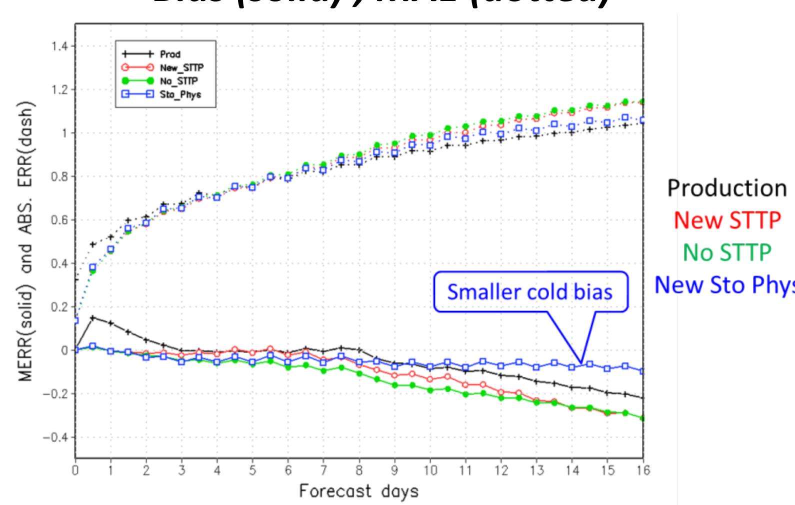


Impact of Model Forcing : GEFS SL T574, *P. Pegion, W. Kolczynski, J. Whitaker, T. Hamill*

Change in 120-h Ensemble Spread



Tropics: 850-hPa Temp
Bias (solid) , MAE (dotted)



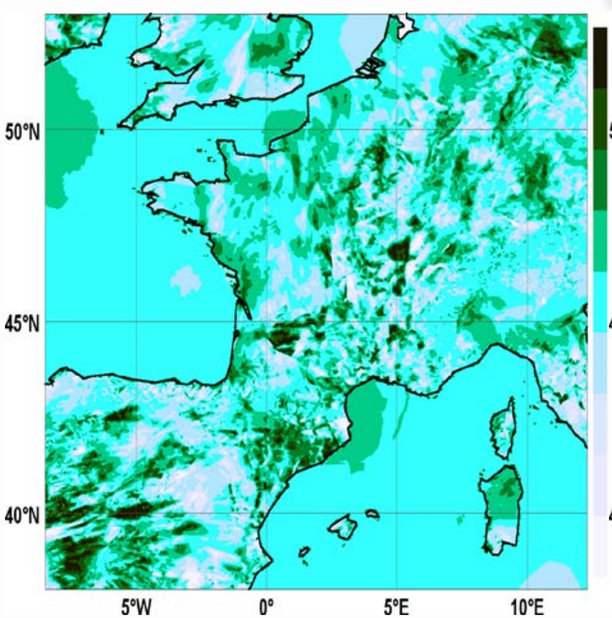
Different schemes address different issues, may be complementary

Accounting for Model Uncertainty

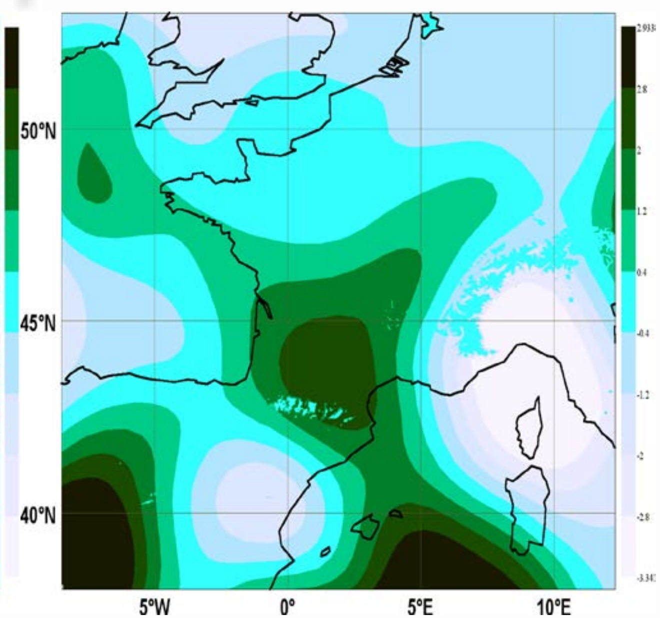
Interaction between EDA and surface perturbations

Ensemble perturbations from the AROME EDA (ensemble data assimilation) are improved when simple random noise is used at the surface, instead.
i.e. a better surface perturbation scheme should be developed in EDA.

