Estimating uncertainty with the ECMWF ensembles

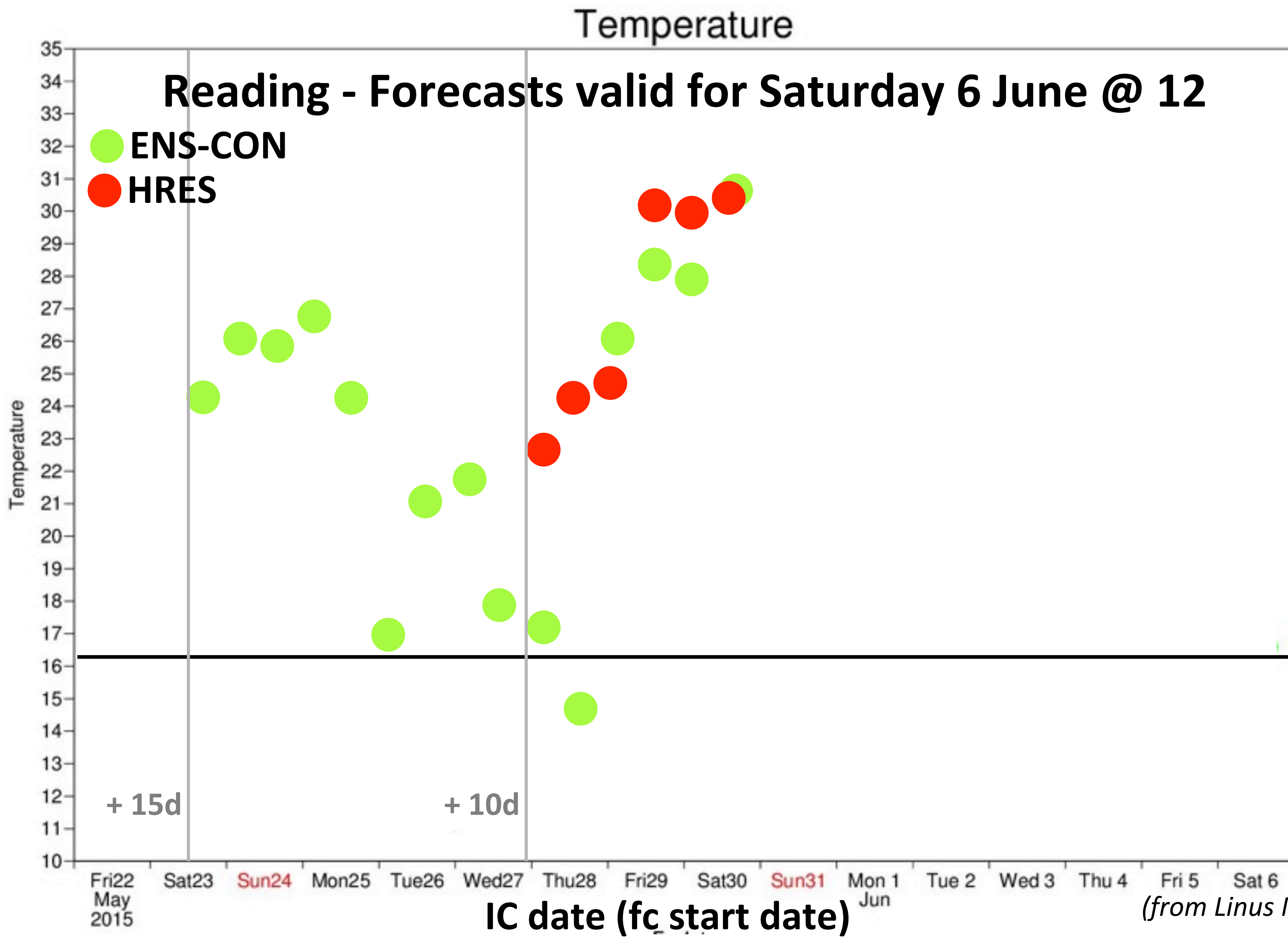
Roberto Buizza

European Centre for Medium-Range Weather Forecasts

Plan

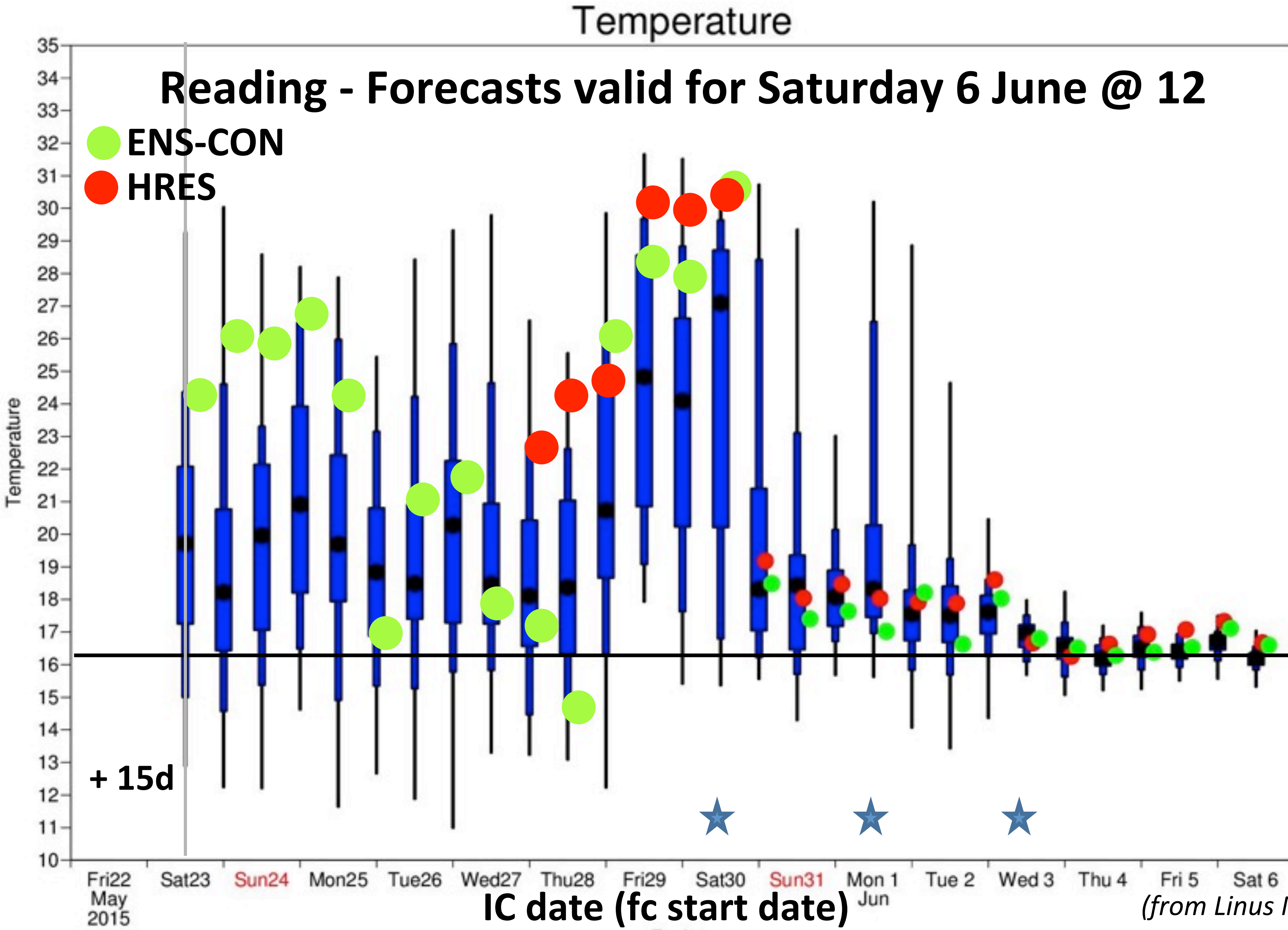
1. Why do we need ensembles?
2. The ECMWF ensembles
3. The resolution upgrade of March 2016
4. Few key performance indices
5. Prediction of an extreme precipitation event in Houston
6. Conclusions

1. Why do we need ensembles?



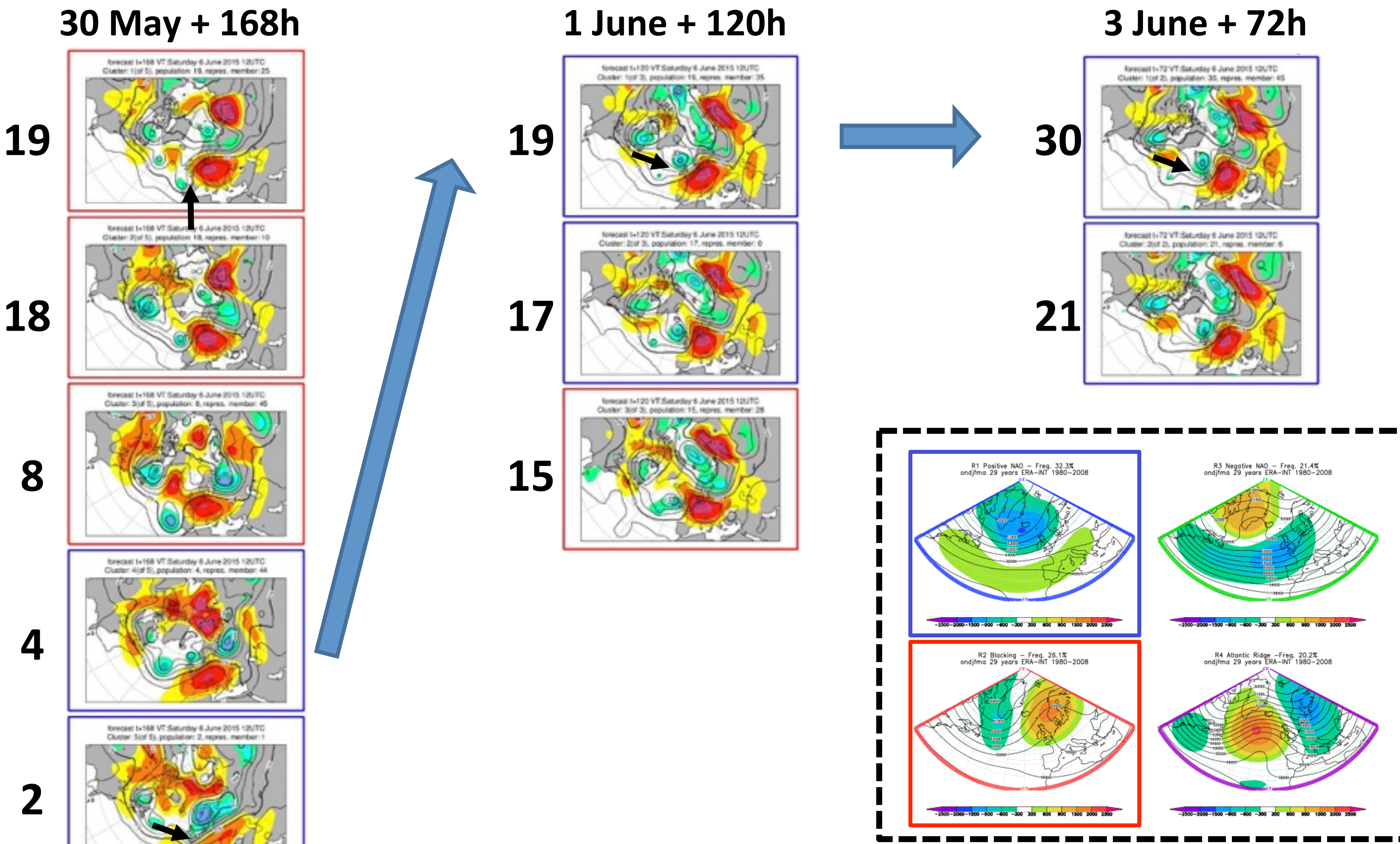
(from Linus Magnusson)

