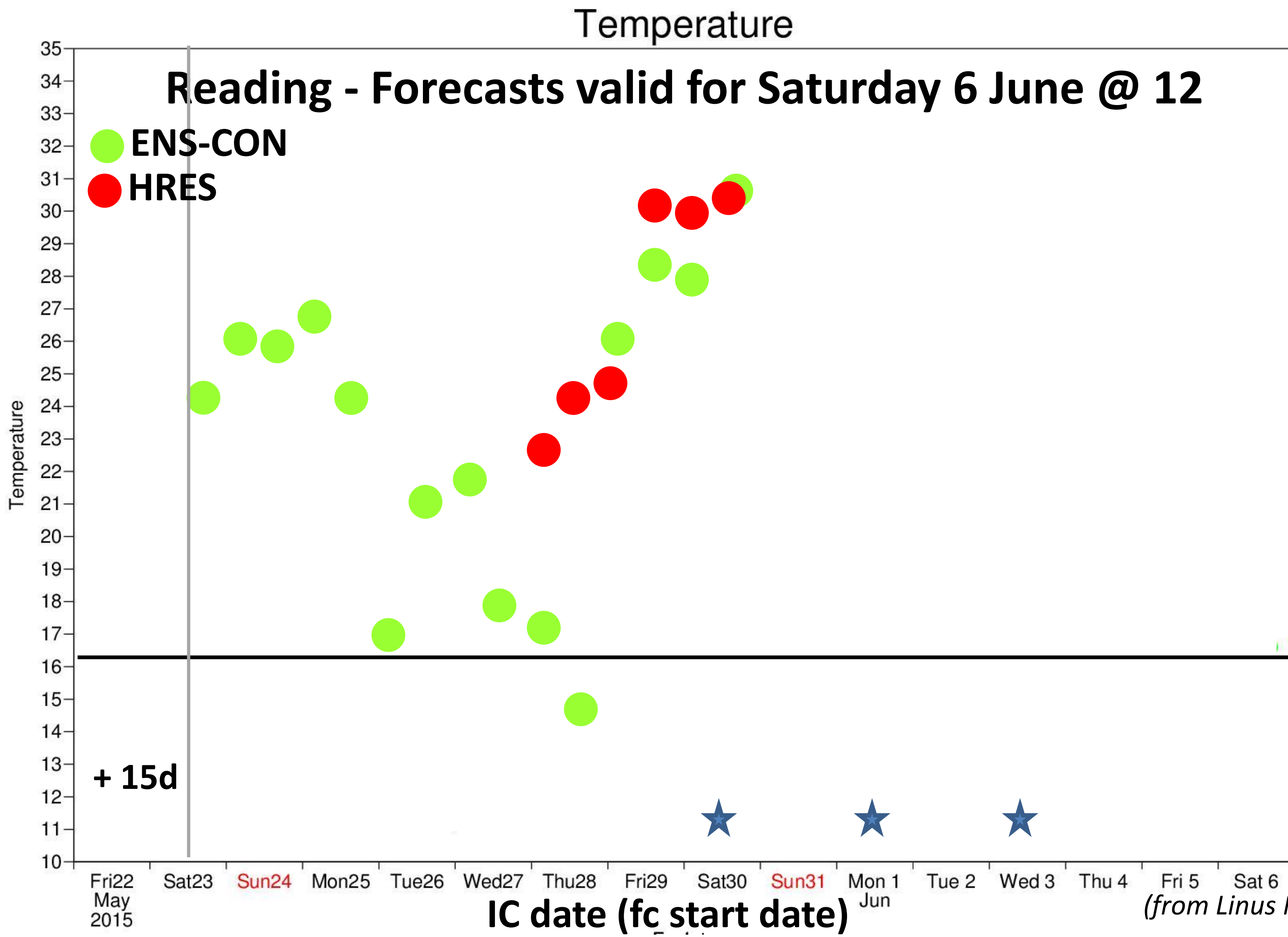
Estimating uncertainty with the ECMWF ensembles

Roberto Buizza

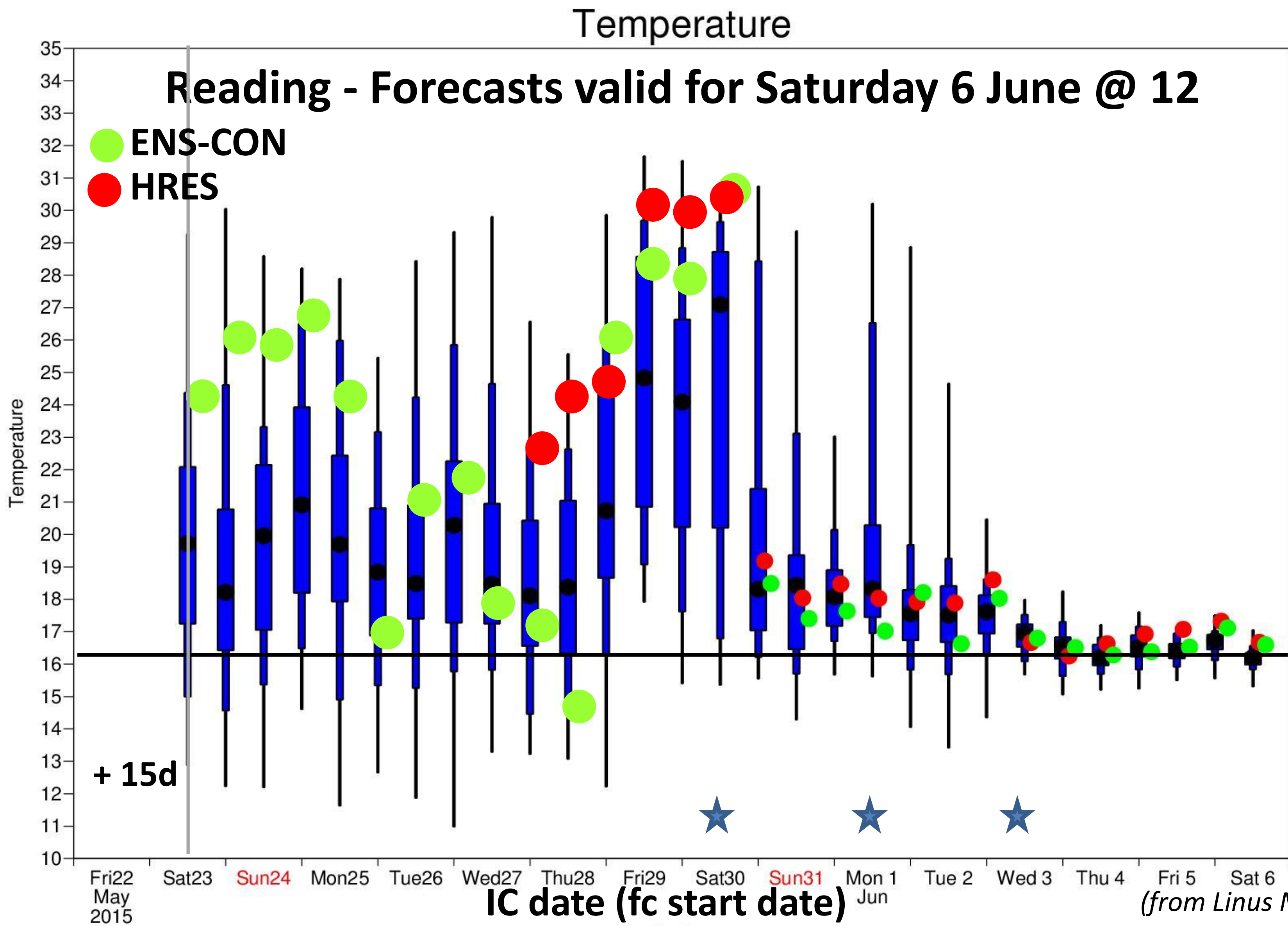
European Centre for Medium-range Weather Forecasts

Why do we need ensembles?



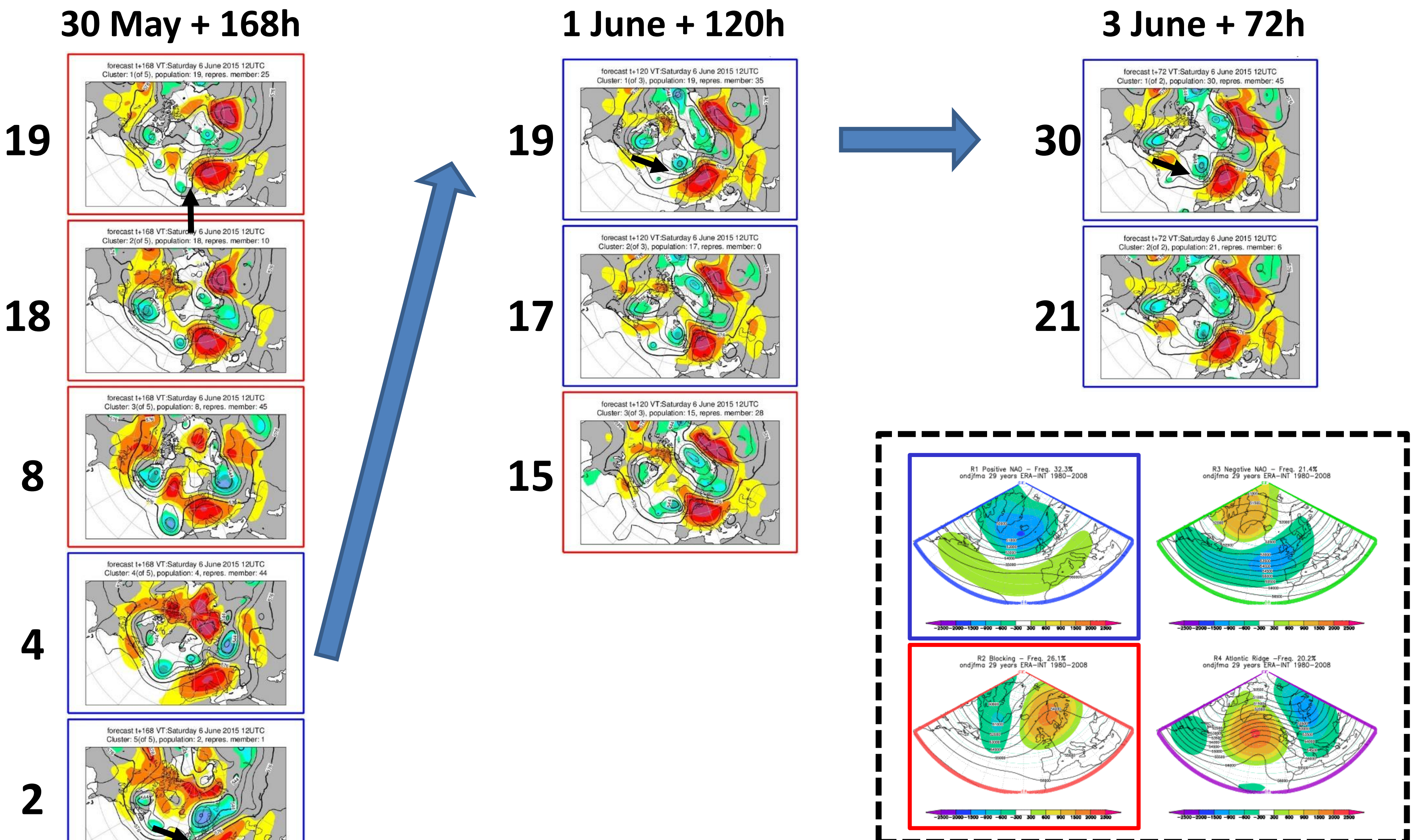
(from Linus Magnusson)

Why do we need ensembles?



(from Linus Magnusson)

Why do we need ensembles?

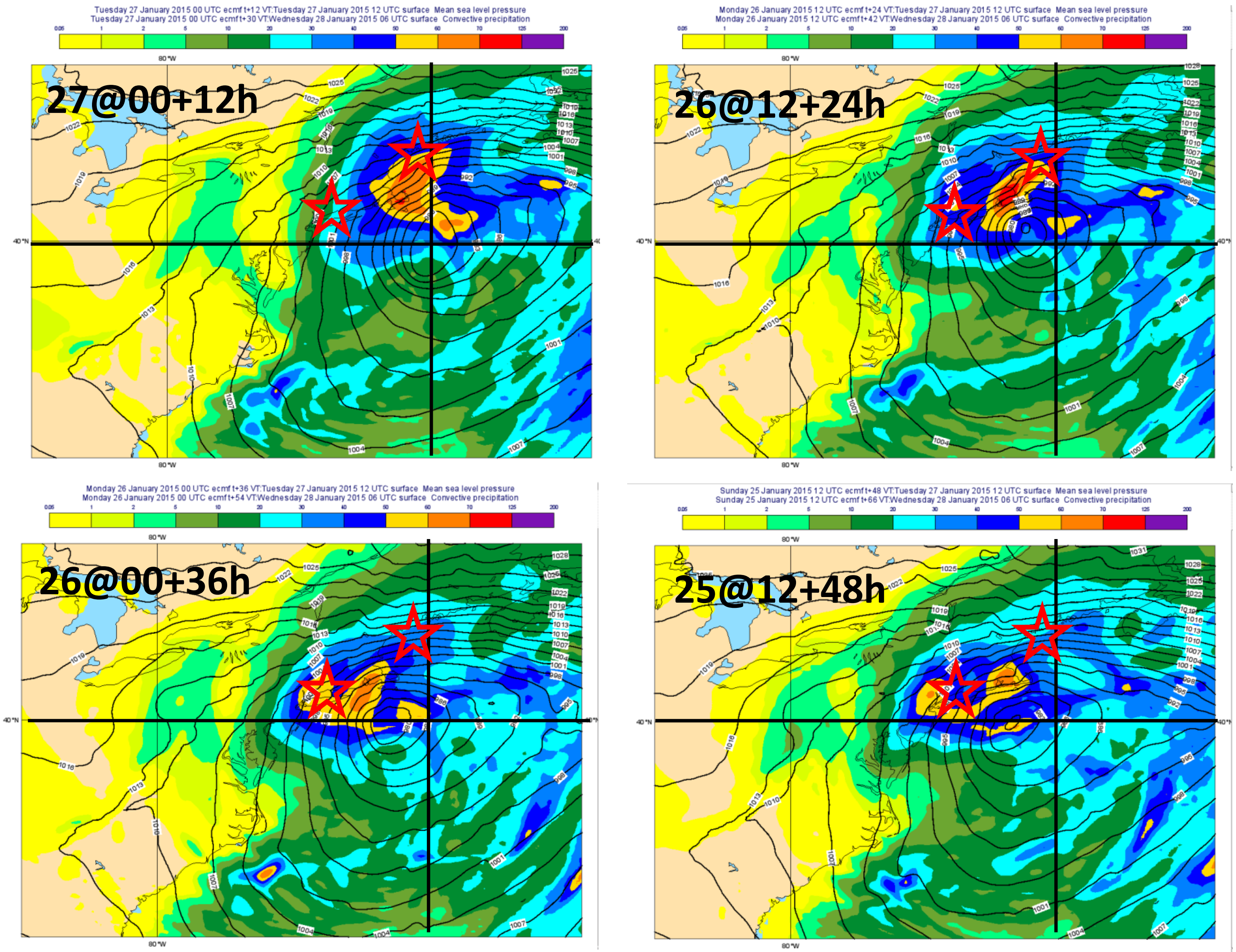


(from Linus Magnusson and Laura Ferranti)

We need ENS even in the short range (US Storm)

Single HRES fcs failed to positioned correctly the storm, and this led to snowfall overestimation for NY of in the 24-36-48h forecasts.