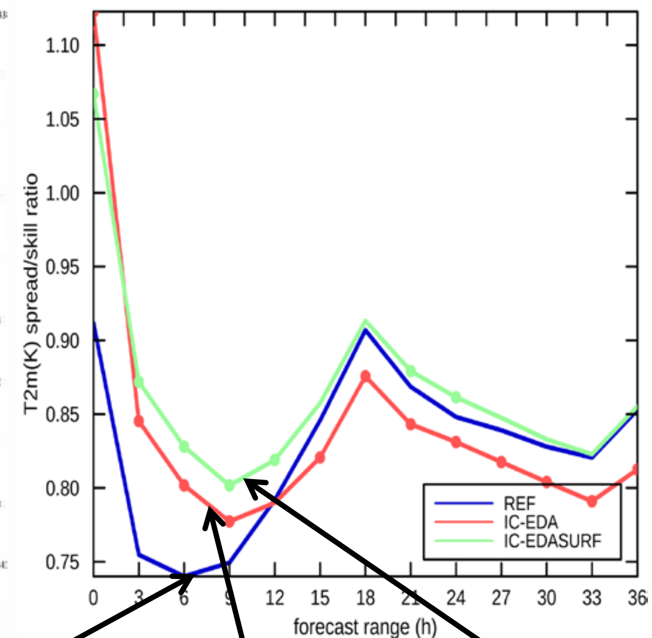
Ts perturbation (EDA)



Ts perturbation (random)