1. Why do we need ensembles?



(from Linus Magnusson)

1. Why do we need ensembles?

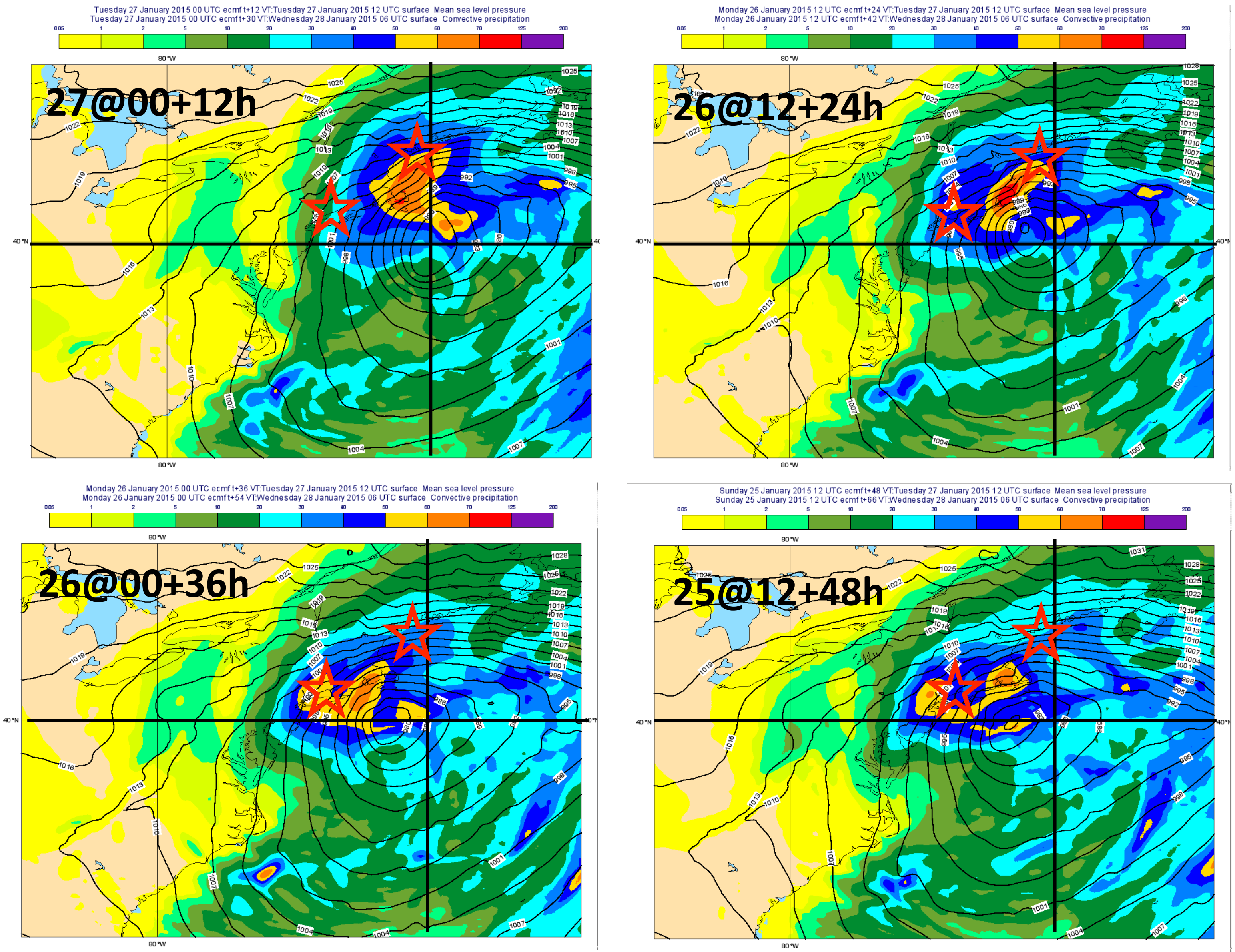


(from Linus Magnusson and Laura Ferranti)

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

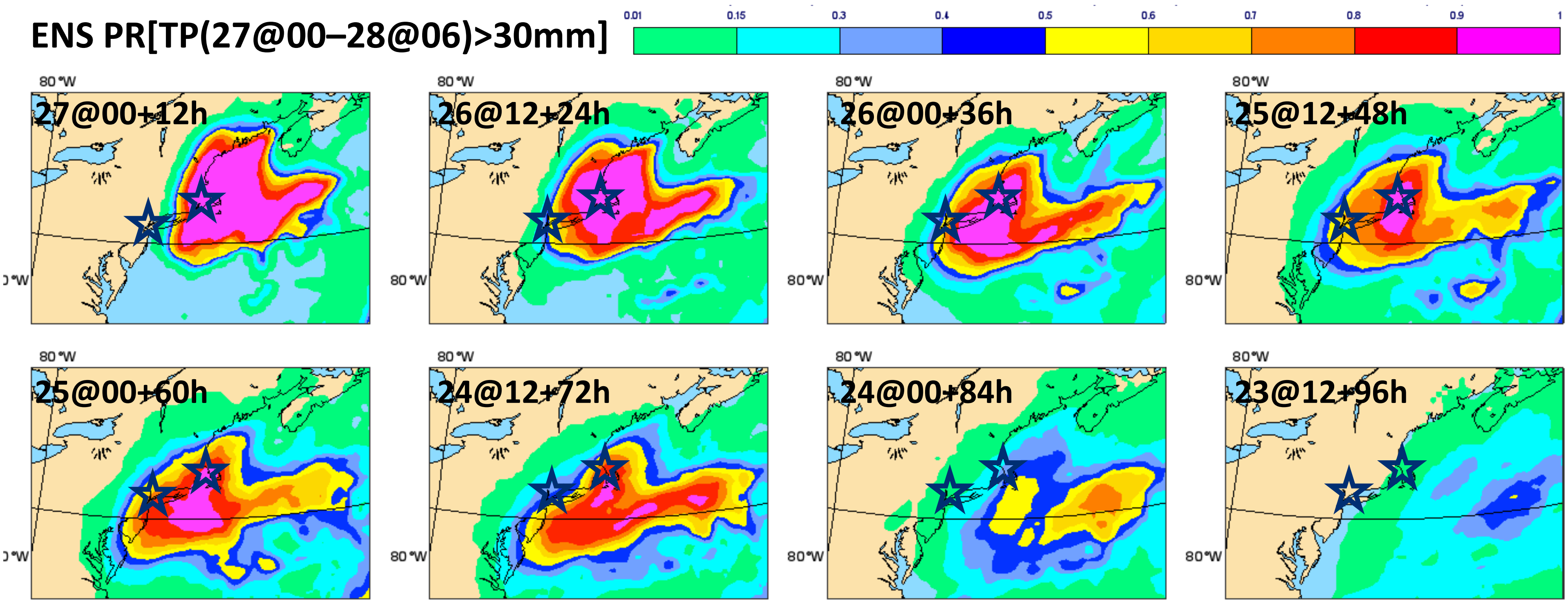
Single HRES fcs failed to positioned correctly the storm, and this led to snowfall overestimation for NY of in the 24-36-48h forecasts.

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



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

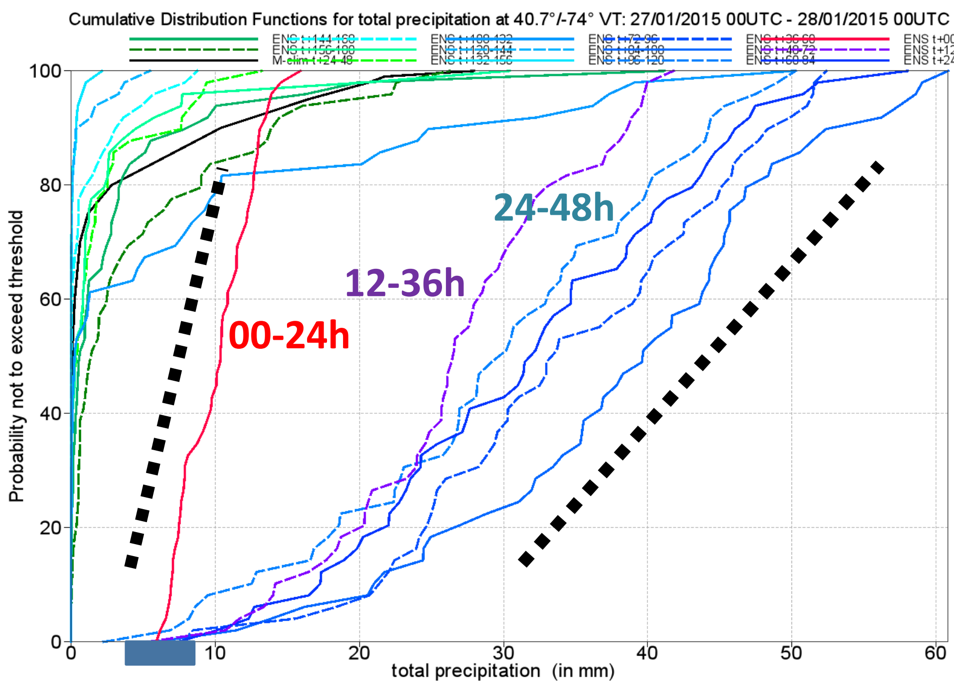
ENS-based probabilistic forecasts can be used to estimate the level of confidence (predictability) of single forecasts. They show that NY was closer to the edge of the area of high probability of +30mm of precipitation, indicating higher uncertainty.