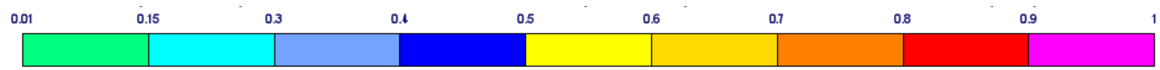
MLSP+TP maps show a 150-200 km eastward shift in the storm centre.

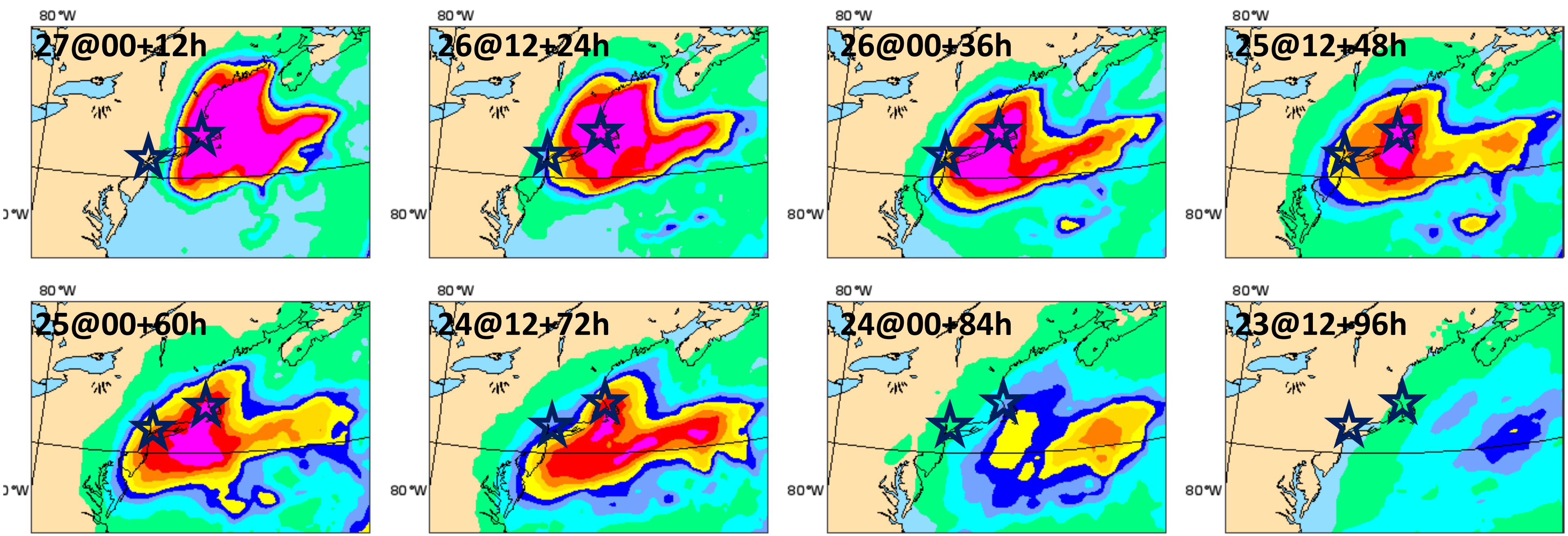


US Storm, 27-28/01/2015: ENS PR fcs

ENS-based probabilistic forecasts can be used to estimate the level of confidence (predictability) of single forecasts. They show that NY was closer to the edge of the area of high probability of +30mm of precipitation, indicating higher uncertainty.

ENS PR[TP(27@00–28@06)>30mm]

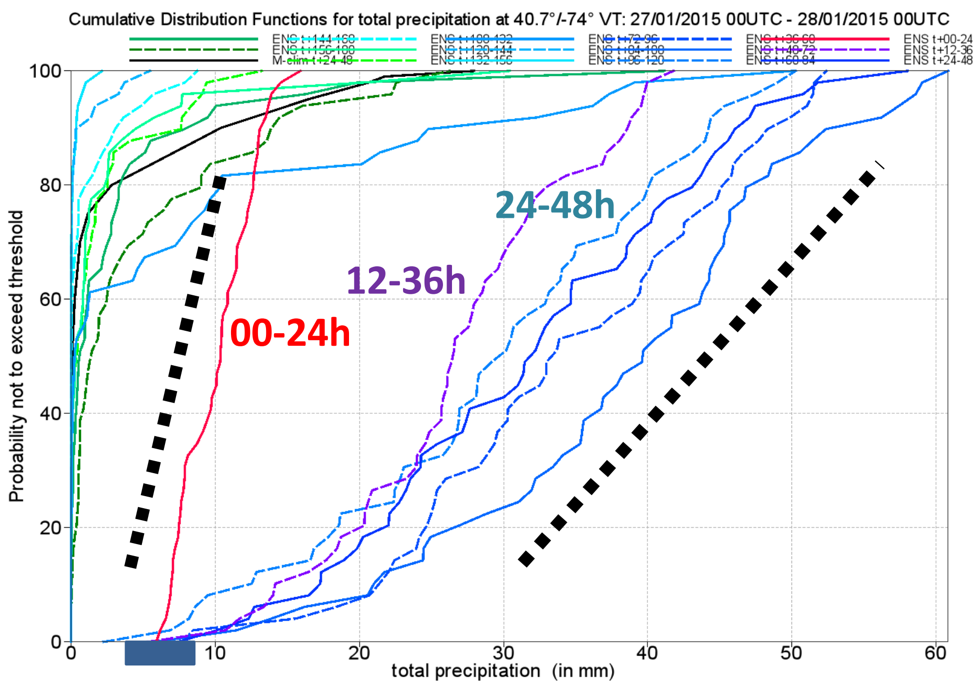




US Storm, 27-28/01/2015: ENS CDF to estimate confidence

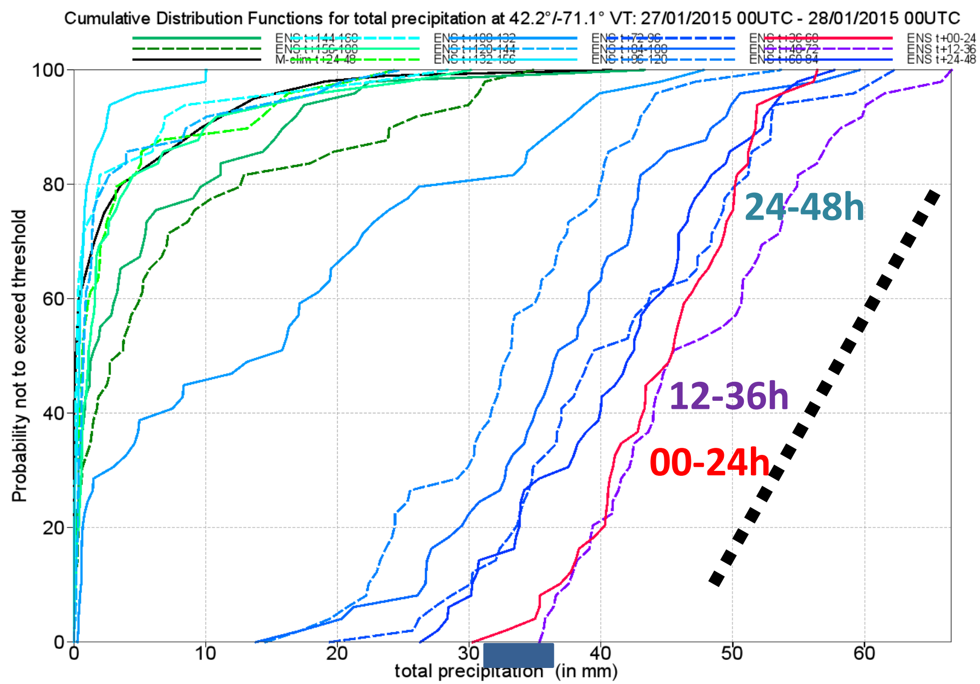
ENS-based probabilistic forecasts expressed in terms of CDF shows that the fcs for NY were more uncertain (the slope of the CDF curves is steeper) than the fcs for Boston.

New York



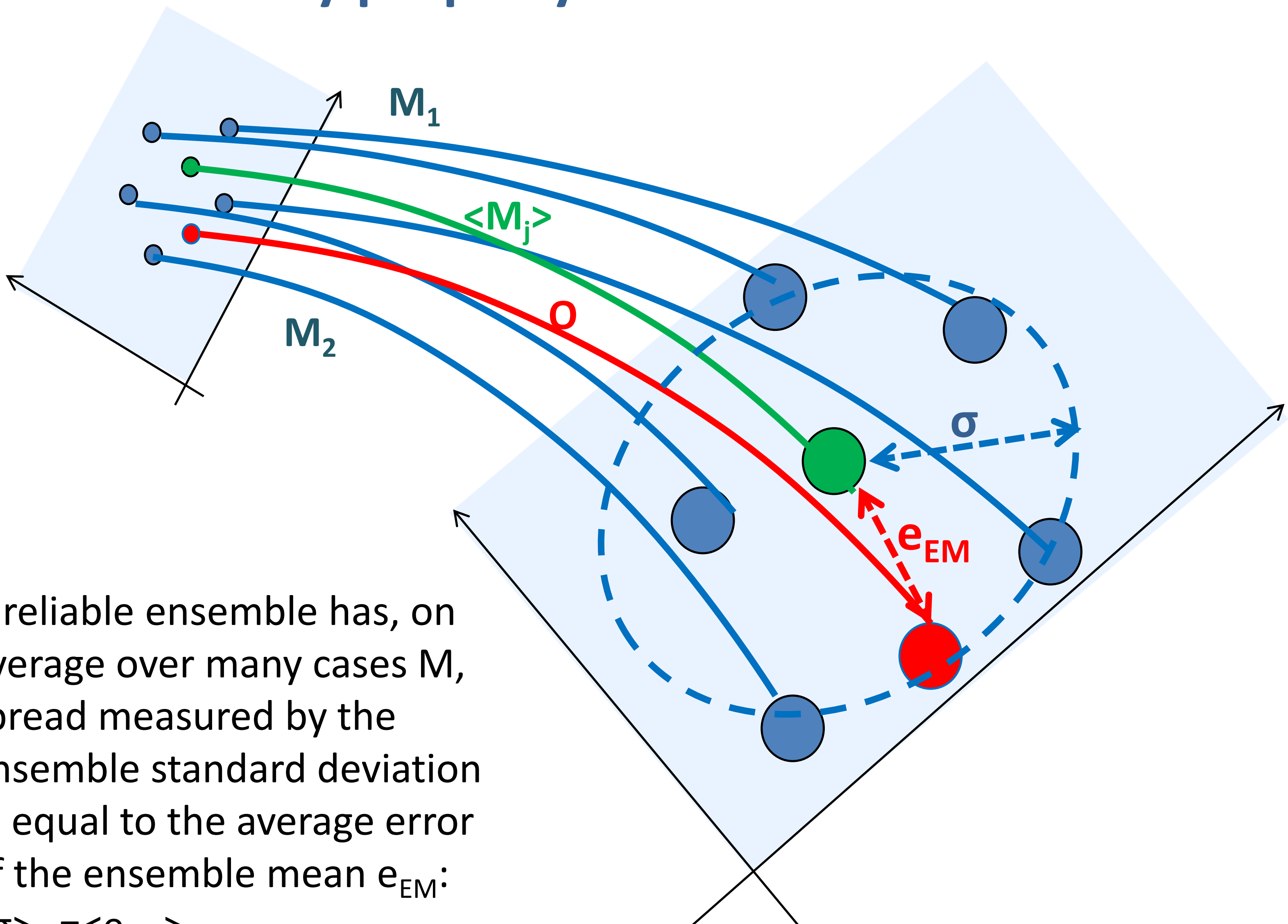
Obs (NEXRAD) 5-10 mm

Boston



Obs (NEXRAD) 30-35 mm

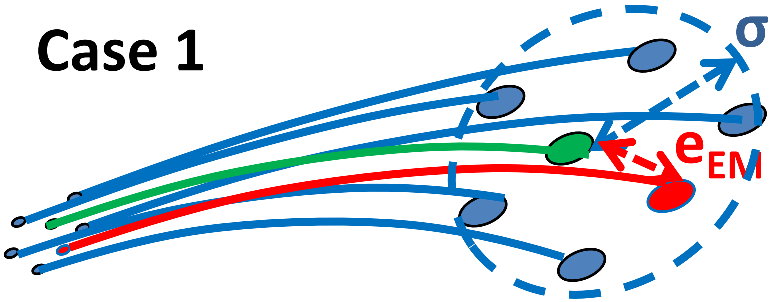
A necessary property for ENS to be valuable: reliability