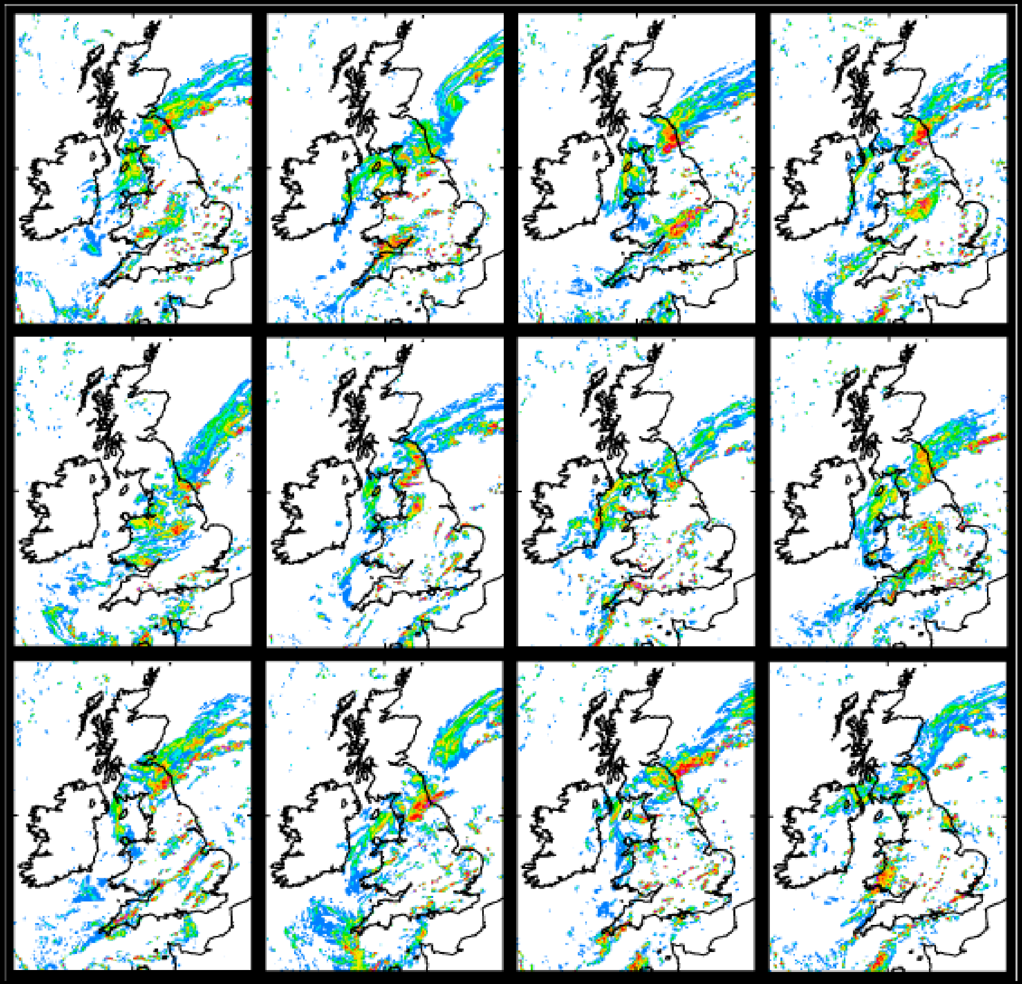


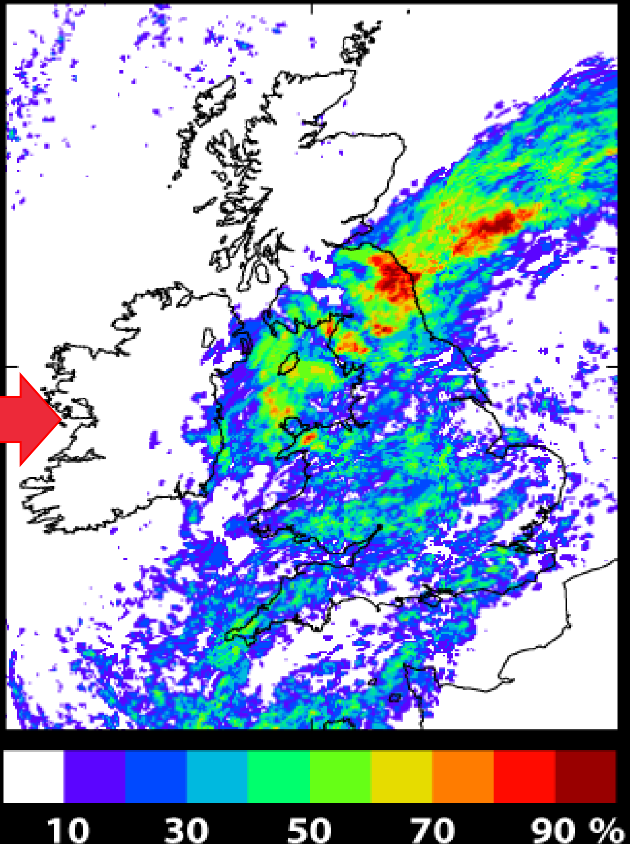
T2m spread/skill ratio
(higher is better)