1. US Storm, 27-28/01/2015: forecast confidence

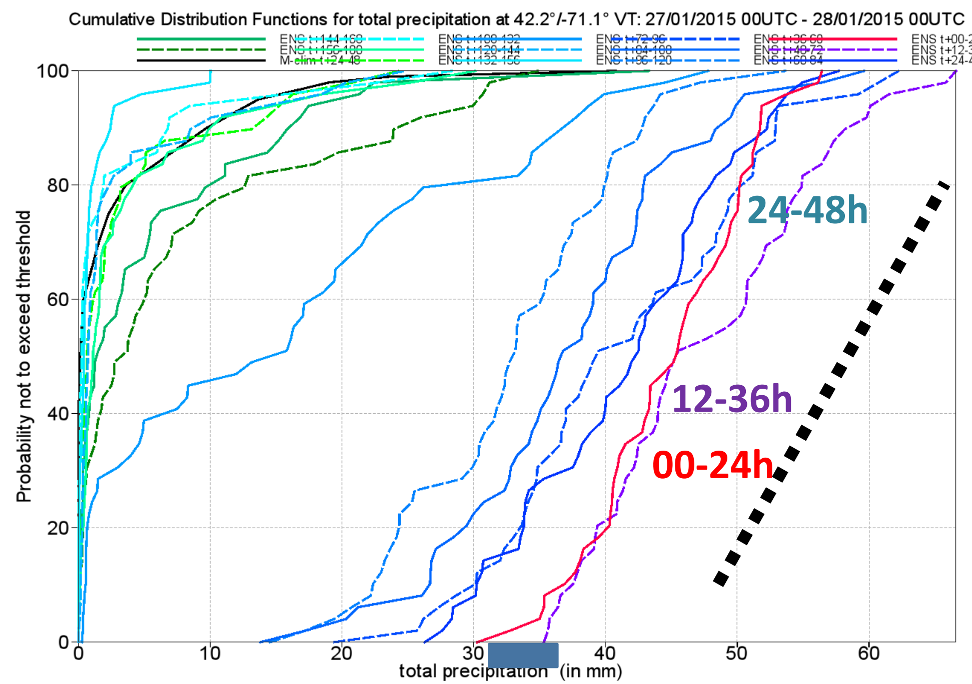
ENS-based probabilistic forecasts expressed in terms of CDF shows that the fcs for NY were more uncertain (the slope of the CDF curves is steeper) than the fcs for Boston.

New York



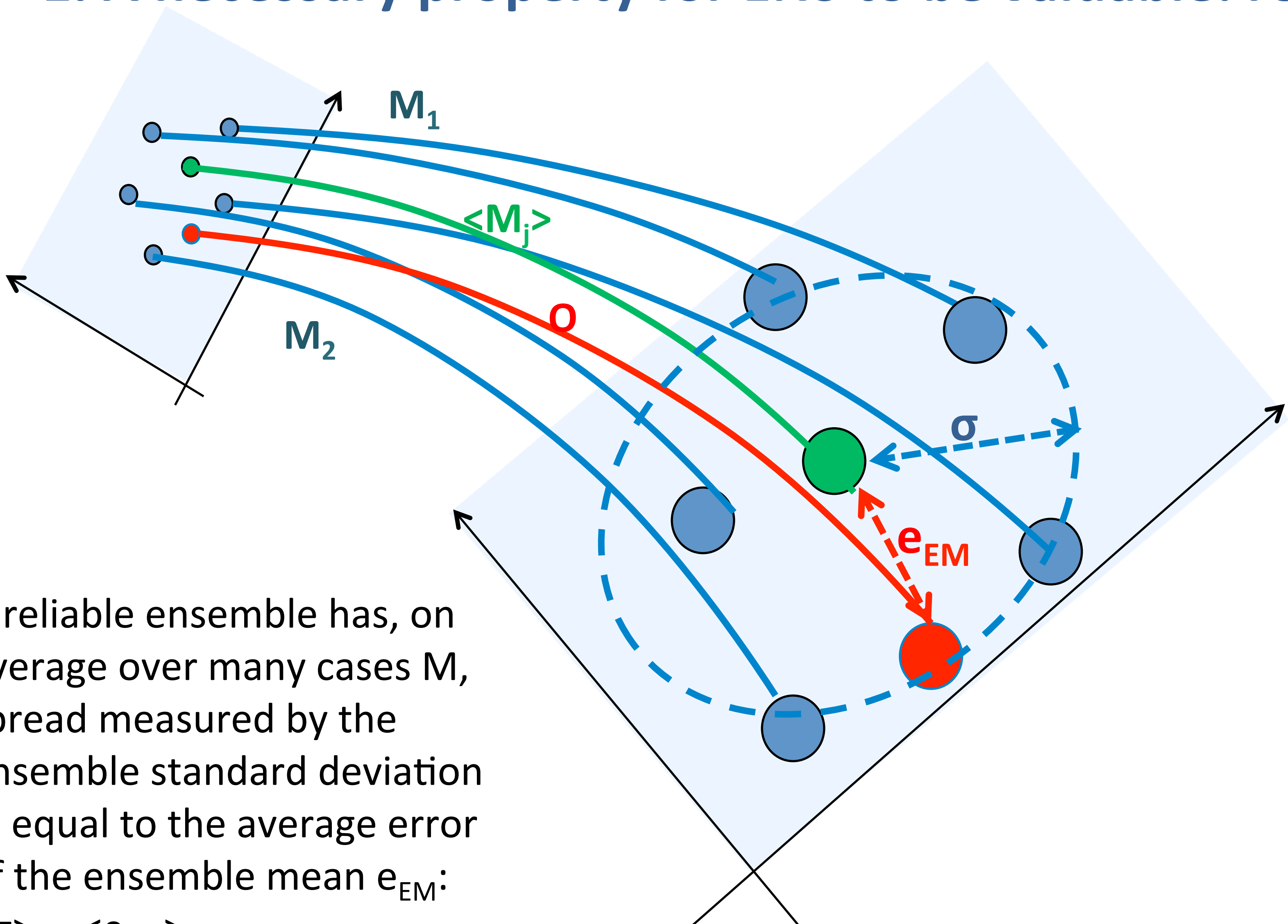
Obs (NEXRAD) 5-10 mm

Boston



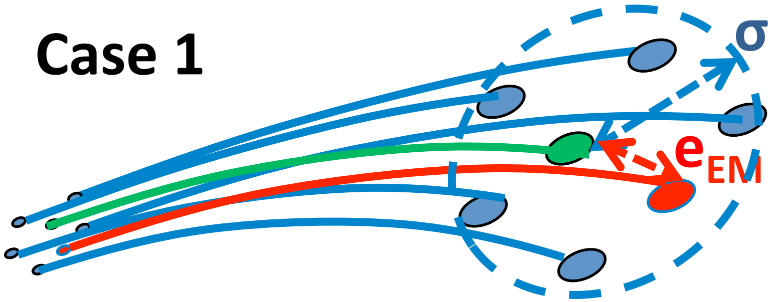
Obs (NEXRAD) 30-35 mm

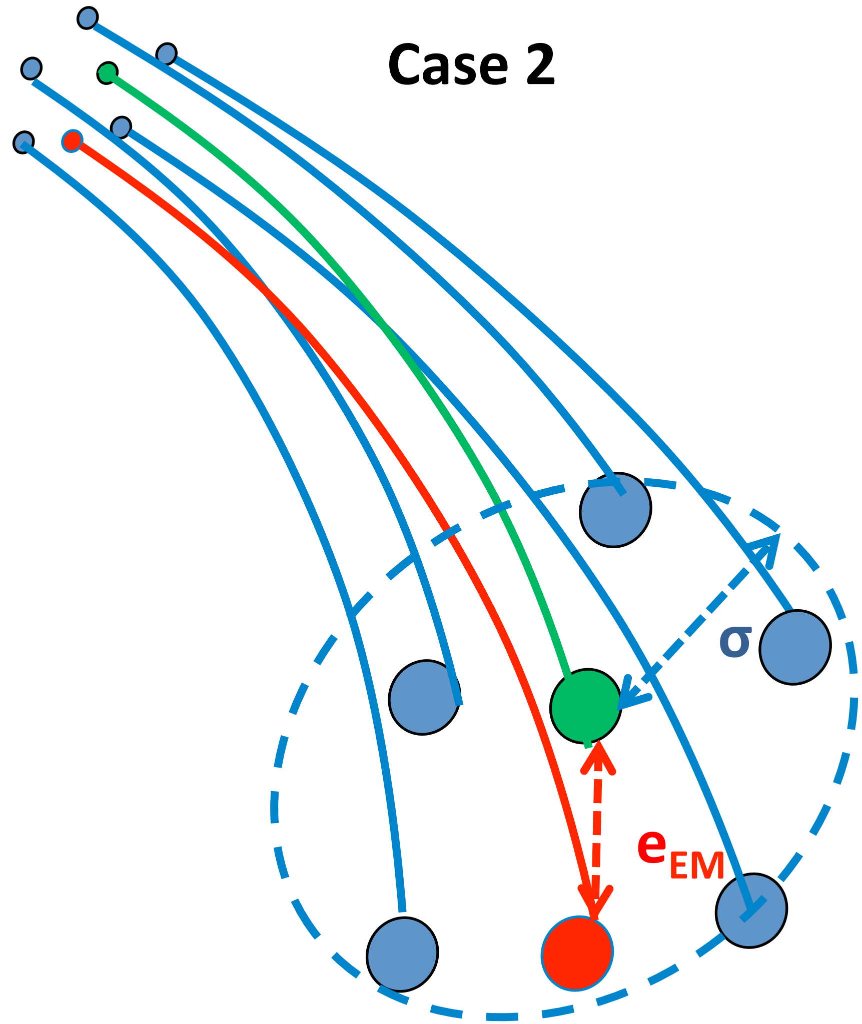
1. A necessary property for ENS to be valuable: reliability