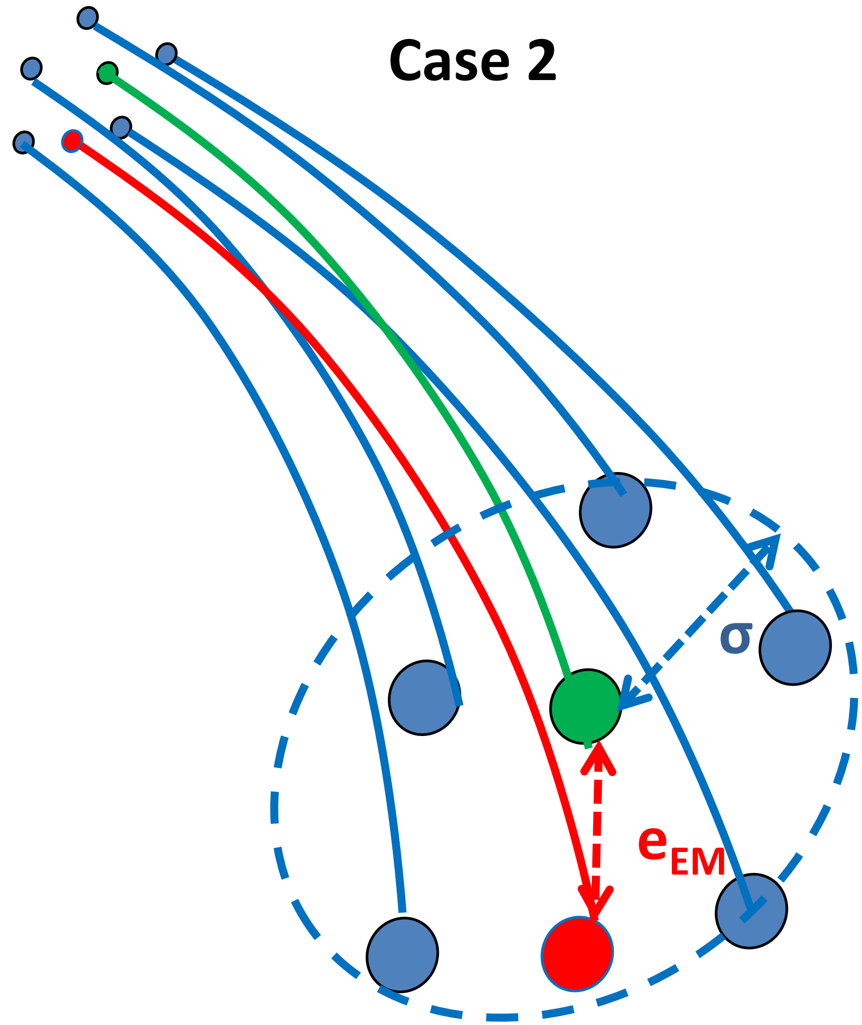
A reliable ensemble has, on average over many cases M, spread measured by the ensemble standard deviation σ, equal to the average error of the ensemble mean e_{EM}:
$\langle\sigma\rangle_M=\langle e_{EM}\rangle_M$

In a reliable ENS, small spread >> high predictability

Case 1



Case 2

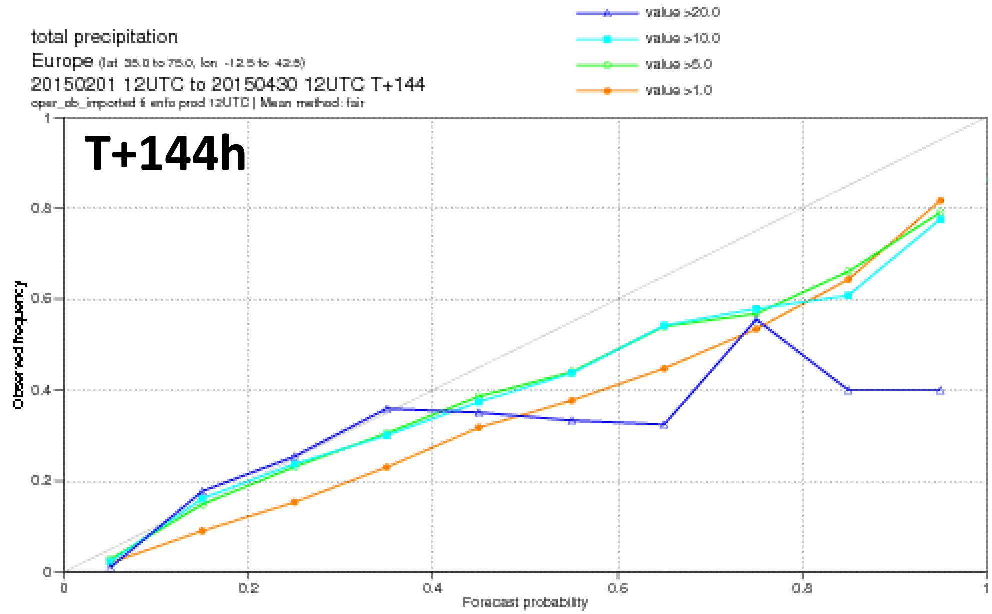


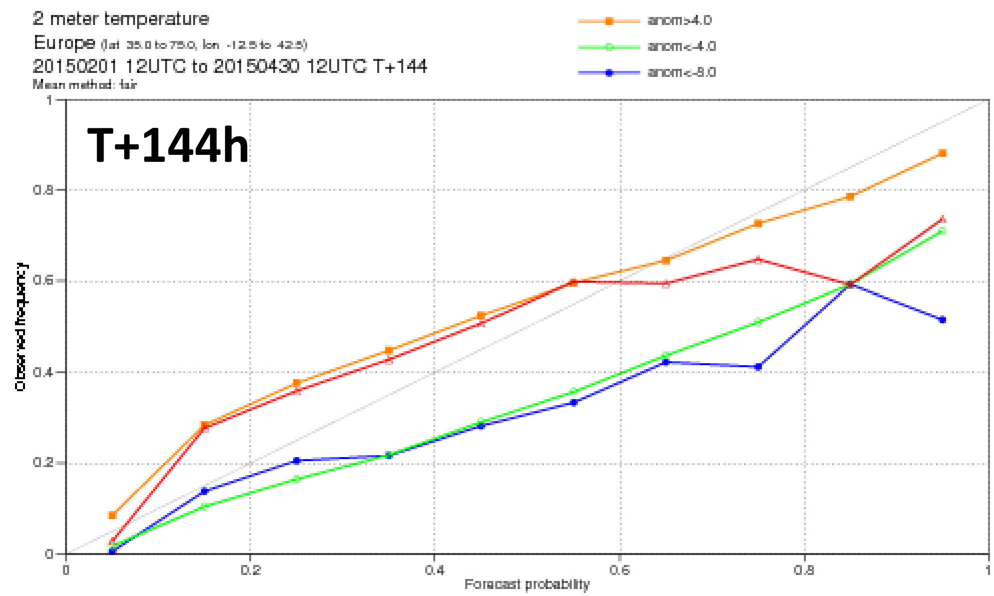
In a reliable ensemble, small ensemble standard deviation indicates a more predictable case, i.e. a small error of the ensemble mean e_{EM}.

Reliability: <fc-prob>~<obs-prob>

One way to check the ensemble reliability is to assess whether the average forecast and observed probabilities of a certain event are similar.

These plots compare the two probabilities at t+144h for the event '24h precipitation in excess of 1/5/10/20 mm' (top) and '2mT gt/lt 4/8 degrees' over Europe for FMA15 (verified against observations).

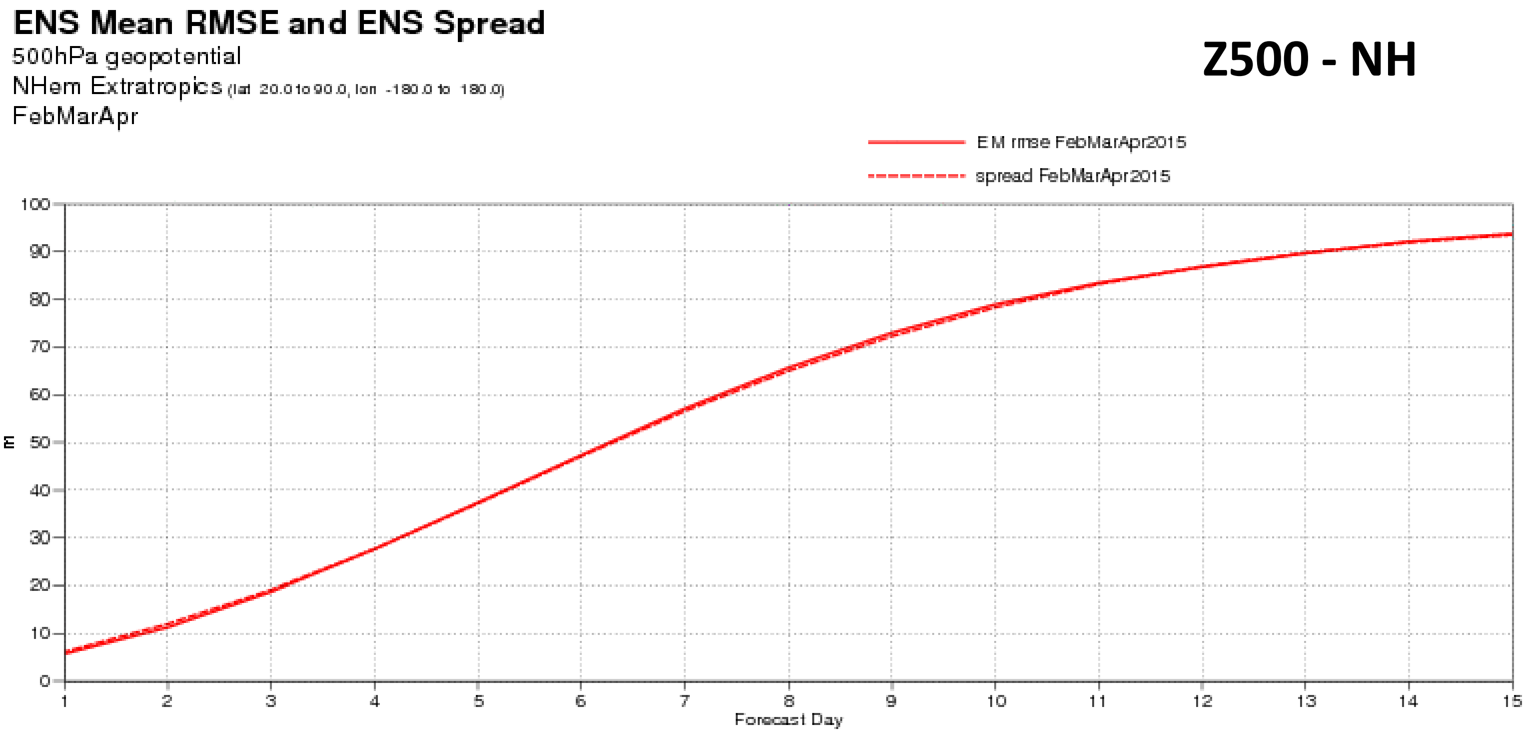


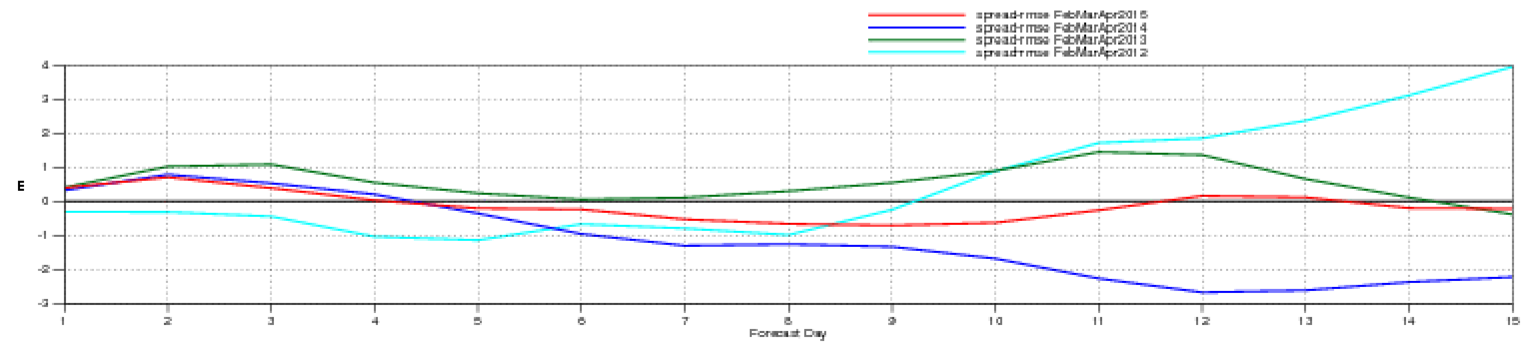


Reliability: <spread>~<rmse(EM)>

One way to check the ensemble reliability is to assess whether the seasonal average of the ensemble standard deviation and of the ensemble-mean error are similar.