METEO FRANCE

MOGREPS-UK
2.2km ensemble

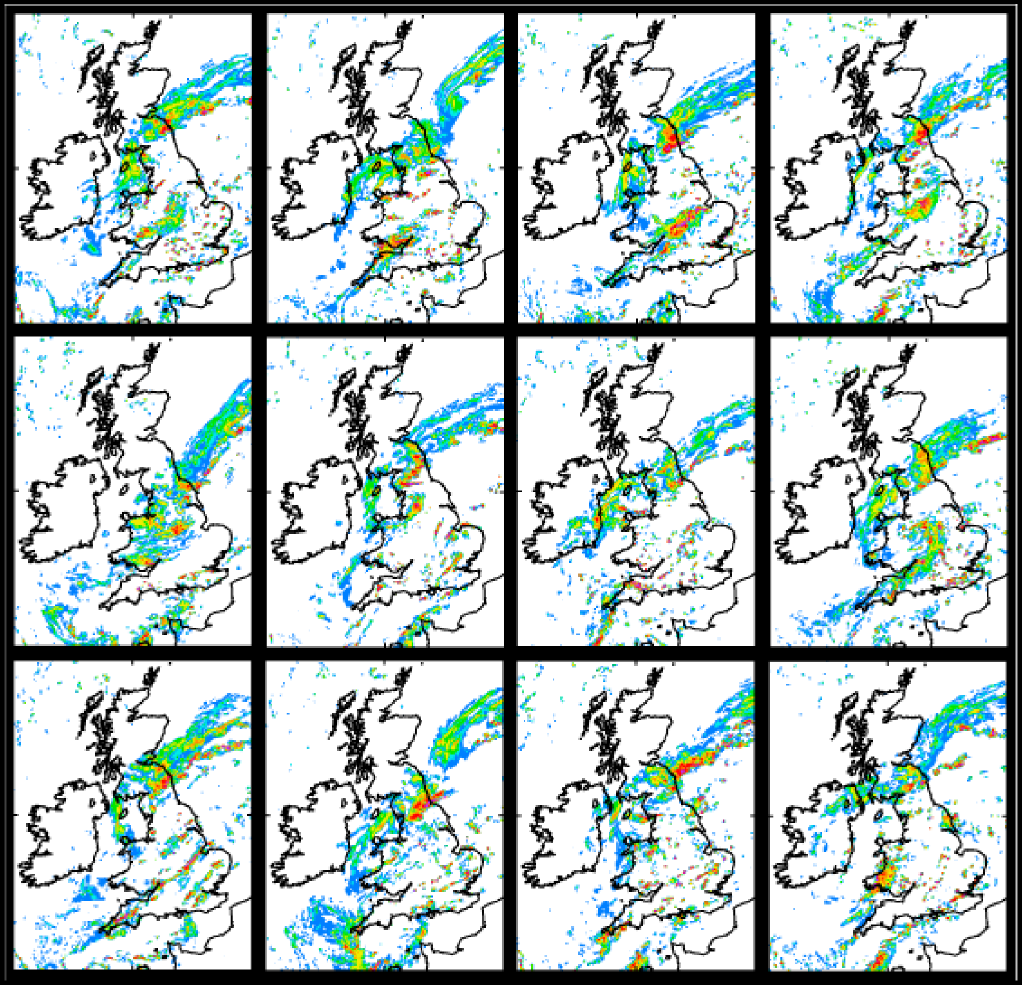


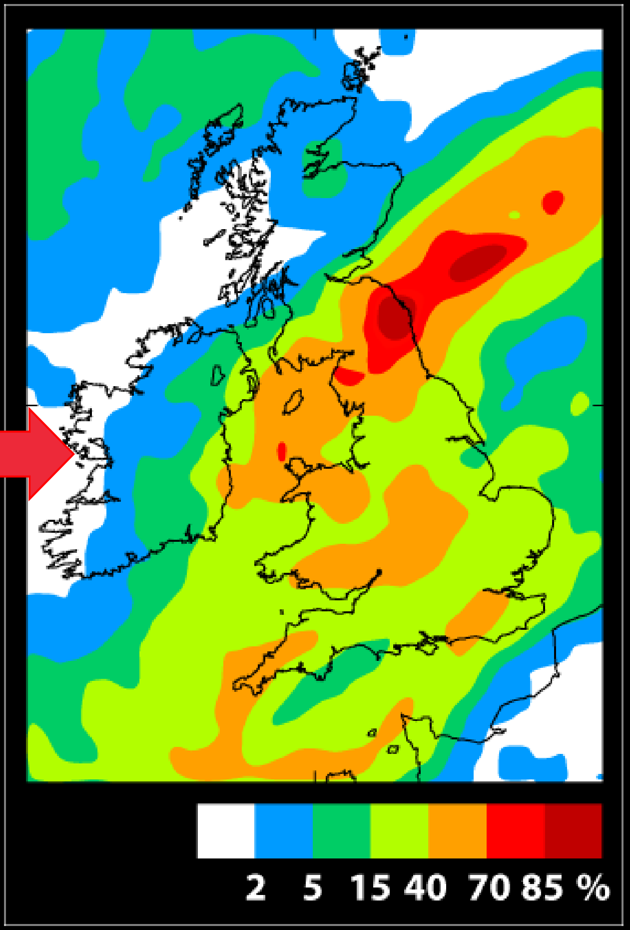


Undersampling leaves "holes" of zero-probability where showers could still occur

© Crown copyright Met Office

MOGREPS-UK … with Neighbourhood processing





Holes filled in

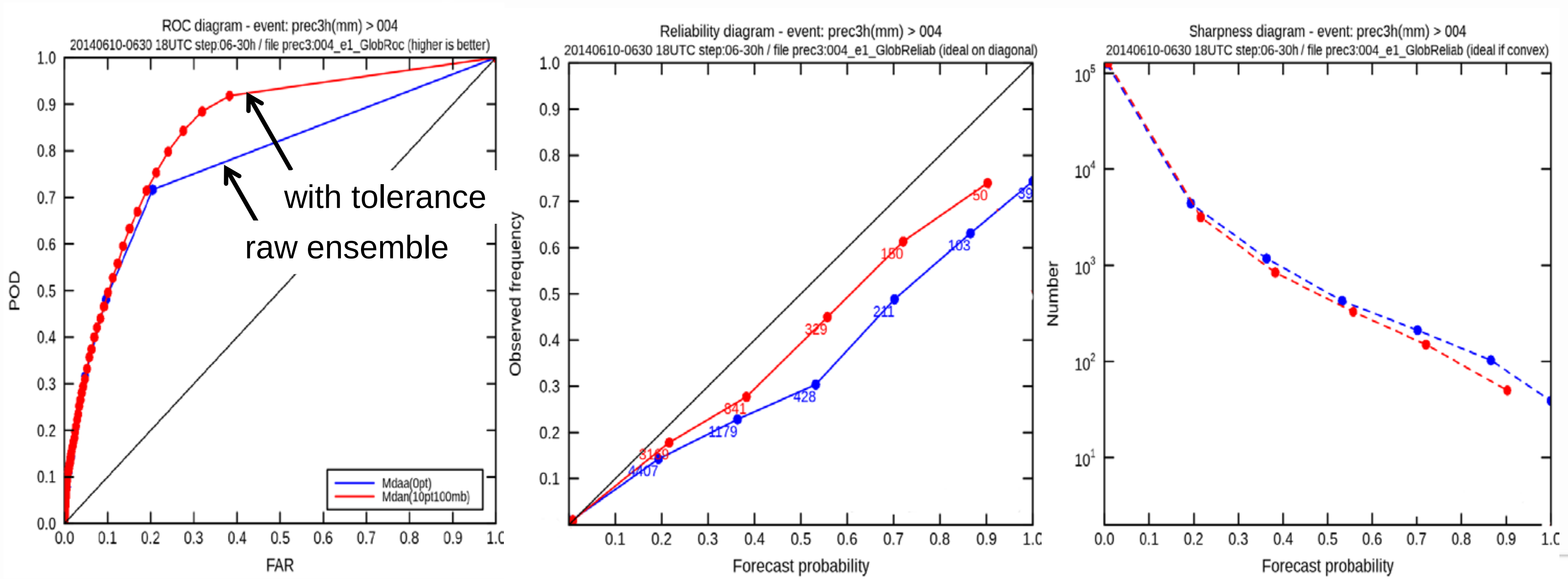
© Crown copyright Met Office

Neighbourhood methods for high precipitation forecasts

Ensemble scores improve when spatial tolerance is introduced in the forecast PDF computation :

- improved reliability & ROC metrics
- negligible loss of sharpness
- largest effect comes from improved membership

Performance is sensitive to details of the method used.