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}:
$\langle\sigma\rangle_M=\langle e_{EM}\rangle_M$

1. In a reliable ENS, small spread >> high predictability



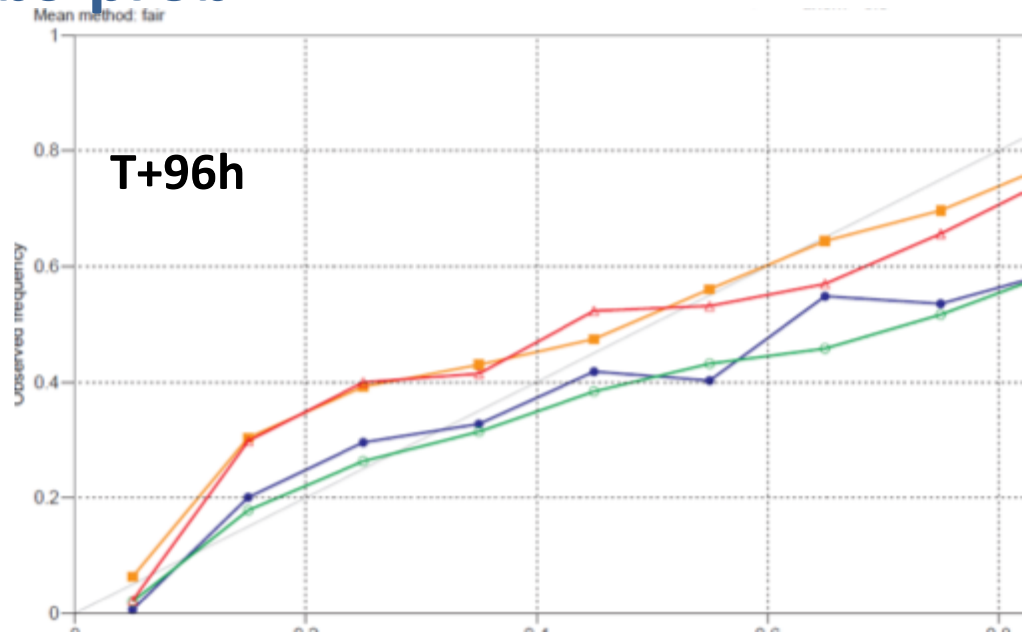


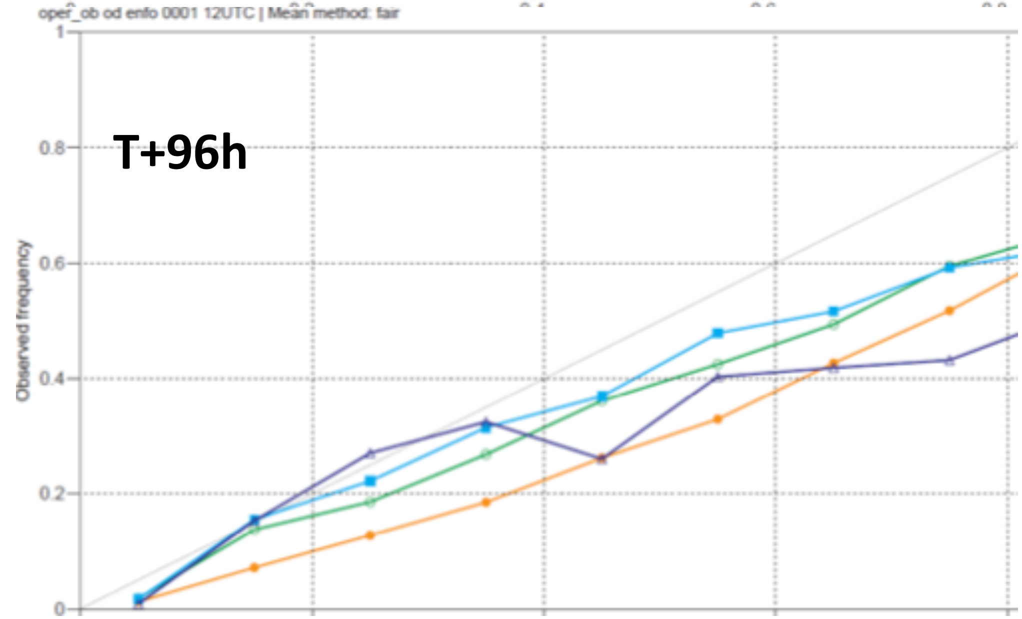
In a reliable ensemble, small ensemble standard deviation indicates a more predictable case, i.e. a small error of the ensemble mean e_{EM}.

1. 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 MAM16 (verified against observations).