Z500 - NH

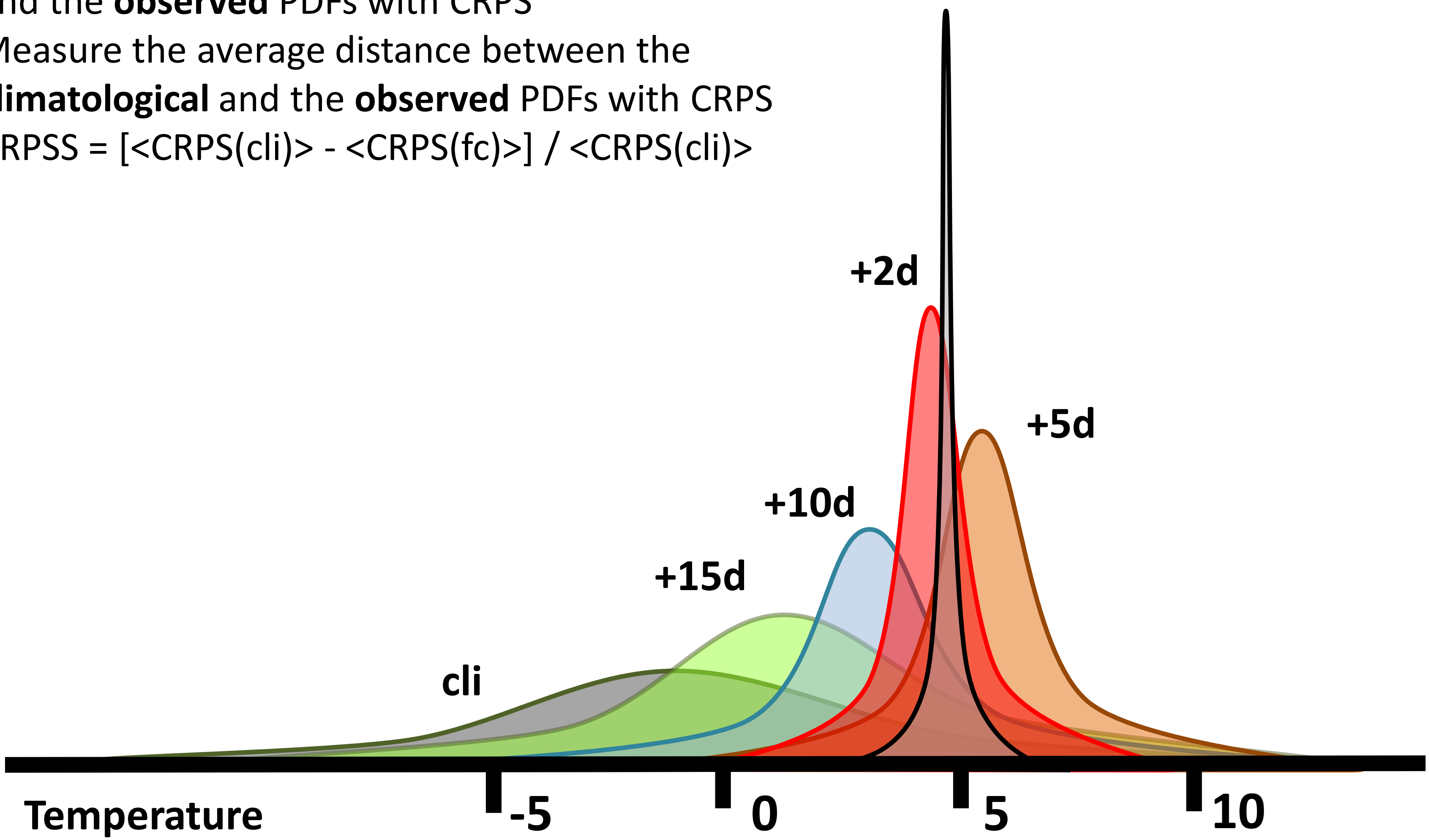




This plot compares these statistics for Z500 over NH in FMA15.

PDF forecast skill: how do we measure it?

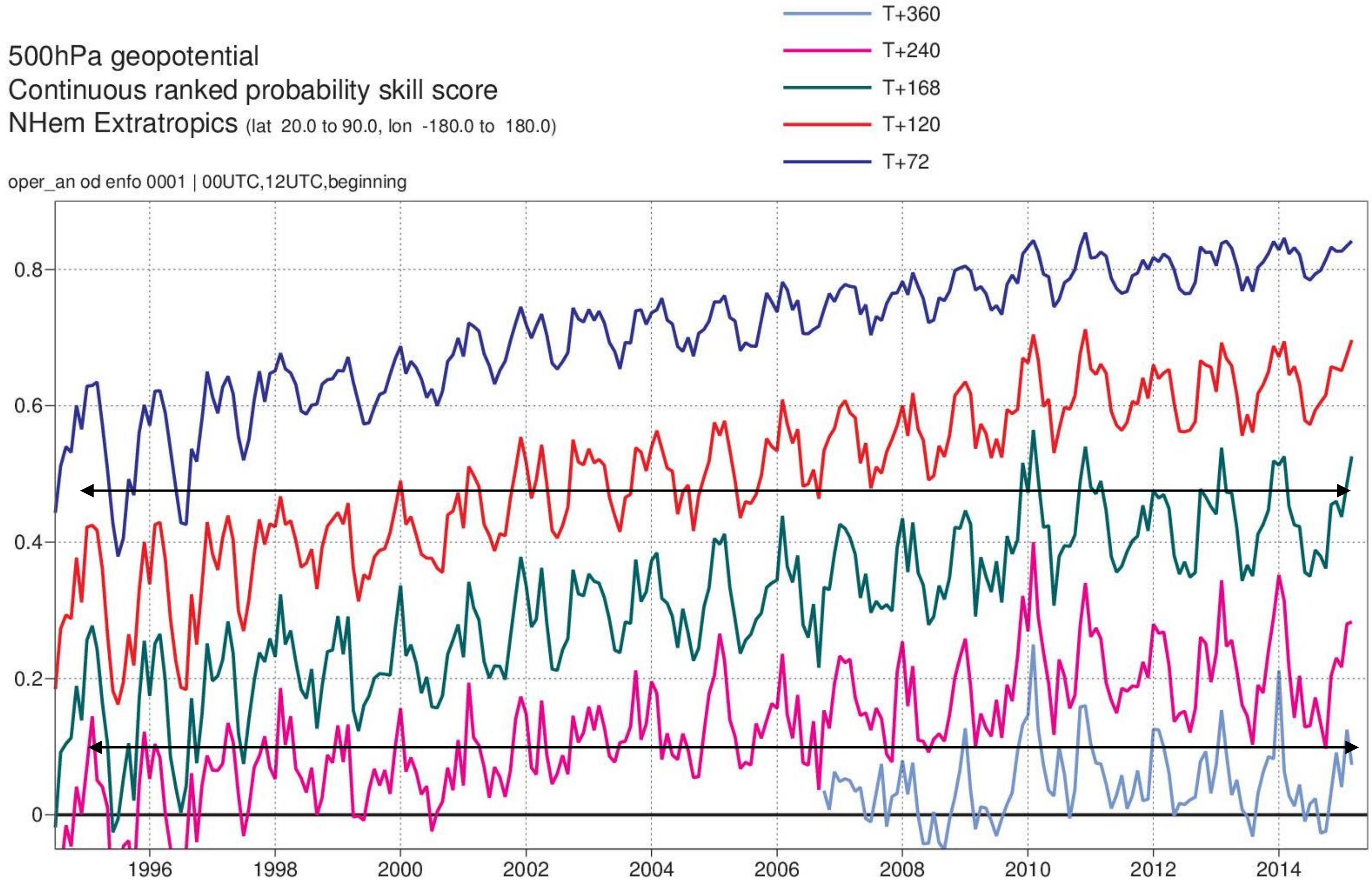
- Measure the average distance between the **forecast** and the **observed** PDFs with CRPS
- Measure the average distance between the **climatological** and the **observed** PDFs with CRPS
- CRPSS = [<CRPS(cli)> - <CRPS(fc)>] / <CRPS(cli)>



Accuracy: ENS CRPSS (Z500, NH)

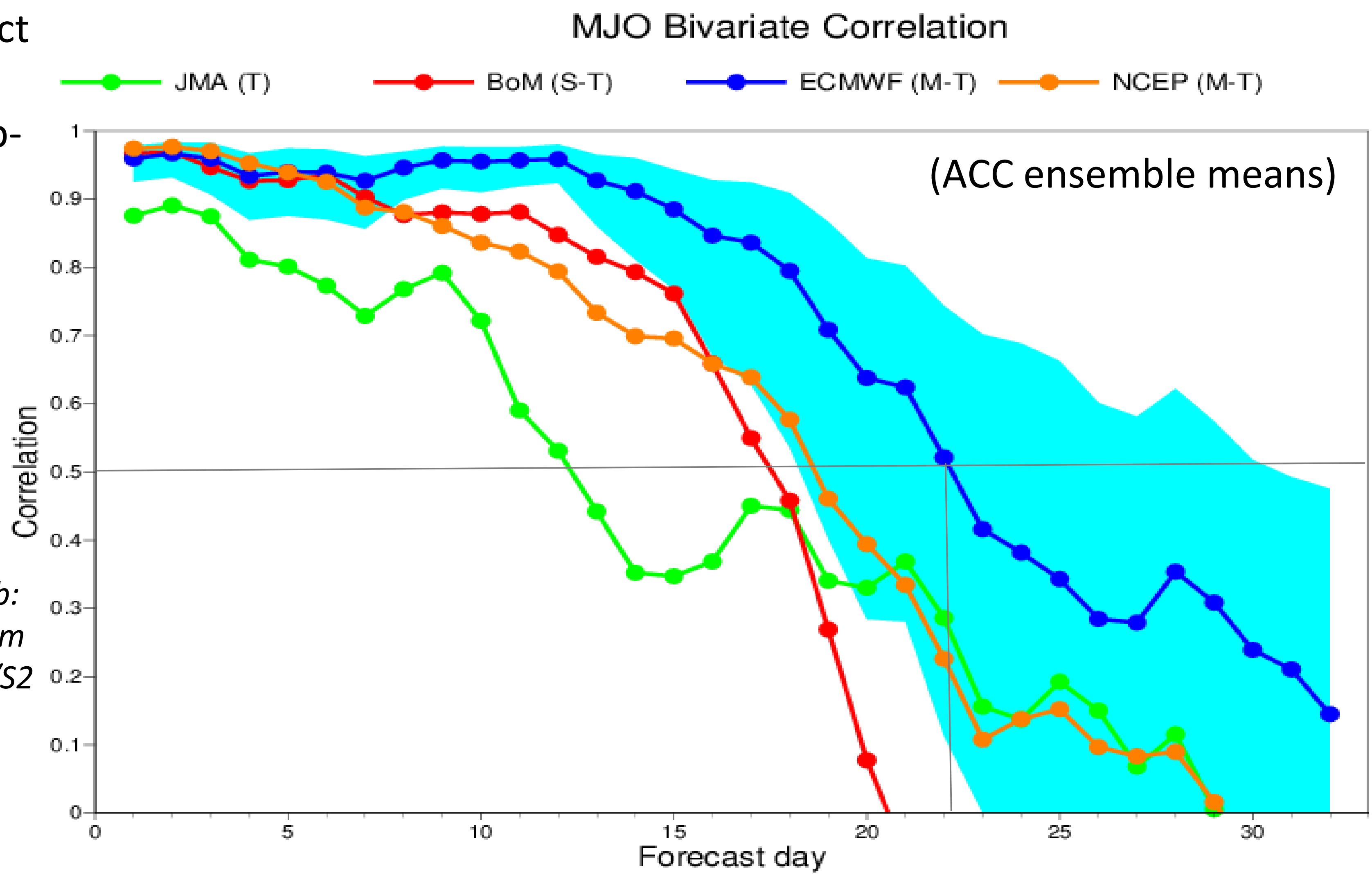
Accuracy of probabilistic forecasts can be measured using the Continuous Rank Probability Score and Skill Score.

CRPSS shows that today 15d fcs are as good as 10d fcs 20 years ago.



ENS has skill in the monthly forecast range

The **S2S** (WWRP & WCRP) project is helping us to understand sub-seasonal to seasonal predictability.



(S2S @ ECMWF web: https://software.ecmwf.int/wiki/display/S2S/Home)

(from Frederic Vitart)

ENS has skill in the monthly forecast range

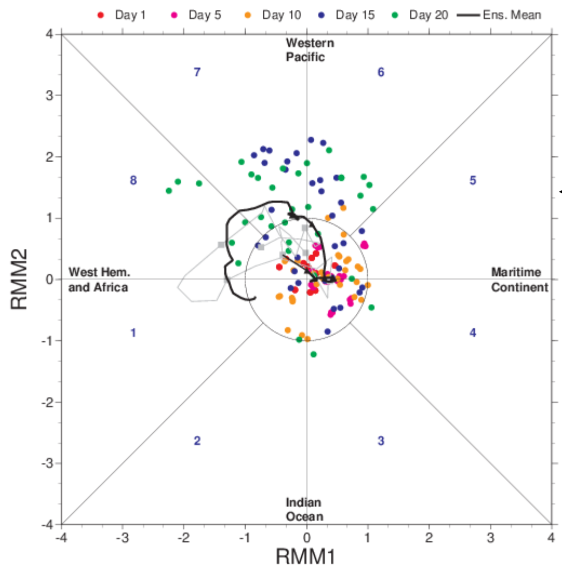
The **S2S** (WWRP & WCRP) project is helping us to understand sub-seasonal to seasonal predictability.