Number of multi-model ensembles are growing

Mesoscale: TIGGE-LAM, NOAA SREF, AEMET-SREPS, SESAR, CAPS, HFIP
Global 1-2 weeks: NAEFS, NUOPC, TIGGE, HIWPP, ICAP
Subseasonal to seasonal: NMME, DEMETER, S2S

Why do multi-model ensemble often outperform single model ensembles? Is the improvement in skill due to larger ensemble size or to combining signals? (extra slide)

International Conference on S2S prediction, 10-13 Feb 2014

Differences in Skill and Predictability in Multi-Model Ensembles

Timothy DelSole

George Mason University, Fairfax, Va and
Center for Ocean-Land-Atmosphere Studies, Calverton, MD

1. Proposed an objective procedure for deciding if the skill of a combined forecast is significantly higher than a single forecast.
2. Skill of each model in NMME is significantly enhanced by combining it with other models, at least for some lead time and target month.
3. The skill improvement comes from combining different signals, not from increasing ensemble size.

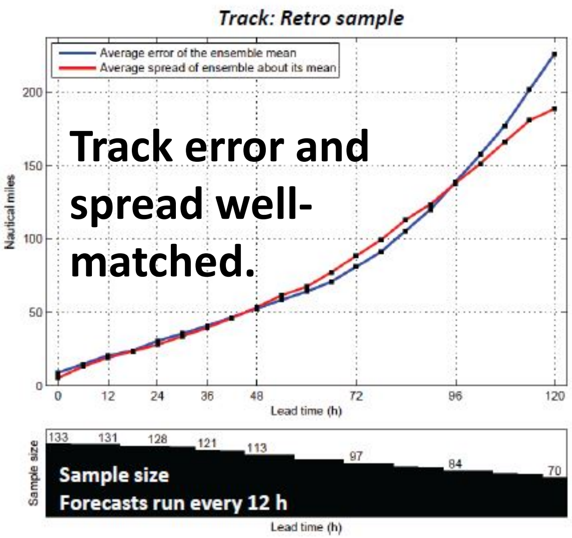
- ***How does one combine multi-model forecasts of unequal skill? Equal weights competitive with more complex schemes (DelSole et al. 2012, Sansom et al. 2013, …)***

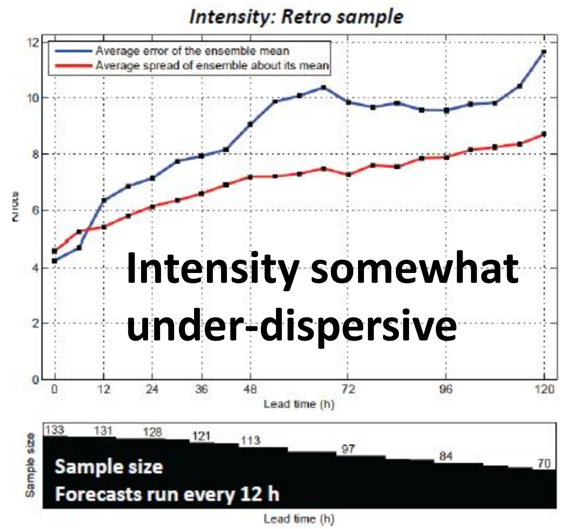
- ***Tradeoffs between independence from multi-models vs. focusing resources on one system.***

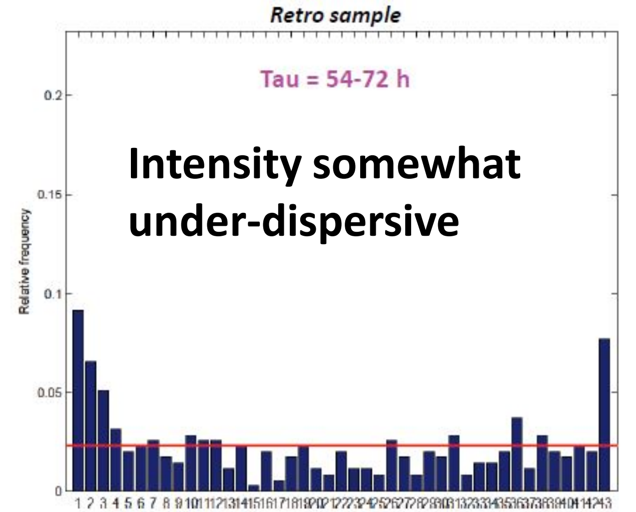
- ***Issues of latency, data transfer reliability, etc.***