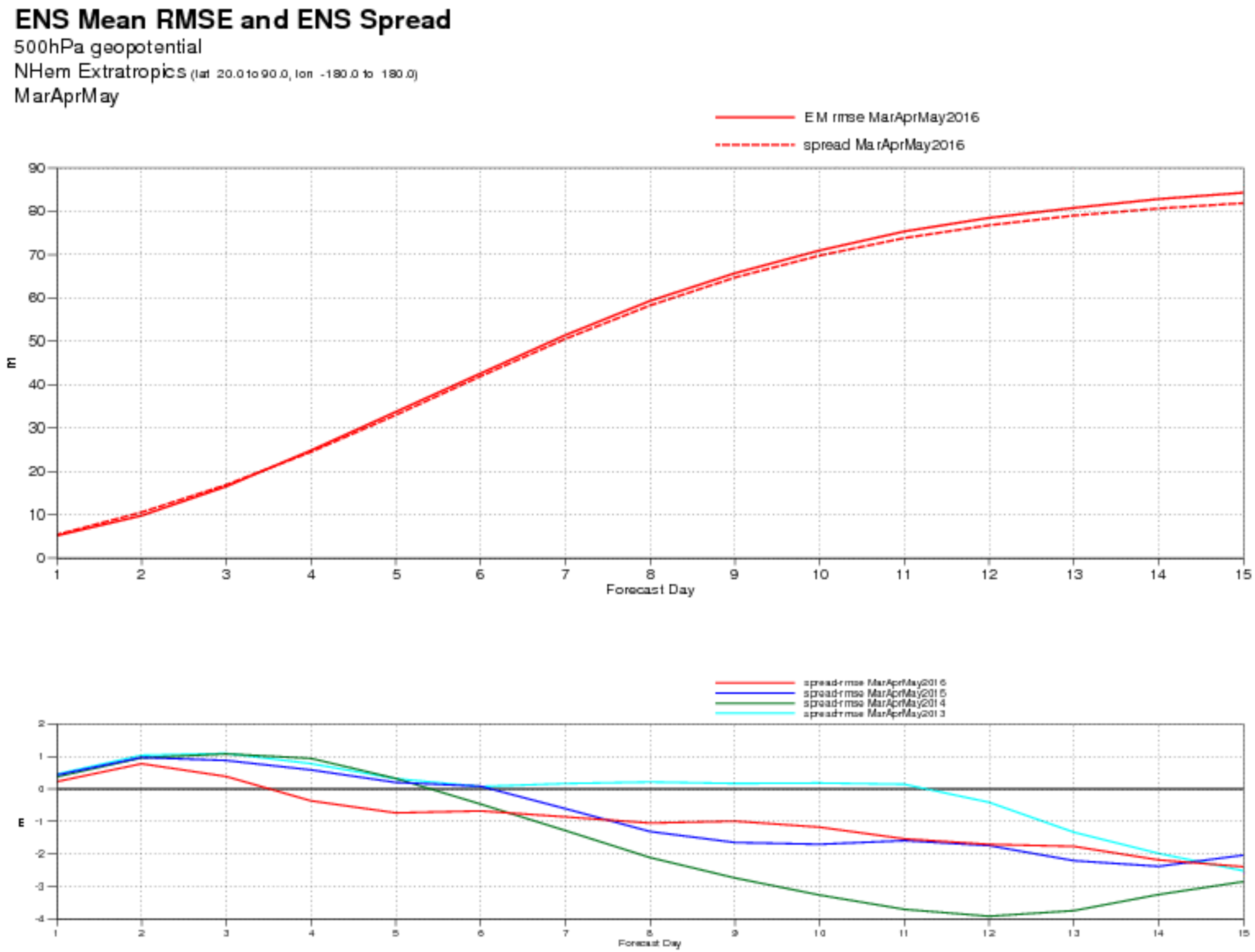




1. ENS reliability: Z500 over NH (MAM16)

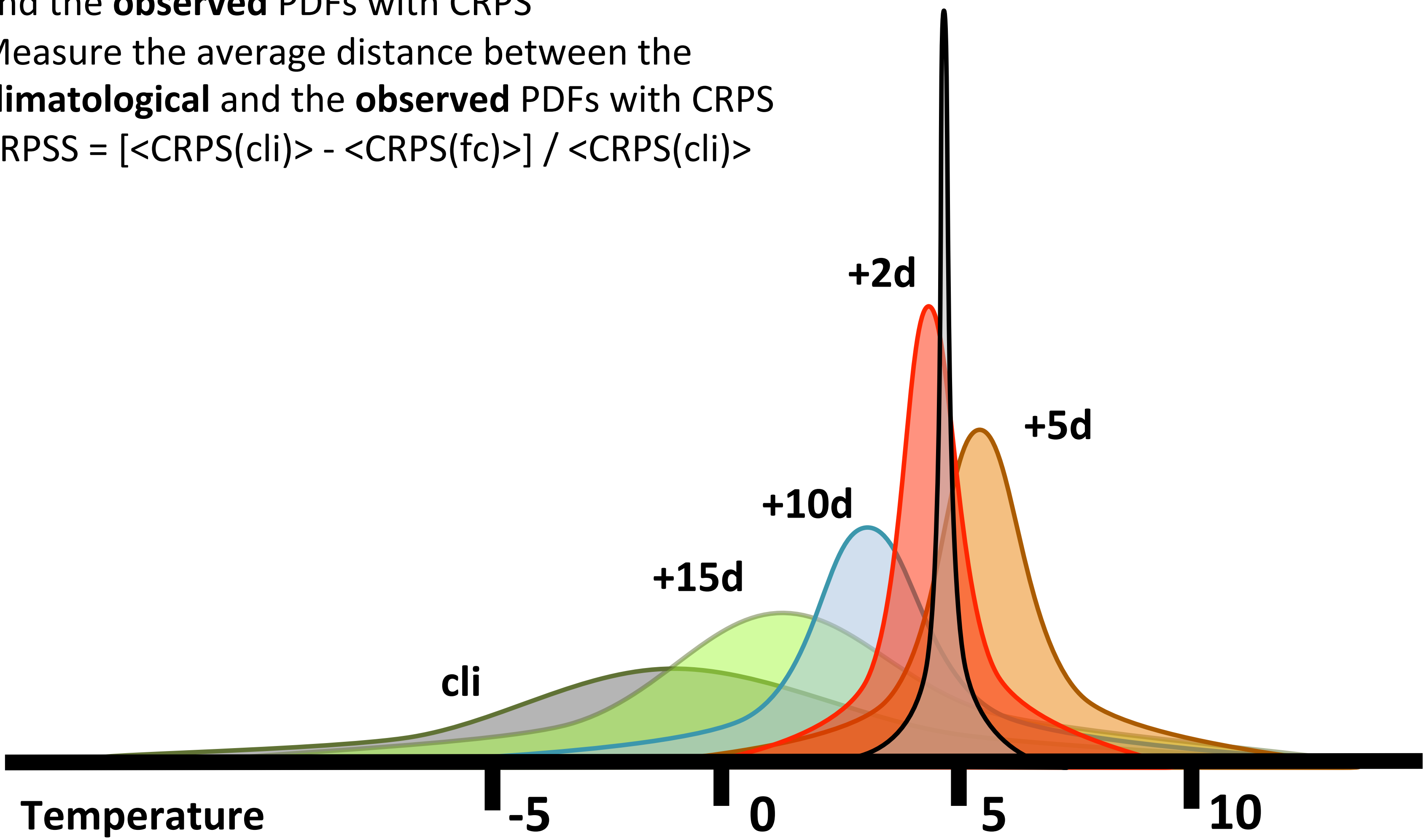
In a reliable ensemble the average spread matches the average error of the ensemble mean.

This plot shows the two curves for Z500 over NH in FMA16.



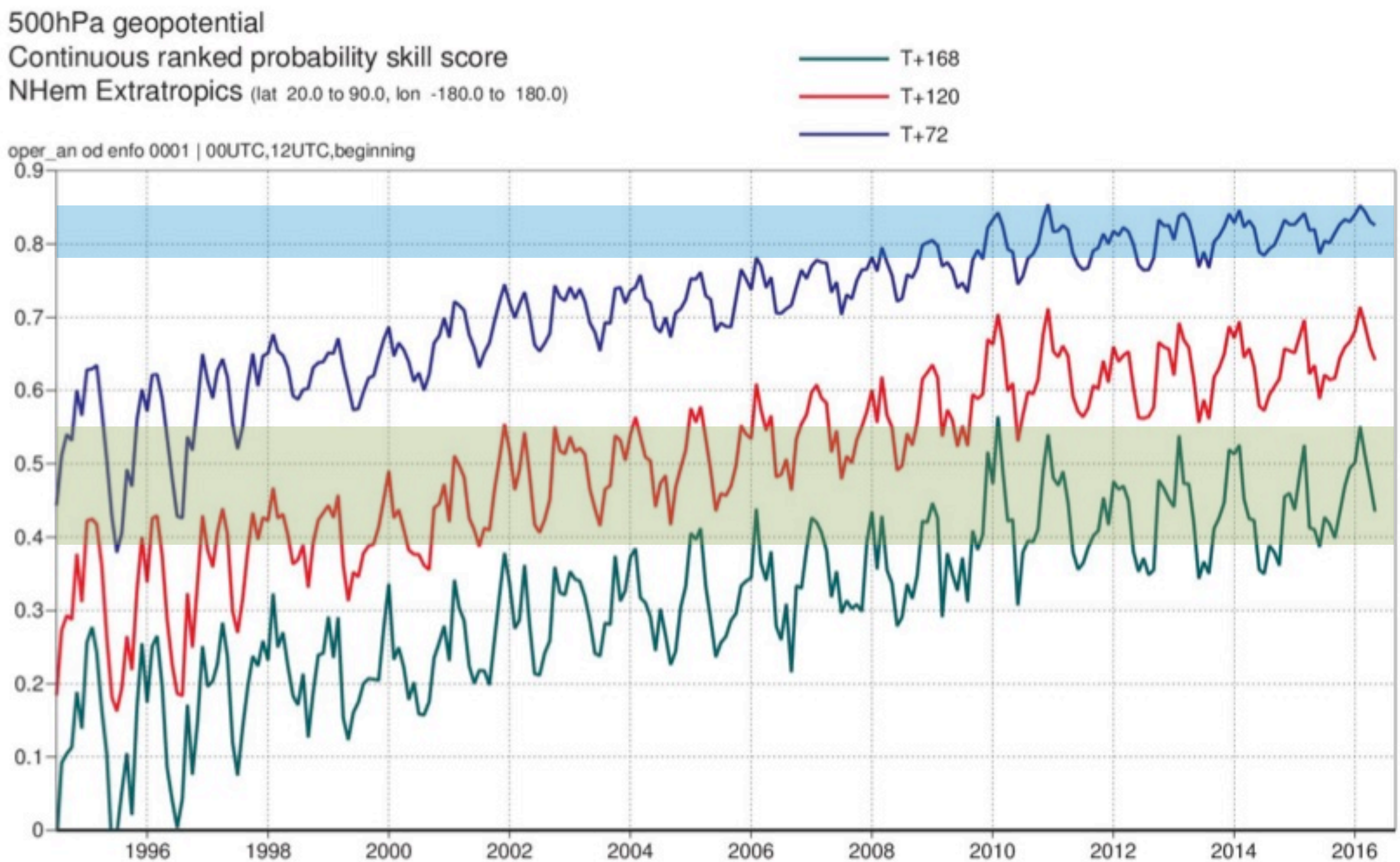
1. PDF forecast skill: how do we measure it?

- 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)>

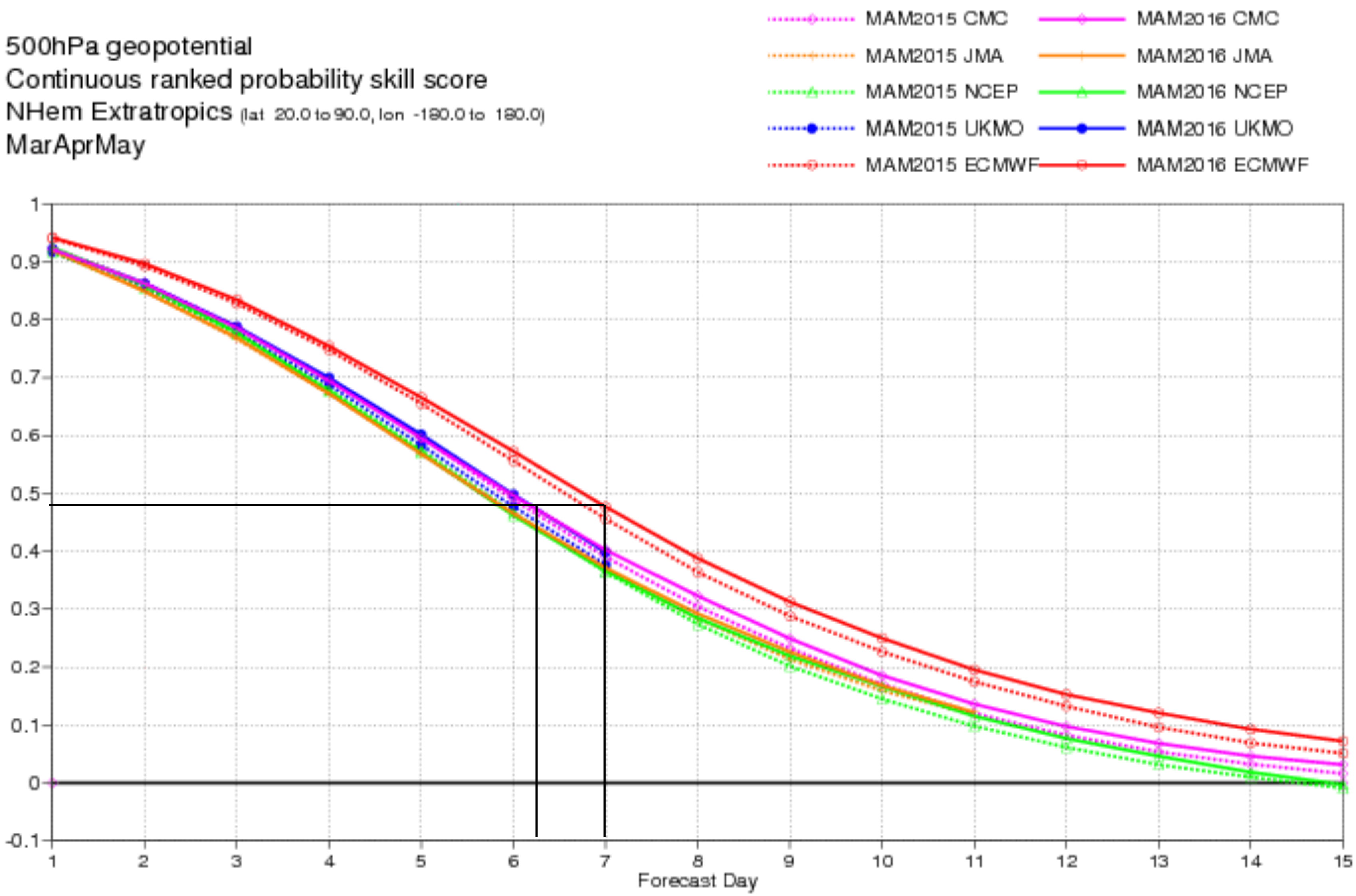


1. ENS skill over the NH

CRPSS is a measure of skill. Today, +7d fcs are as good as +3d fcs 20y ago!