(S2S @ ECMWF web: https://software.ecmwf.int/wiki/display/S2S/Home)

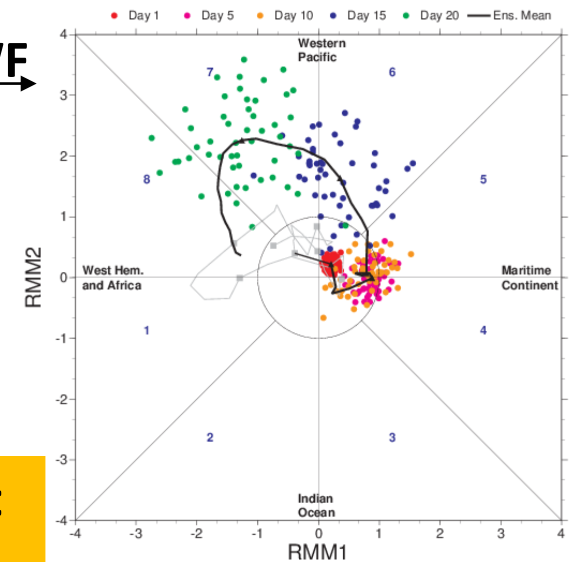
(from Frederic Vitart)

MJO forecast 26/02/2015 S2S Database

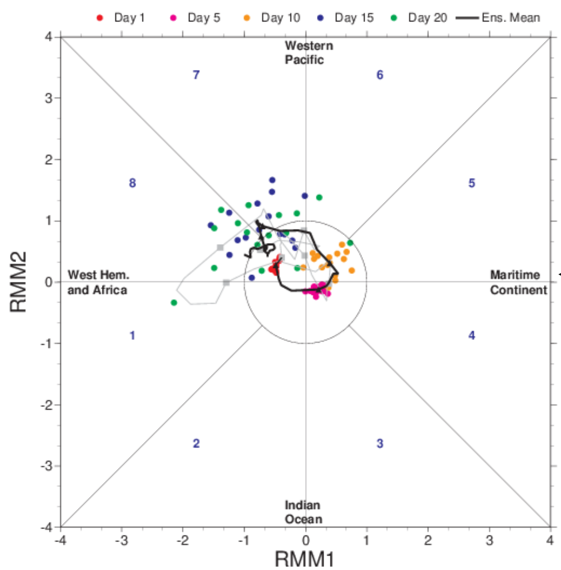
CAWCR



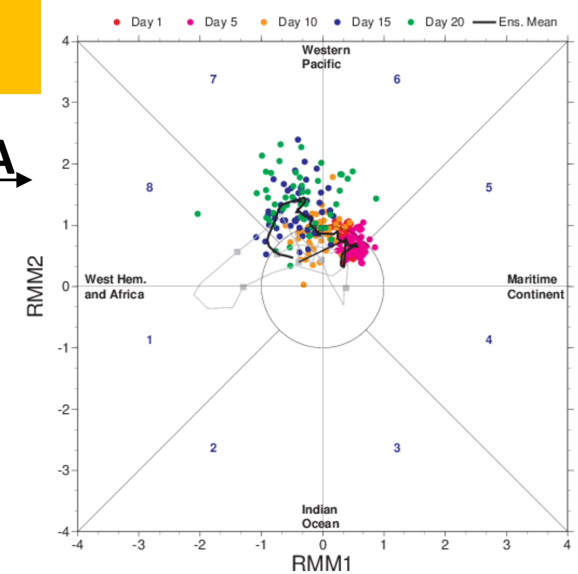
ECMWF



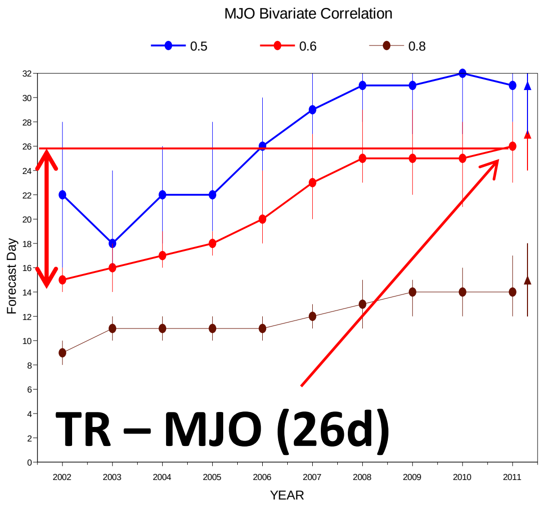
NCEP

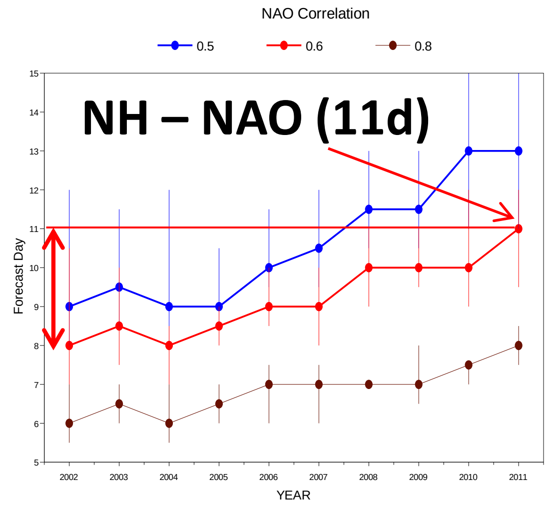


JMA



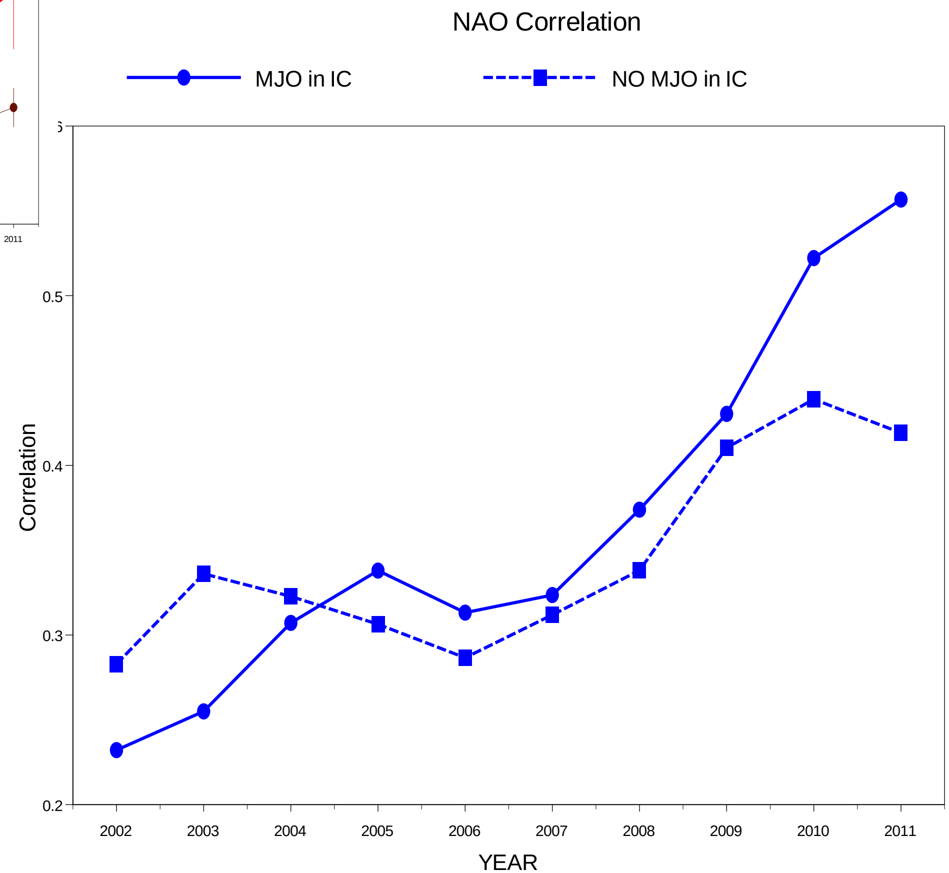
The MJO is important because it has impact on NAO





The skill of monthly forecasts have been continuously improving both in the tropics for the MJO (top left) and the extra-tropics for the NAO (top right).

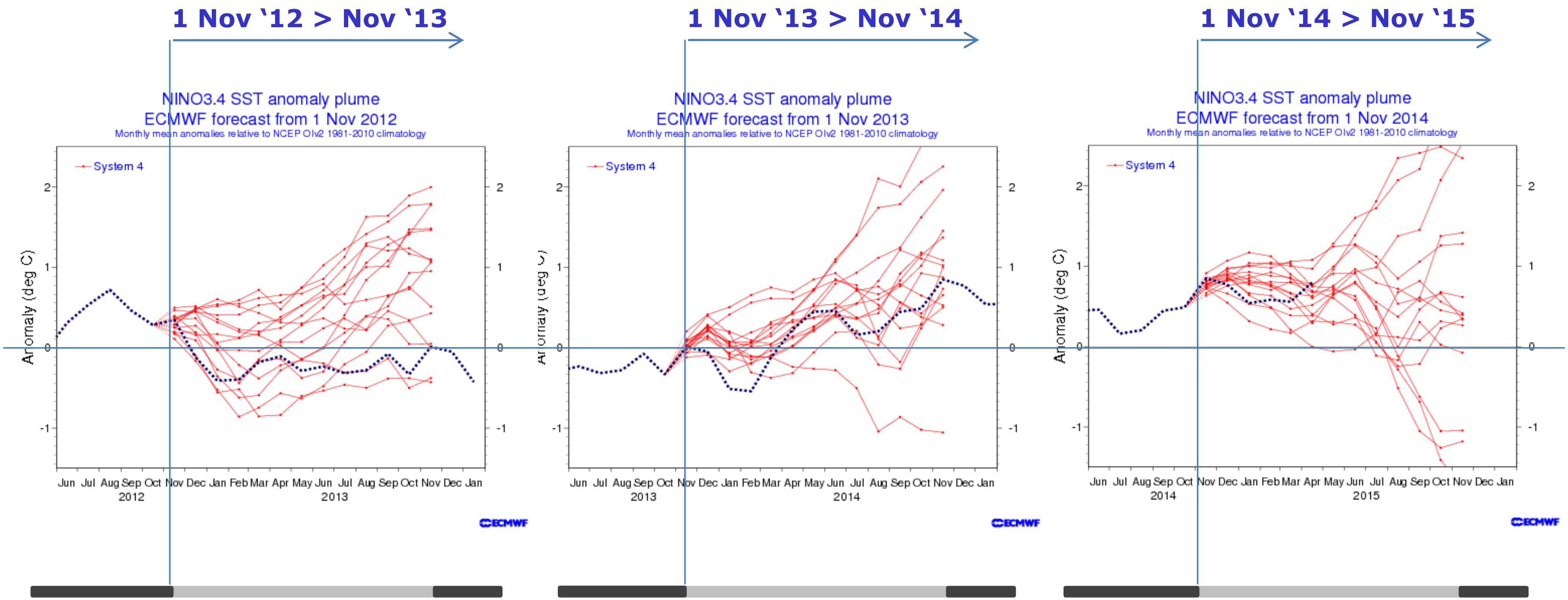
Improvements in the physics have led to better teleconnection between tropics and extra-tropics (right).



(from Frederic Vitart)

The seasonal ensemble S4 provides probabilities up to 1y

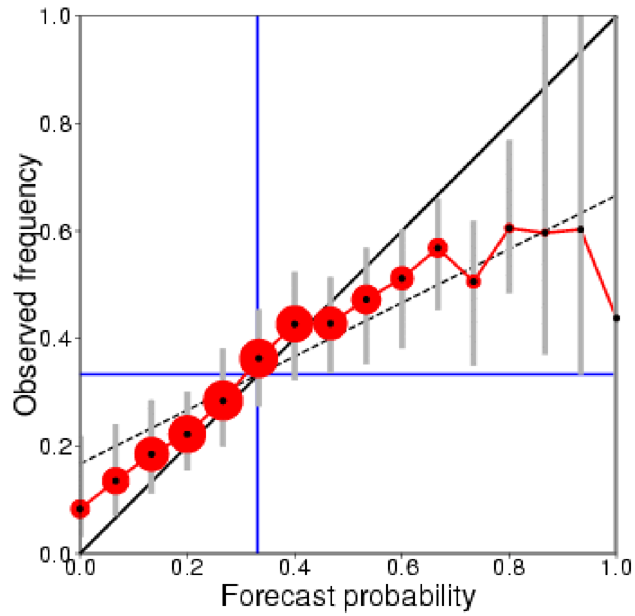
The tropics remain the area where seasonal prediction has the highest skill, as indicated e.g. by the accuracy of 1-year forecasts of SST anomaly in the Nino3.4 area.



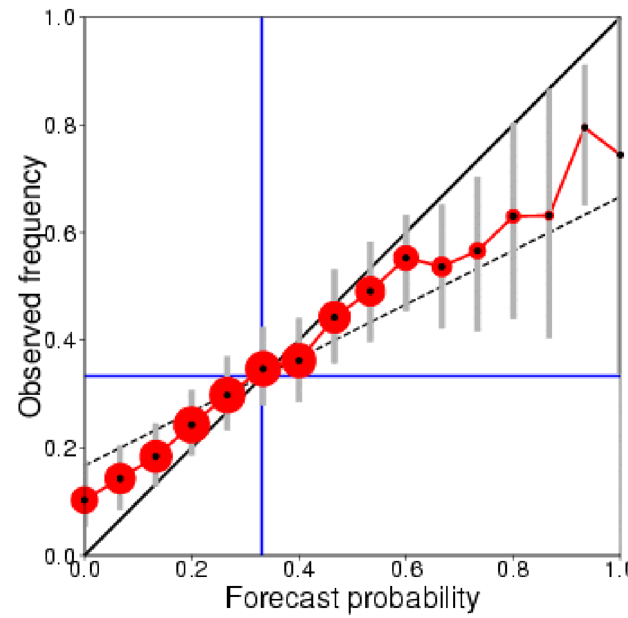
How does S4 perform in terms of reliability? 2mT

On average (30 years), 4-6 month probabilistic predictions of 2mT over North America and Europe started in Feb for MJJ (t+4-6m) are reliable and skilful compared to climatology (BSS>0).