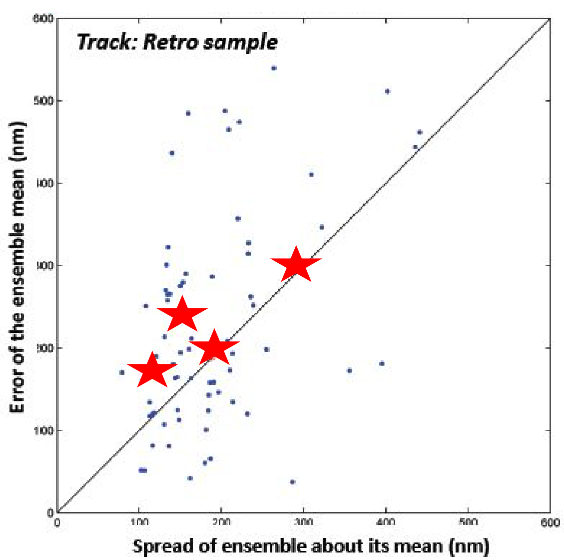
NOAA Hurricane Forecast Improvement Program Multi-Model Ensemble

HWRF EPS (27/9/3 km, 42 levels) – 20 members
GFDL EPS (55/18/6 km, 42 levels) – 10 members
COAMPS-TC EPS (27/9/3 km, 40 levels) – 10 members

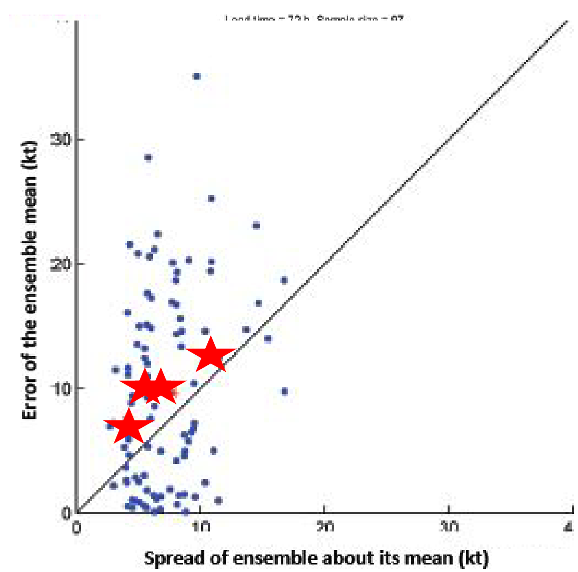








Larger 120-h track error (left) and larger 72-h intensity error (right) associated with larger ensemble spread, on average