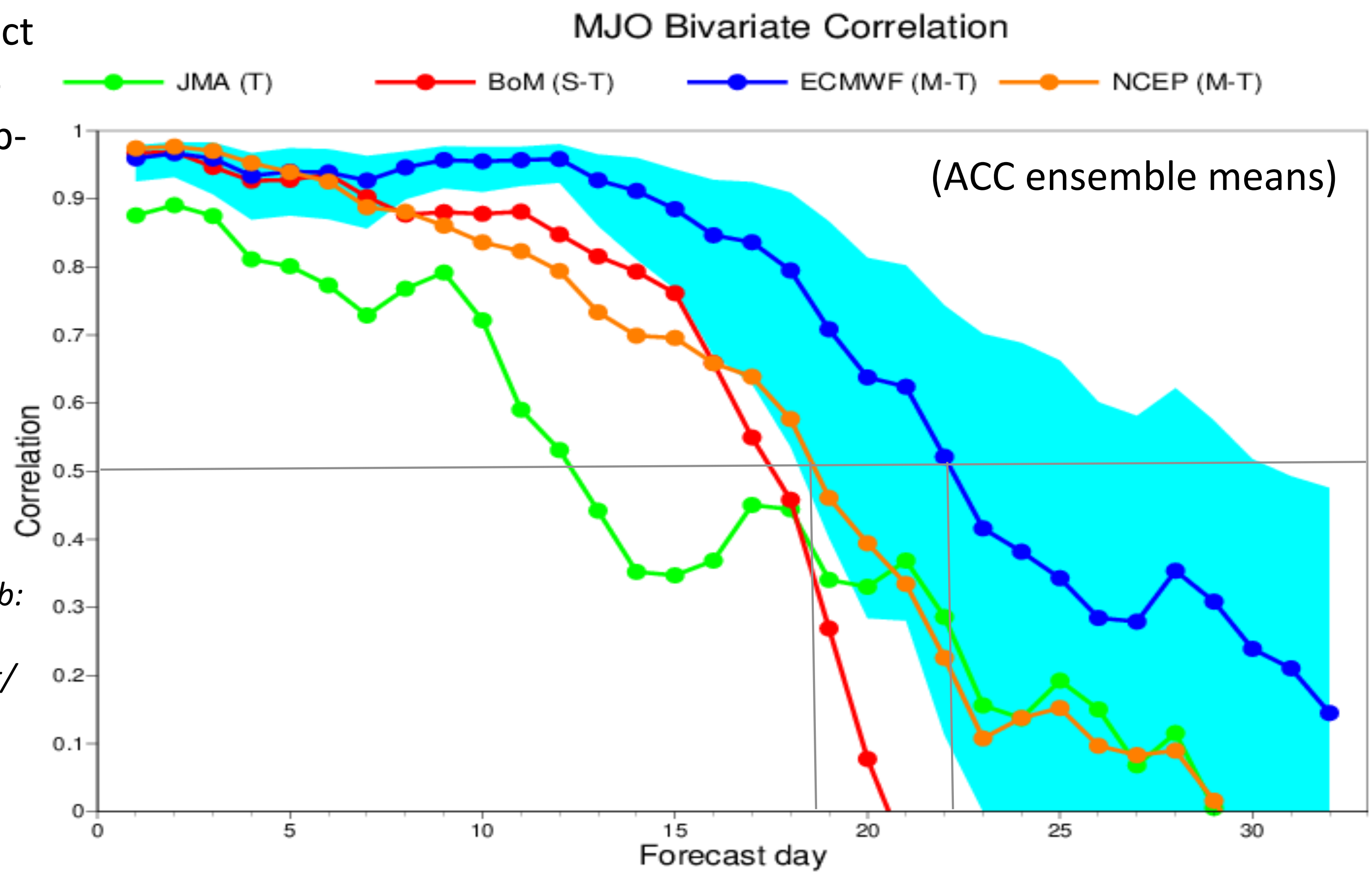


1. The TIGGE ensembles: Z500 over NH, MAM16



1. The S2S ensembles: MJO skill

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

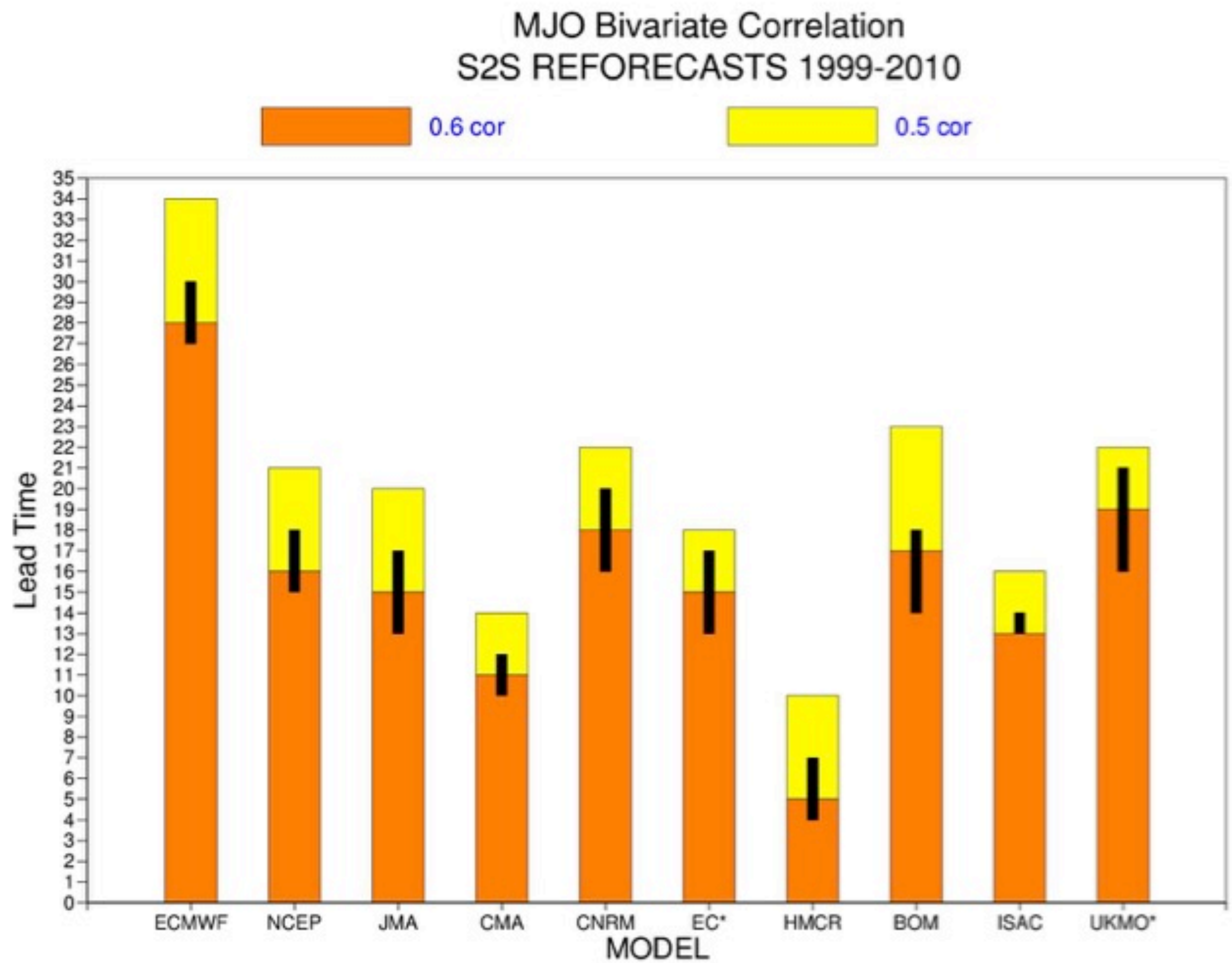


(S2S @ ECMWF web: https://software.ecmwf.int/wiki/display/S2S/Home)

(from Frederic Vitart)

1. The S2S ensembles: MJO skill

Forecast lead-time when the ACC for the prediction of the Madden-Julian-Oscillation (MJO) reaches 0.6 correlation (orange bars) and 0.5 correlation (yellow bars).
The black lines indicate the 95% confidence interval of the time when the 0.6 correlation is reached.
Results are based on the reforecast from 1999 to 2010 from all the models, verified against ERA-Interim analyses.