EU – 1 Feb + 4-6m



NA – 1 Feb + 4-6m

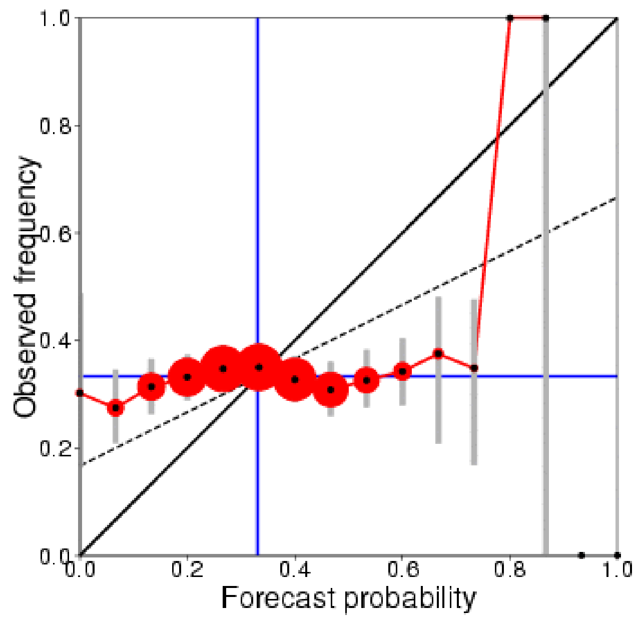


BSS PR(2mT>U3)	EU	NA
1 Feb > MJJ (t+ 4-6m)	0.064	0.050
1 Apr > MJJ (t+2-4m)	0.058	0.074

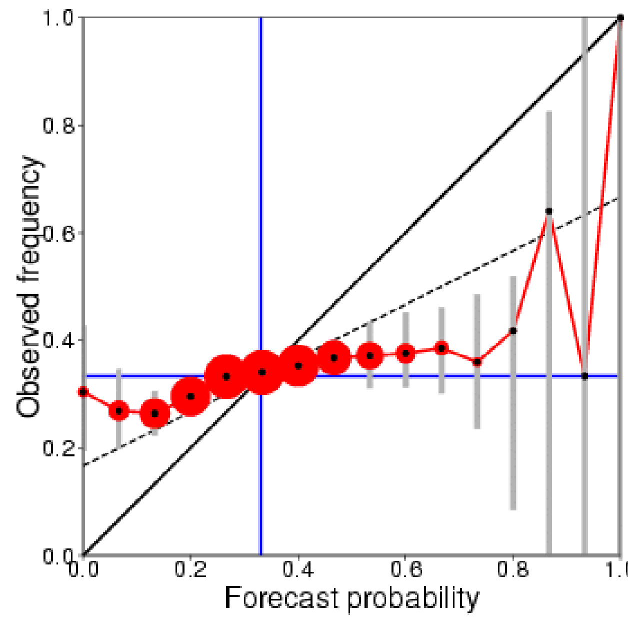
… but for precipitation, reliability is very poor!

On average (30 years), even shorter-range 2-4 month probabilistic predictions of TP over North America and Europe started in Apr for MJJ (t+2-4m) are **not** reliable and **less** skilful than climatology.

EU – 1 Apr + 2-4m

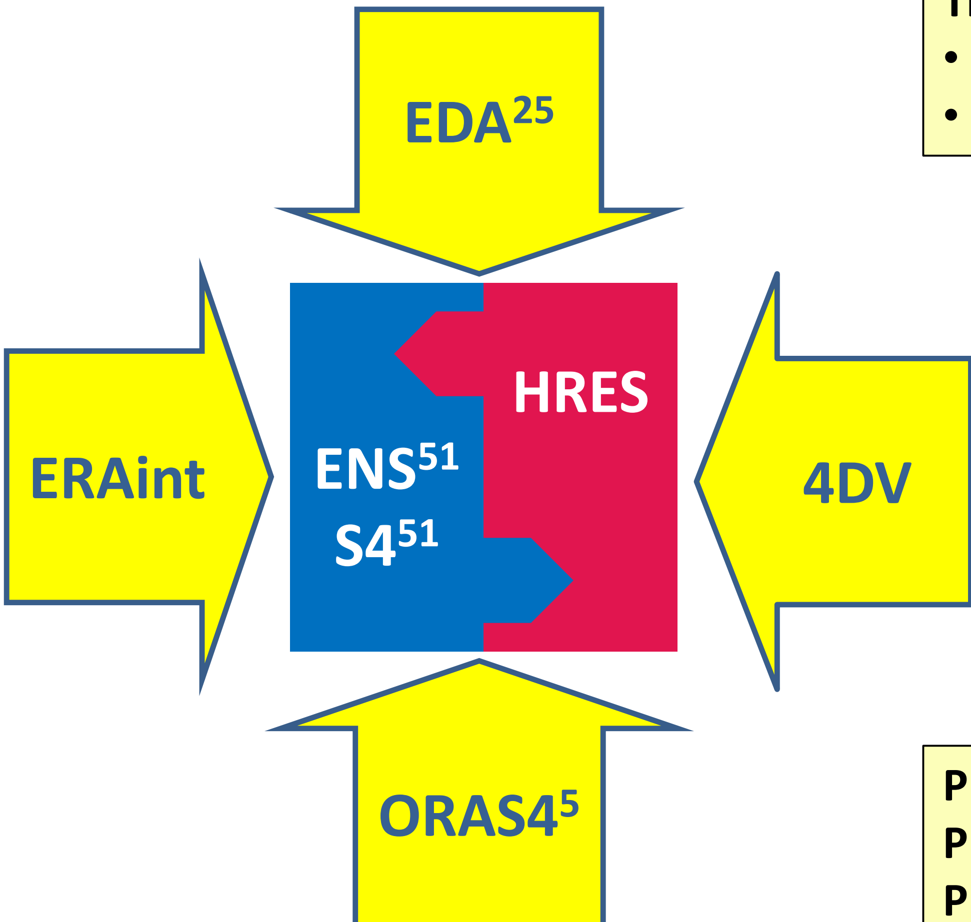


NA – 1 Apr + 2-4m



BSS PR(TP<L3)	EU	NA
1 Feb > MJJ (t+ 4-6m)	-0.049	-0.052
1 Apr > MJJ (t+2-4m)	-0.072	-0.052

Ensembles: the ECMWF way



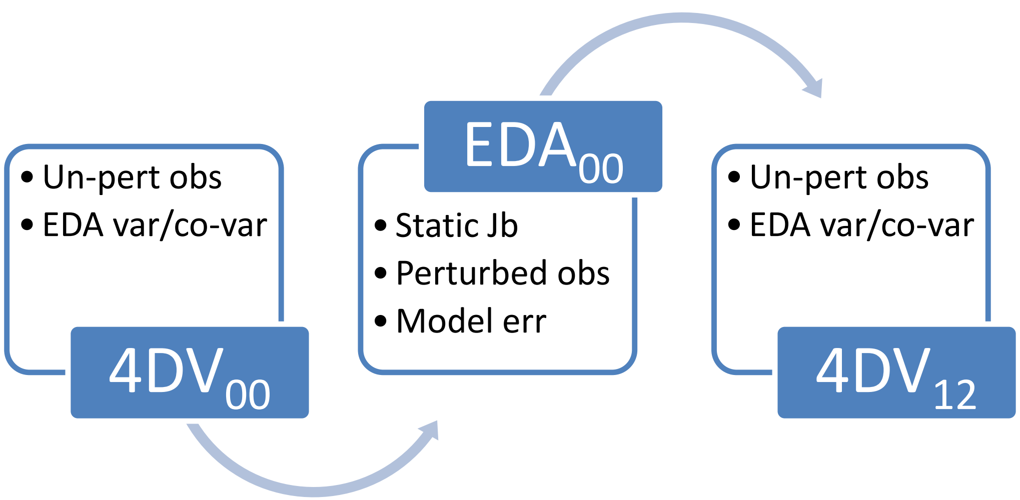
They simulate the effect of:

- **Observation/initial uncertainties**
- **Model uncertainties (2 stochastic schemes)**

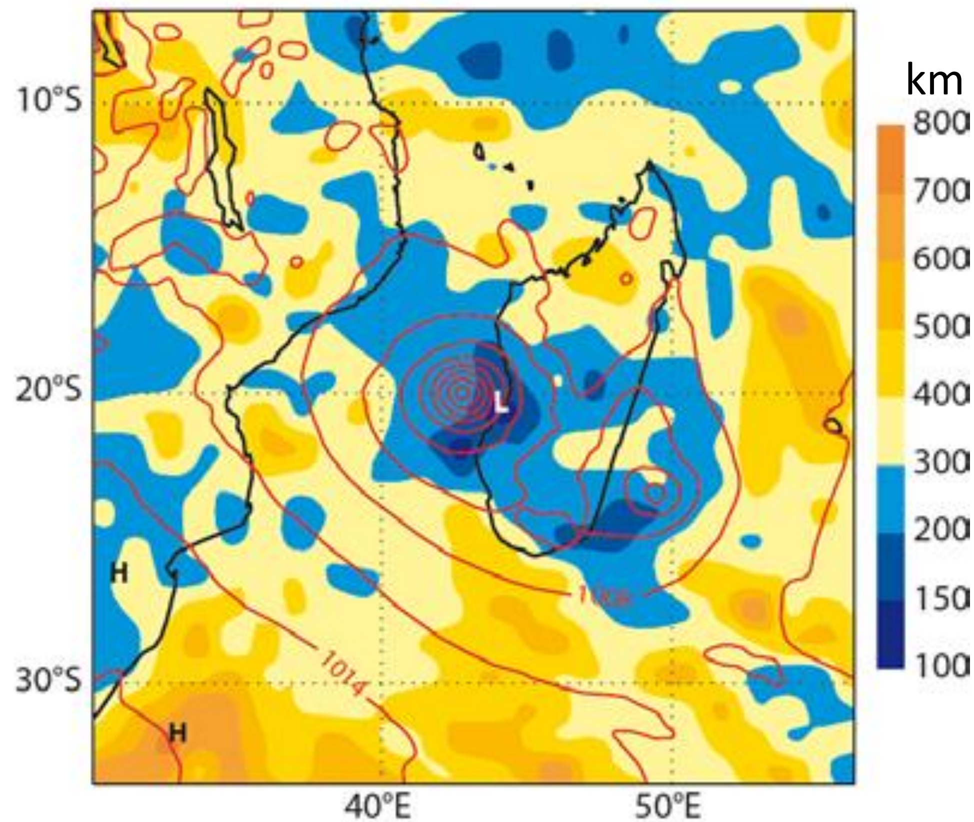
PDF(0) << 4DV+EDA^{25}+ORAS4^{5}
PDF(0) << ERA+ORAS4^{5} (the past)
PDF(T) << HRES+ENS^{51}/S4^{51}

The EDA is used to estimate flow dependent stats

The 25-member **Ensemble of Data Assimilations** provides the 4DV-HRES with flow dependent background error statistics.



Background error correlation length scale for long(p_{msl}) and p_{msl}

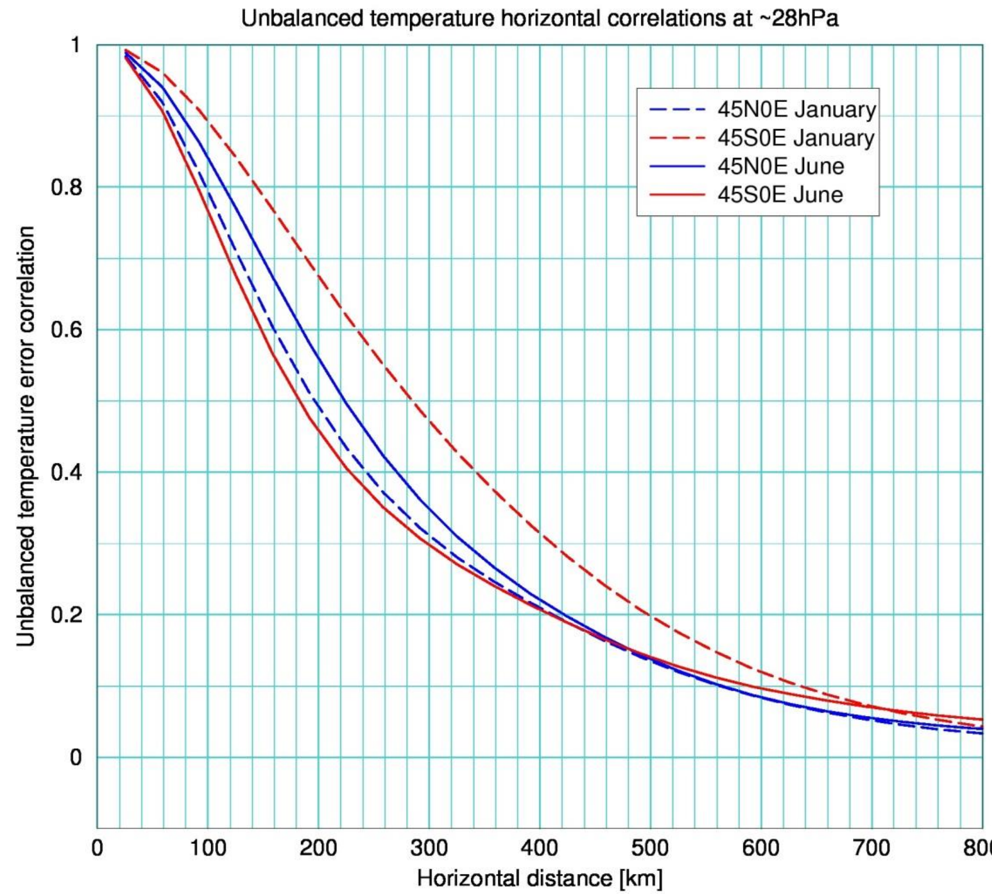


(from Massimo Bonavita)

EDA flow-dependent stats are key to assimilate obs

This figure shows the differences between EDA-based temperature correlations at ~28 hPa for two points, (45°N;0°E) (blue) and (45°S;0°E) (red) for a day in January (dashed) and June (solid).

The plot shows, e.g., that the SH winter (red solid) and summer (red dashed) temperature correlations are significantly different, a feature that cannot be accurately presented by climatological correlations.