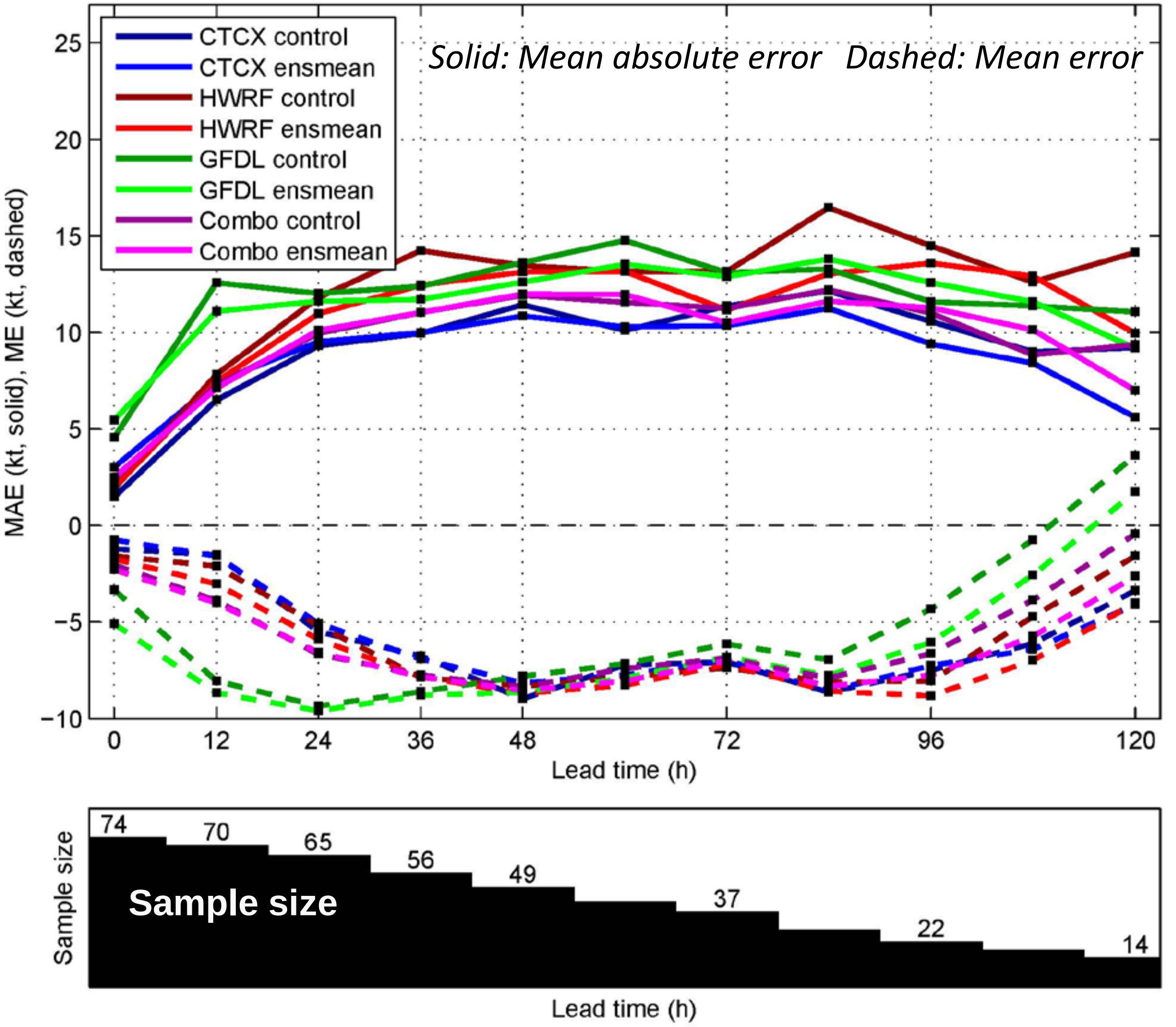
NRL Analysis

Extra slides

Hurricane Multi-Model Ensembles

NOAA Hurricane Forecast Improvement Program multi-model ensemble.

HWRF EPS (27/9/3 km, 42 levels) – 20 members
GFDL EPS (55/18/6 km, 42 levels) – 10 members
COAMPS-TC EPS (27/9/3 km, 40 levels) – 10 members



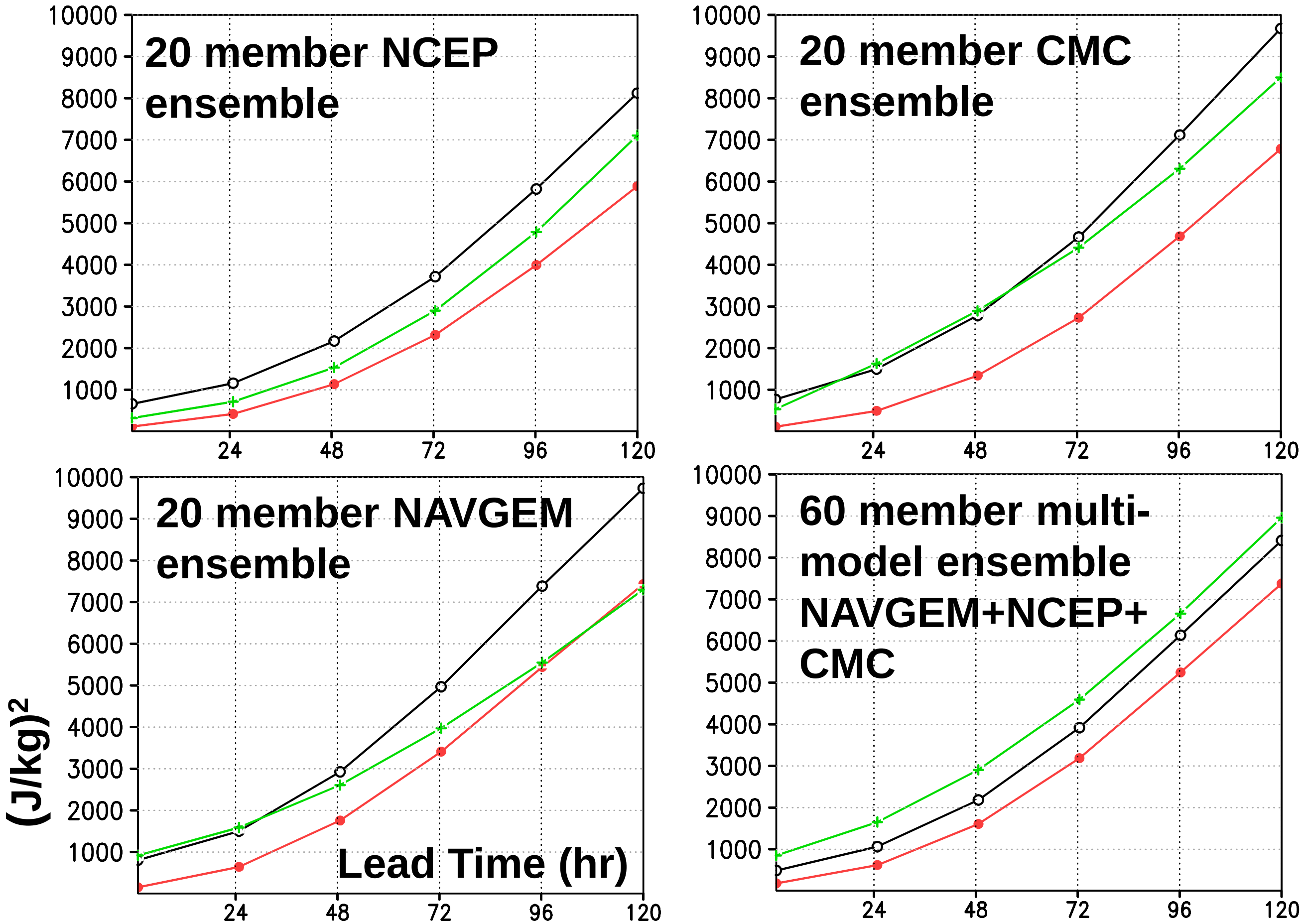
For individual model, ensemble mean has improved accuracy relative to the control