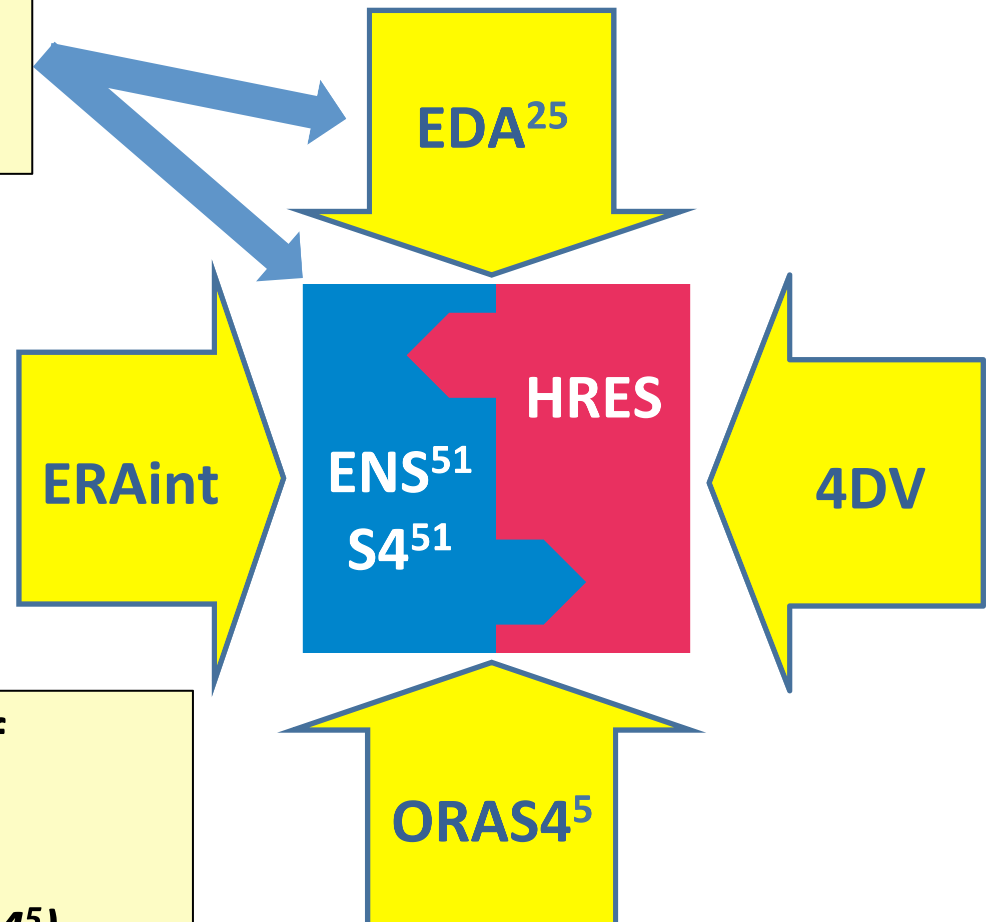


Plan

1. Why do we need ensembles?
2. The ECMWF ensembles
3. The resolution upgrade of March 2016
4. Few key performance indices
5. Prediction of an extreme precipitation event in Houston
6. Conclusions

2. Ensembles are used at analysis and forecast time

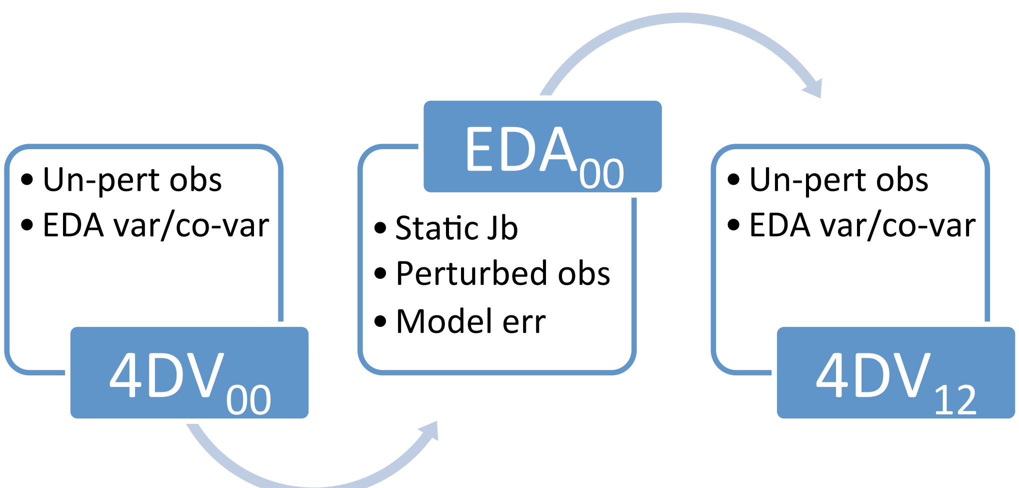
The ensembles simulate the effect of:
- **Observation/initial uncertainties**
- **Model uncertainties (2 stochastic schemes)**



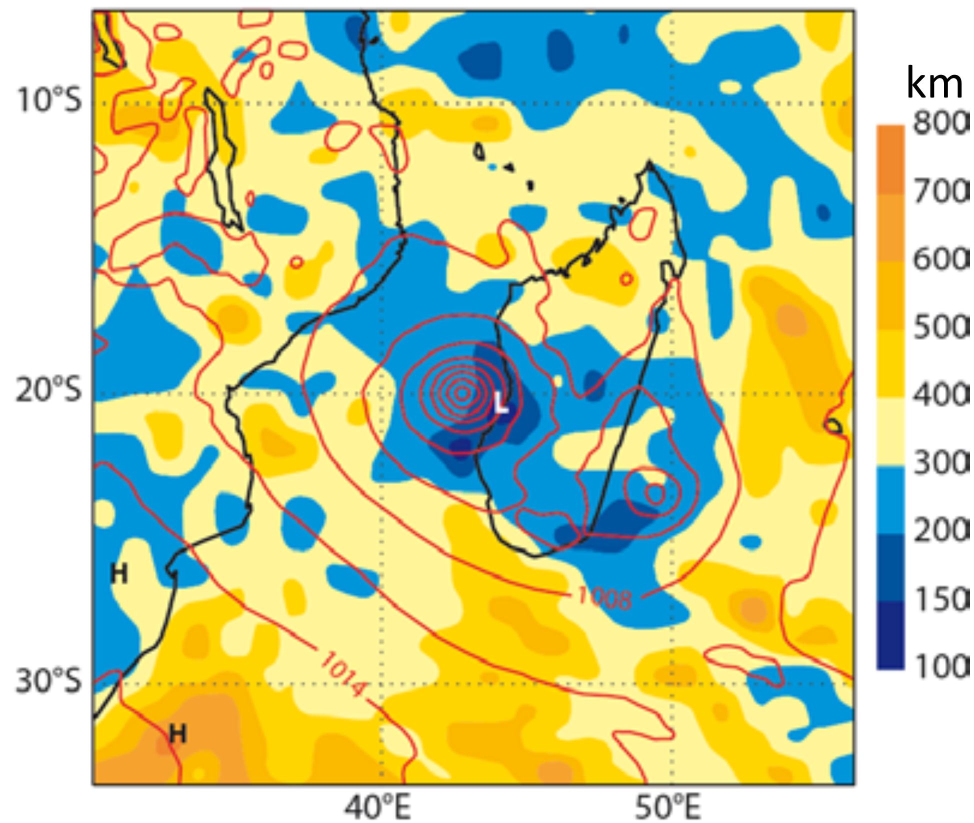
The ensembles give us estimates of the PDF of analyses and forecast states:
PDF(0) << 4DV+EDA^{25}+ORAS4^{5}
(for the refc: PDF(0) << ERAint+ORAS4^{5})
PDF(T) << HRES+ENS^{51}/S4^{51}

2. The EDA is used to provide 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)

2. The medium-range/monthly ensemble (ENS)

ENS includes **51 forecasts** with resolution:

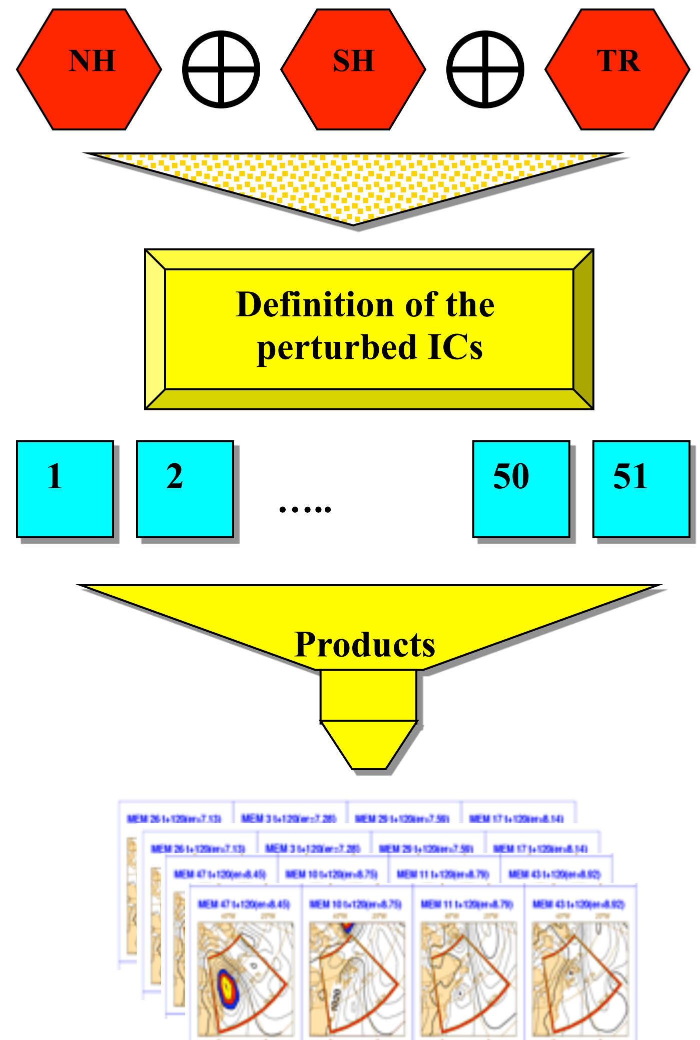
- T_{co}639L91 (**~18km**, 91 levels) from **day 0 to 15**
- T_{co}319L91 (**~36km**, 91 levels) from **day 15 to 46** (only at 00UTC on Mon and Thu).

Initial uncertainties are simulated by adding to the unperturbed analyses a combination of **T42L91 singular vectors**, computed to optimize total energy over 48h hours, and perturbations generated by the ECMWF **T_{co}639L137 EDA** (Ensembles of Data Assimilation).

Model uncertainties are simulated by adding stochastic perturbations to the tendencies due to parameterized physical processes (**SPPT and SKEB schemes**).

The unperturbed analysis is given by the T_{co}1279L137 4DVAR.

ENS runs daily at 00 and 12 UTC, with a TOA at 0.01 hPa.