(from Massimo Bonavita)

How can we keep improving the ensembles?

Work is progressing on many areas to further improve the ECMWF ensembles:

1. **Modelling** (including model uncertainty simulation): improve all model components (land, atmosphere and ocean) and increase resolution; upgrade the stochastic schemes that simulate model uncertainty

2. **Initial Conditions** estimation: integrate further the analysis and forecast ensembles (EDA and ENS) and re-assess the potential benefit of starting ENS directly from EDA analyses; assess the impact of using a more strongly coupled DA

3. **Predictability:** identify sources of predictability, and ways to extract predictable signals for the ensemble PDF

4. **Ensemble methods**: assess whether different ensemble configurations (IC/model unc, membership, truncation, refc suite, …) could lead to more accurate and reliable PDF fcs

Forthcoming resolution increase: all IFS components

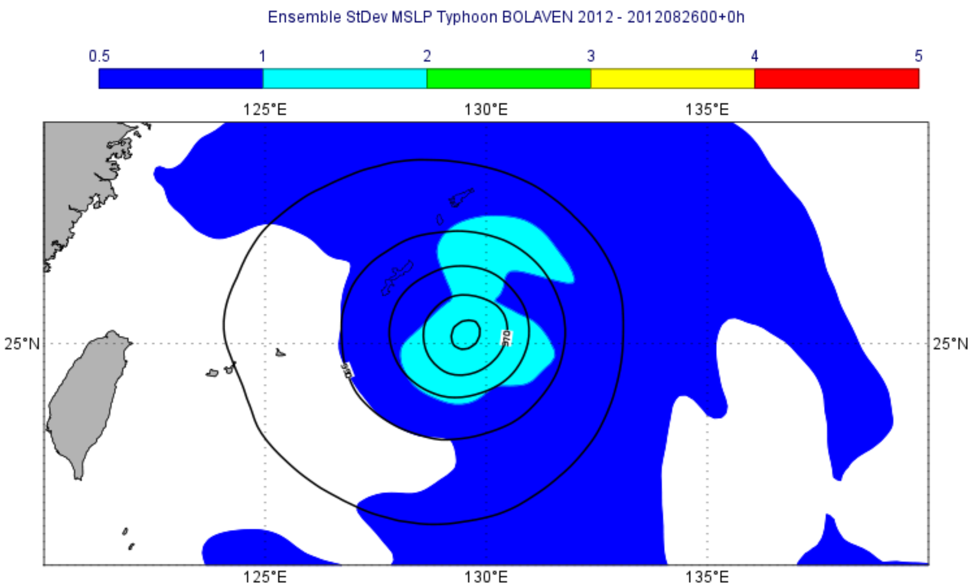
	Operational suite	Future resolution (preferred option)
HRES	T_L1279 (~16km) / L137	T_{CO}1279 (~9 km) / L137
4DVAR	T_L1279+3xT_L255 / L137	T_{CO}1279+T_{CO}255/319/399 / L137
EDA	T_L399 / L137 (25 members)	T_{CO}639 / L137 (25 members)
ENS/MOFC IFS:	51 members TL639 (~ 32 km) (0 - 10d) +T_L319 (10 - 46d) / L91	51 members T_{CO}639 (~ 18km) (0 - 10d) +T_{CO}319 (10 – 46d) / L91
OCEAN:	NEMO ORCA100z42	NEMO ORCA025z75

T_{CO} – Cubic octahedral Gaussian reduced grids

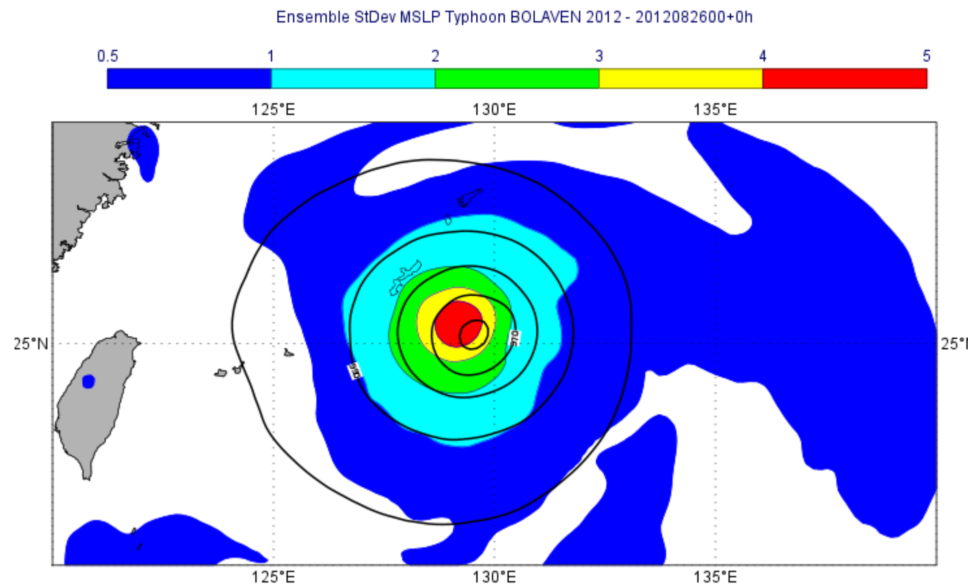
Predictability: impact of using a higher-resolution EDA

An increase in the resolution of the EDA outer loops from T399 to T639 is expected to improve the estimation of the analysis uncertainty and thus the ENS initial conditions. This figure shows the impact on the ENS spread (std) in the case of typhoon Bolaven (26 Aug 2012).

Ensemble with perturbations from 399 EDA



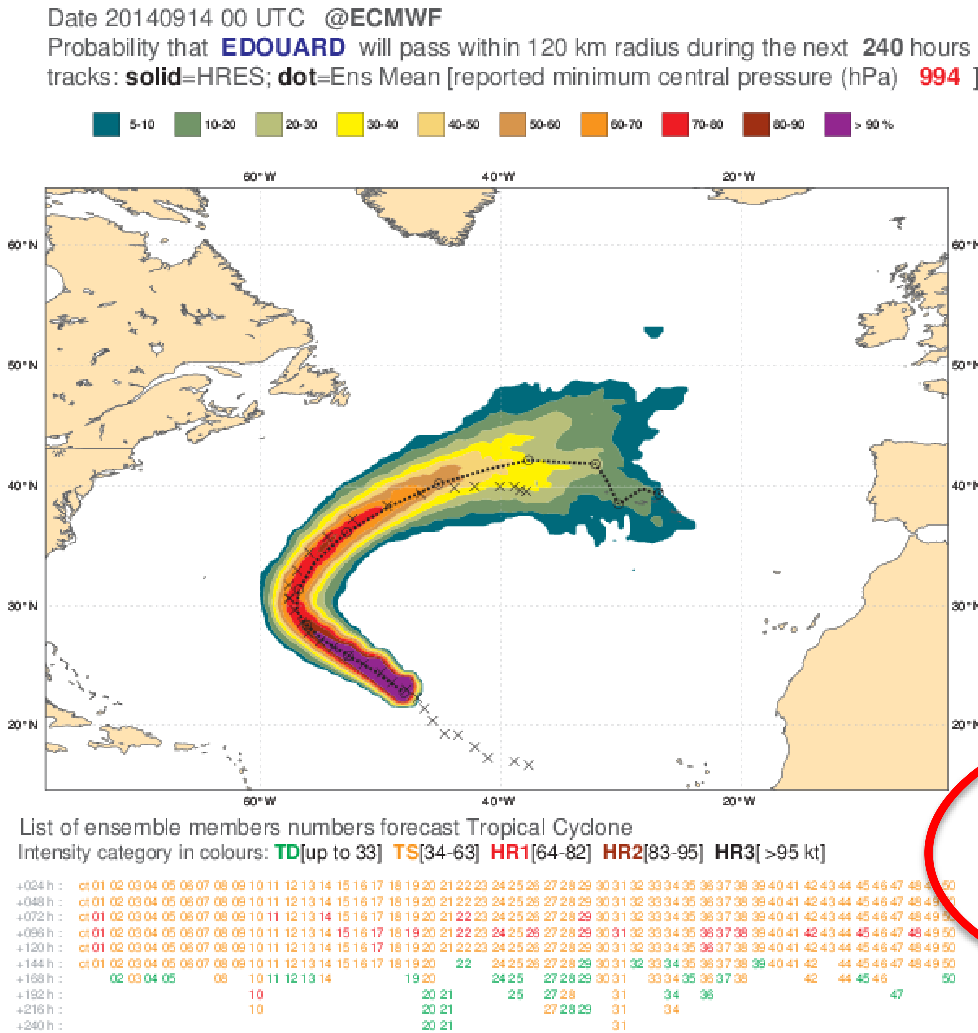
Ensemble with perturbations from 639 EDA

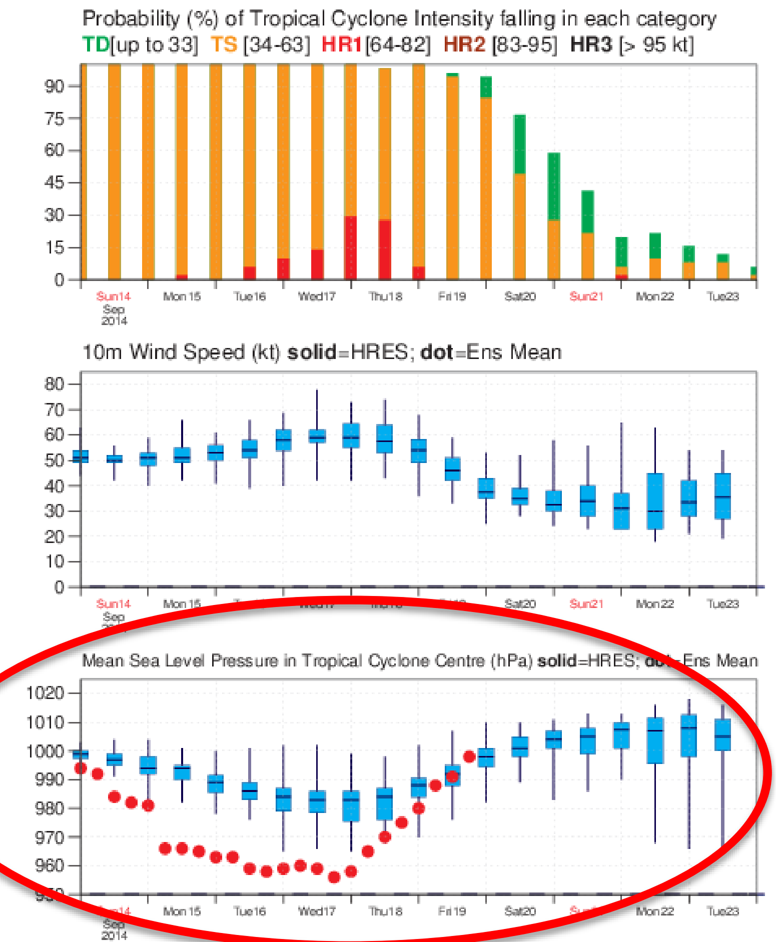


(From Simon Lang)

Very recent results: impact of increasing EDA & ENS resol?