Combined ensemble mean has accuracy similar to consensus of three control members

Control forecasts:
COAMPS-TC: C00C
HWRF: HW00
GFDL: GP00
Combo : Consensus of C00C, HW00, and GP00

Ensemble mean requirements:
COAMPS-TC: 9 of 11 members
HWRF: 17 of 21 members
GFDL: 8 of 10 members
Combo: 34 of 42 members

NRL: Quantifying Model Inadequacy from Multi-model Ensembles *(E. Satterfield)*



TV= TOTAL ERROR VARIANCE

TVS=PORTION OF TV THAT PROJECTS ONTO THE SPACE OF ENSEMBLE PERTURBATIONS

VS=ENSEMBLE VARIANCE

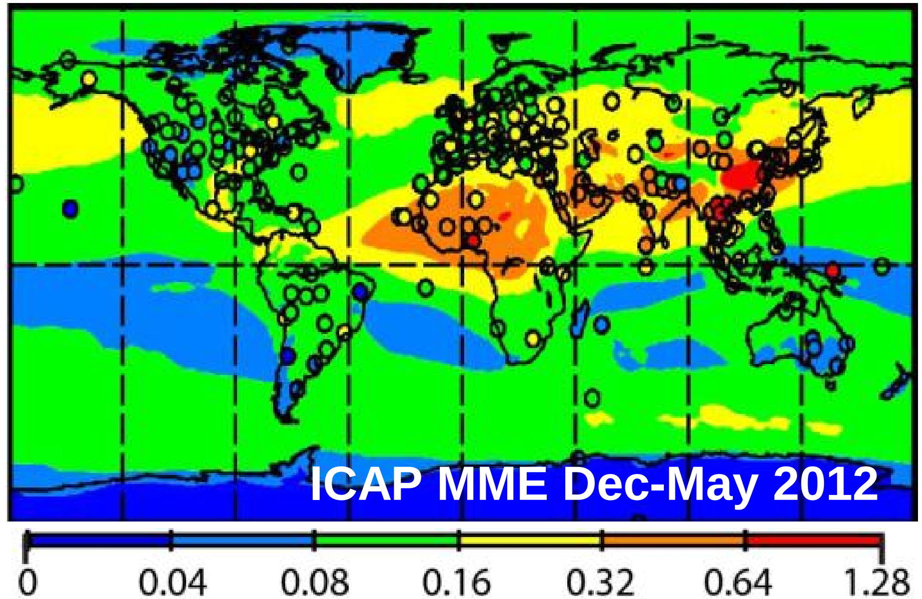
For a perfect ensemble TV= TVS=VS. For an ensemble that correctly represents the second moment of the probability distribution of the state, VS=TV would hold.

- Project explores aspects of multi-model ensemble prediction systems with the goal of improving single model ensemble forecasts
- Improving the quality of the Navy ensemble will lead to improved probabilistic prediction and uncertainty estimation at longer lead times
- **It will also improve the flow dependent error covariance estimates at shorter lead times used in Hybrid DA schemes.**

NRL Developed ICAP Global Multi-model Aerosol Forecast Ensemble:

BSC, ECMWF, FNMOC/NRL, JMA, NASA, NOAA, UKMO

- **The International Cooperative for Aerosol Prediction (ICAP) is a grass roots organization of aerosol forecast developers to share best practices and speak with a common voice on aerosol observation needs for DA.**
- **Ensemble open to any consistent quasi-operational global aerosol model. Currently working on AOT and surface concentrations for multi species and dust only versions, but looking towards 3 full dimensions.**
- **Specific error metrics are kept by centers, ensemble products distributed via GODAE server.**
- **As expected from a multi model ensemble, the ICAP MME has the best RMSE scores and a more consistent bias distribution over the globe.**



Ensemble Mean AOT (550 nm)