OPE: T_L1279 An, T_L399 EDA, T_L639 ENS



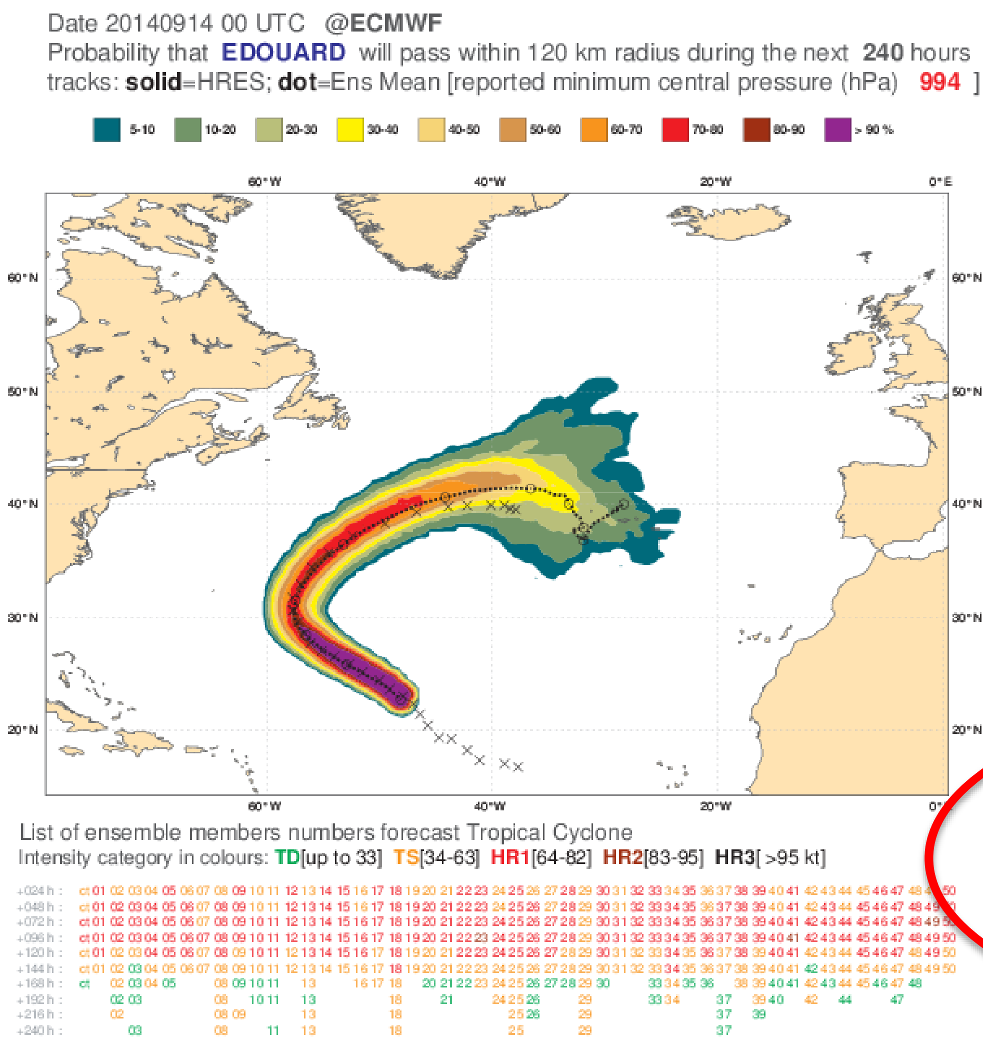


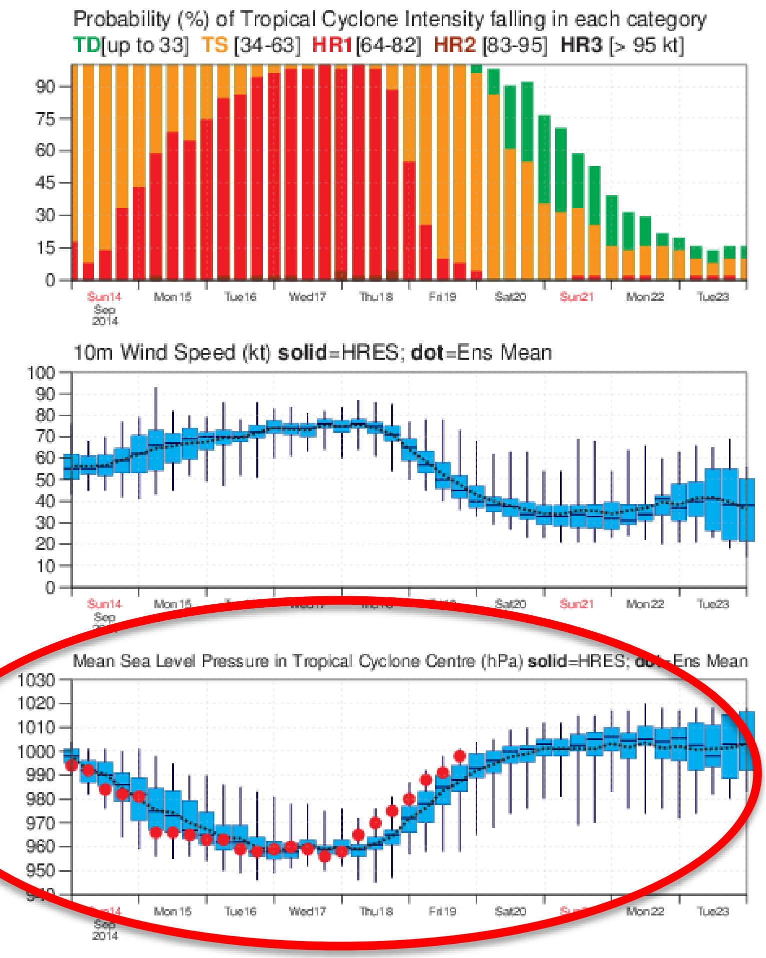
(From Simon Lang)

Too small initial spread and deepening not captured

Very recent results: impact of increasing EDA & ENS resol?

OPE: T_L1279 An, T_L639 EDA, T_{CO}639 ENS

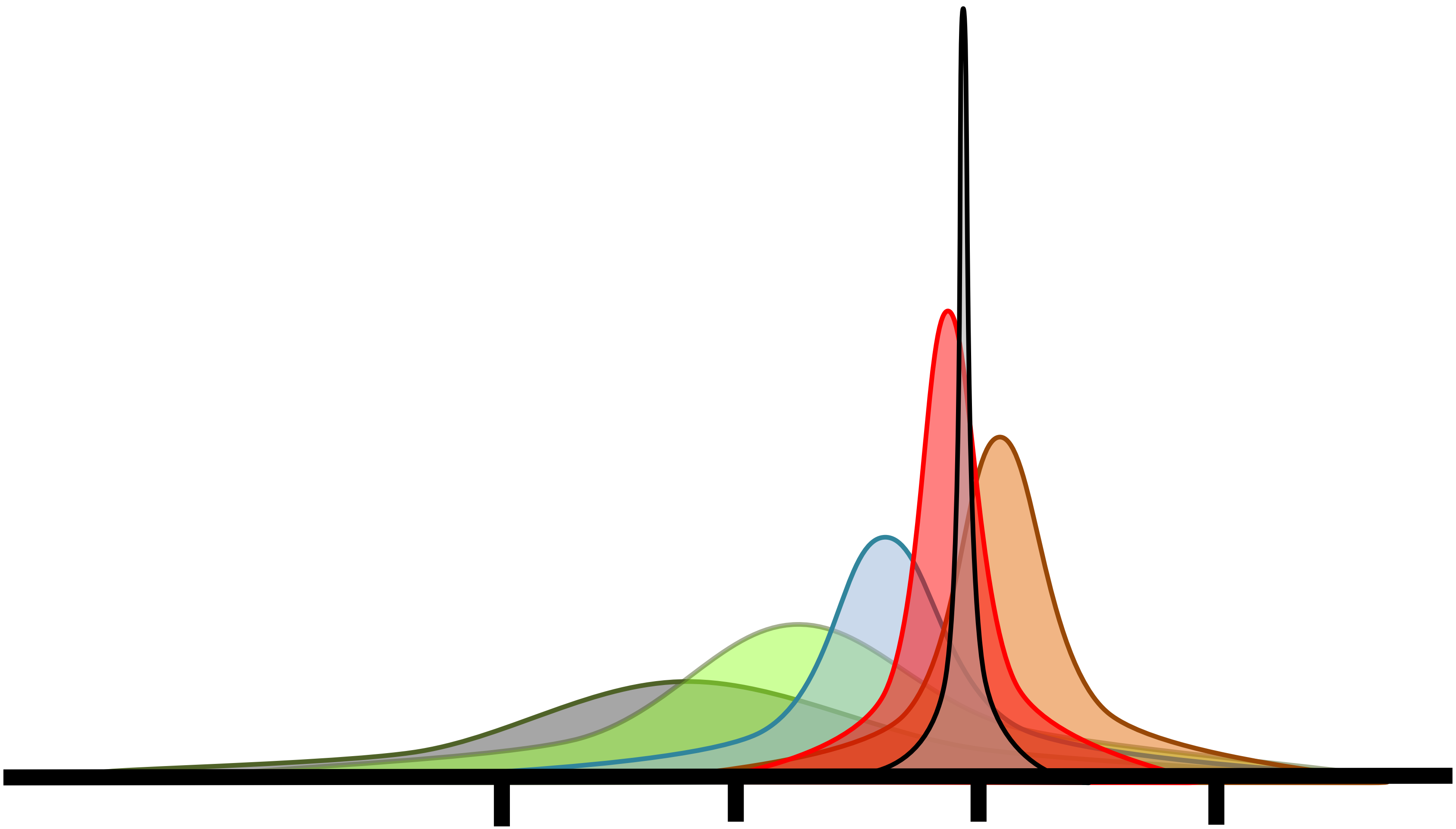




(From Simon Lang)

Deepening is captured and spread reflects uncertainty

.. to conclude .. think ensembles!!!



…. extra slides

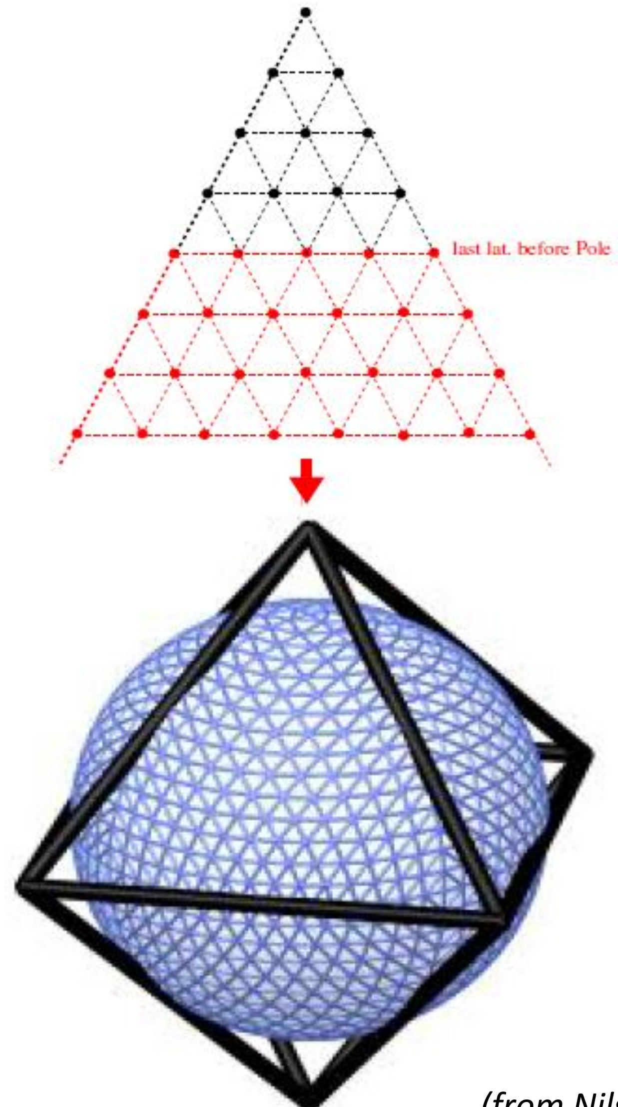
Forthcoming resolution upgrade: octahedral cubic grids

It is a reduced Gaussian grid with the same number of latitude circles ($NDGL$) than the standard Gaussian grid ($\leftrightarrow$ Gaussian weights) but with a new rule to compute the number of points per latitude circle.

Number of points per latitude

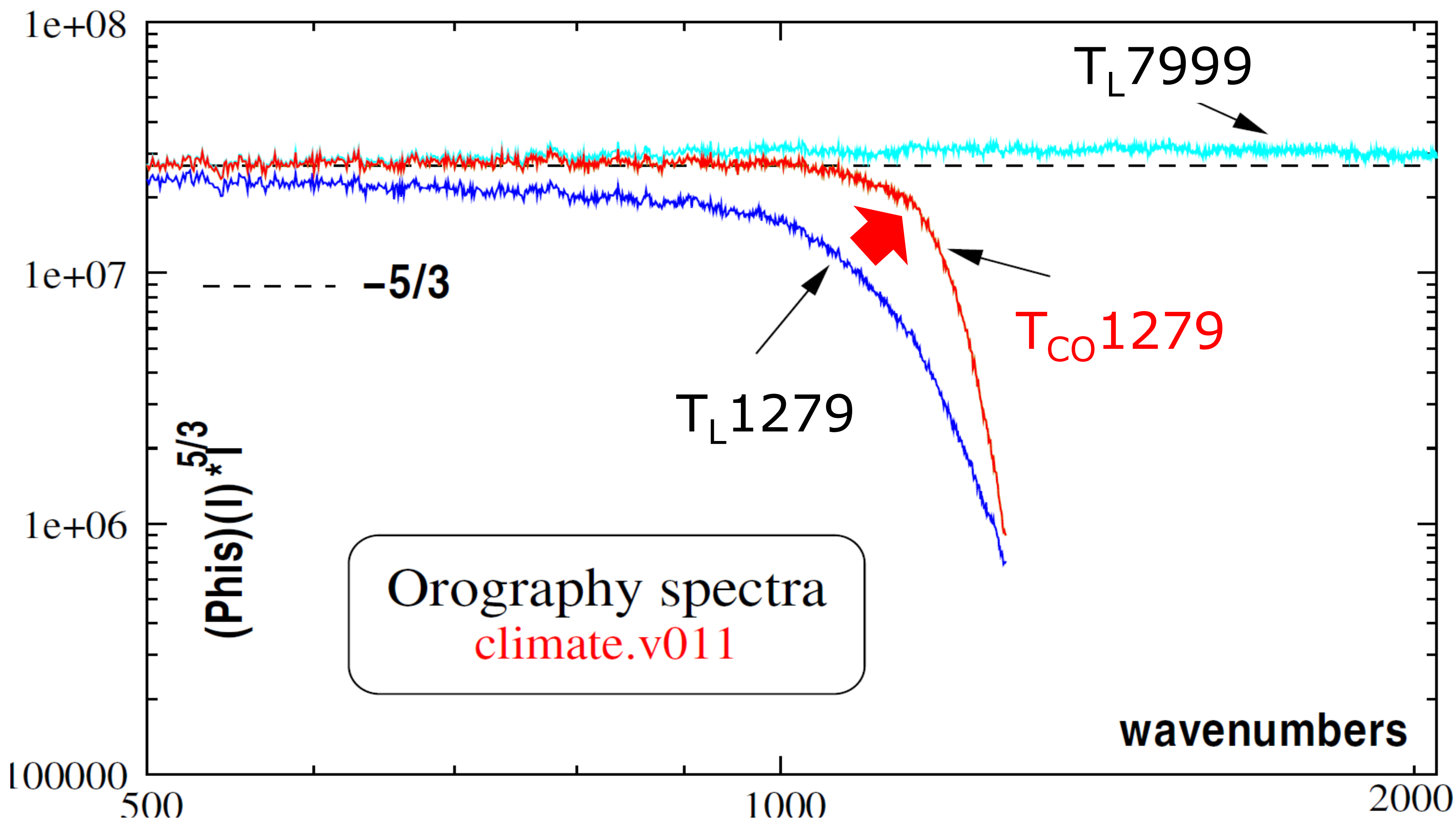
NLOEN(lat_N)=20 $\rightarrow$ Poles
NLOEN(lat_i)=NLOEN(lat_{i-1})+4

TL1279 :2.14 Mpoints
TC1023 :5.45 Mpoints
TC1279 :8.51 Mpoints
TCo1279 :6.59 Mpoints



(from Nils Wedi)

Forthcoming resolution increase: orographic variance



(from Nils Wedi)