2. The ECMWF operational ensemble (ENS) today

Each ensemble forecast is given by the time integration of perturbed equations

$$e_j(d,T) = e_j(d,0) + \int_0^T [A(e_j,t) + P(e_j,t) + \underline{\delta P_j(e_j,t)}]dt$$

$$\delta P_j(\lambda,\varphi,p) = r_j(\lambda,\varphi)P_j(\lambda,\varphi,p) + F_\Psi(\lambda,\varphi,p)$$

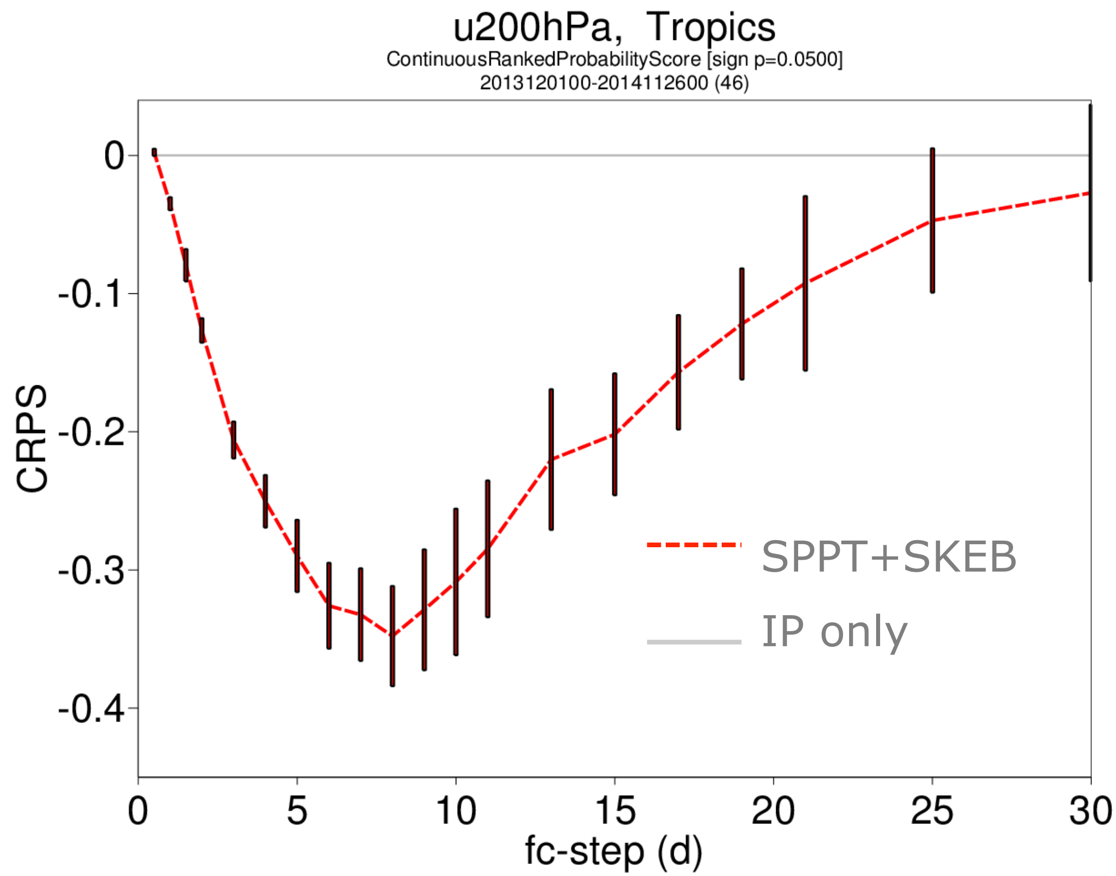
SPPT: Stochastically Perturbed Parameterized Tendencies
(to represent uncertainty associated with parameterisations)

SKEB: Stochastic Kinetic Energy Backscatter
(to represent unresolved upscale energy transfer)

2. Including model uncertainties improves ENS forecasts

Including model perturbations via the SPPT and SKEB schemes gives rise to better ensembles:

- The ensemble is **more reliable** (less under-dispersive);
- **Improved probabilistic skill** is observed at a range of lead times (e.g. see in plot improved CRPS for high-level winds in the Tropics)



Plan

1. Why do we need ensembles?
2. The ECMWF ensembles
3. The resolution upgrade of March 2016
4. Few key performance indices
5. Conclusions

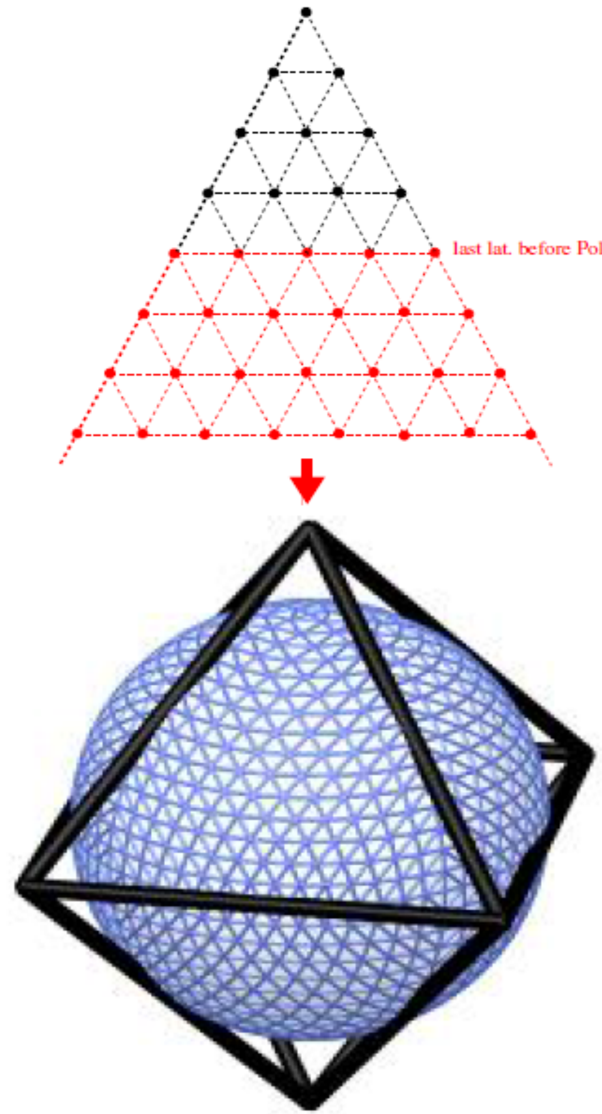
3. The octahedral cubic grid

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.

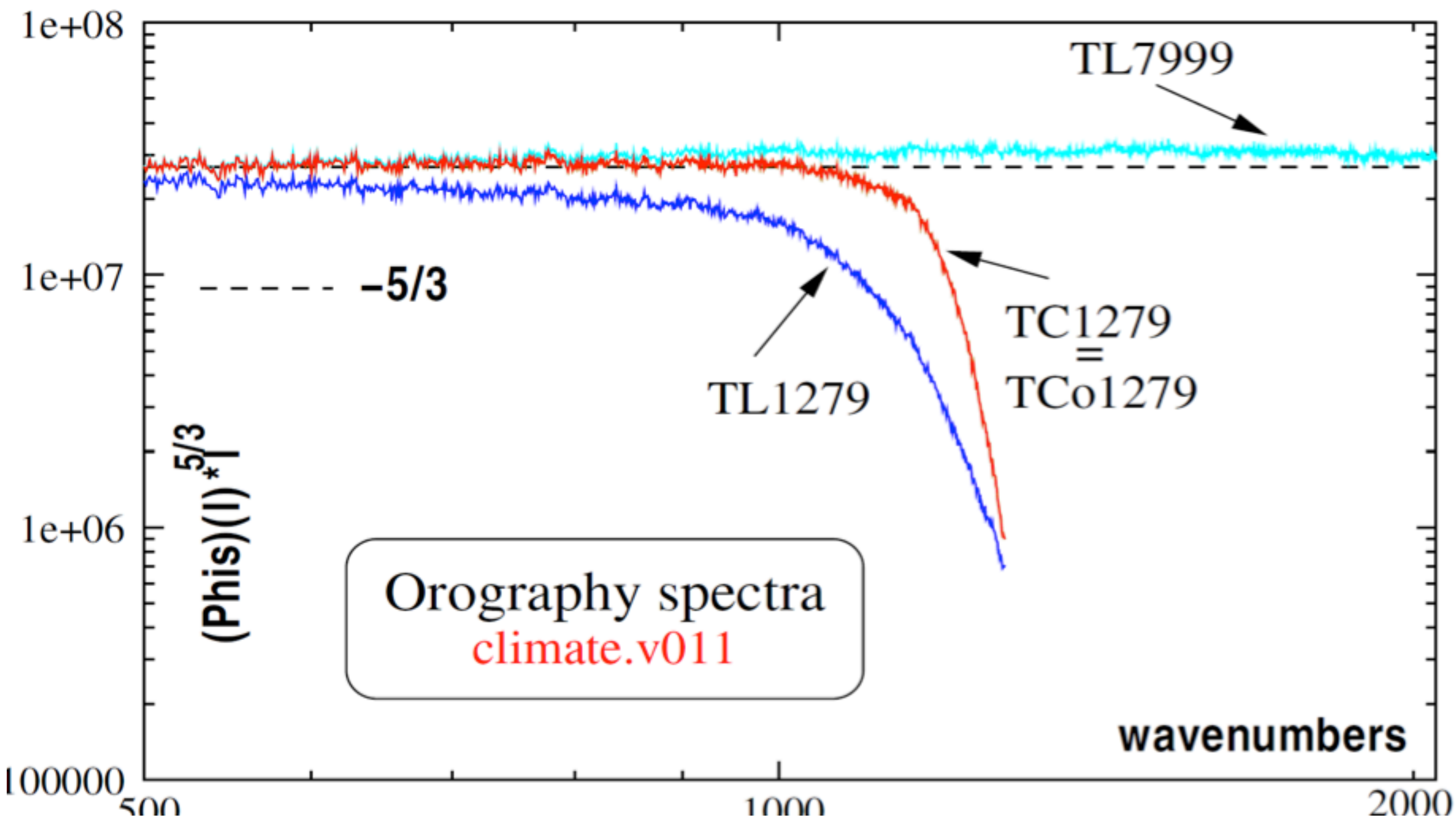
Number of points per latitude

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



3. Orographic variance of different grids



3. March 2016: the resolution of all suites was increased

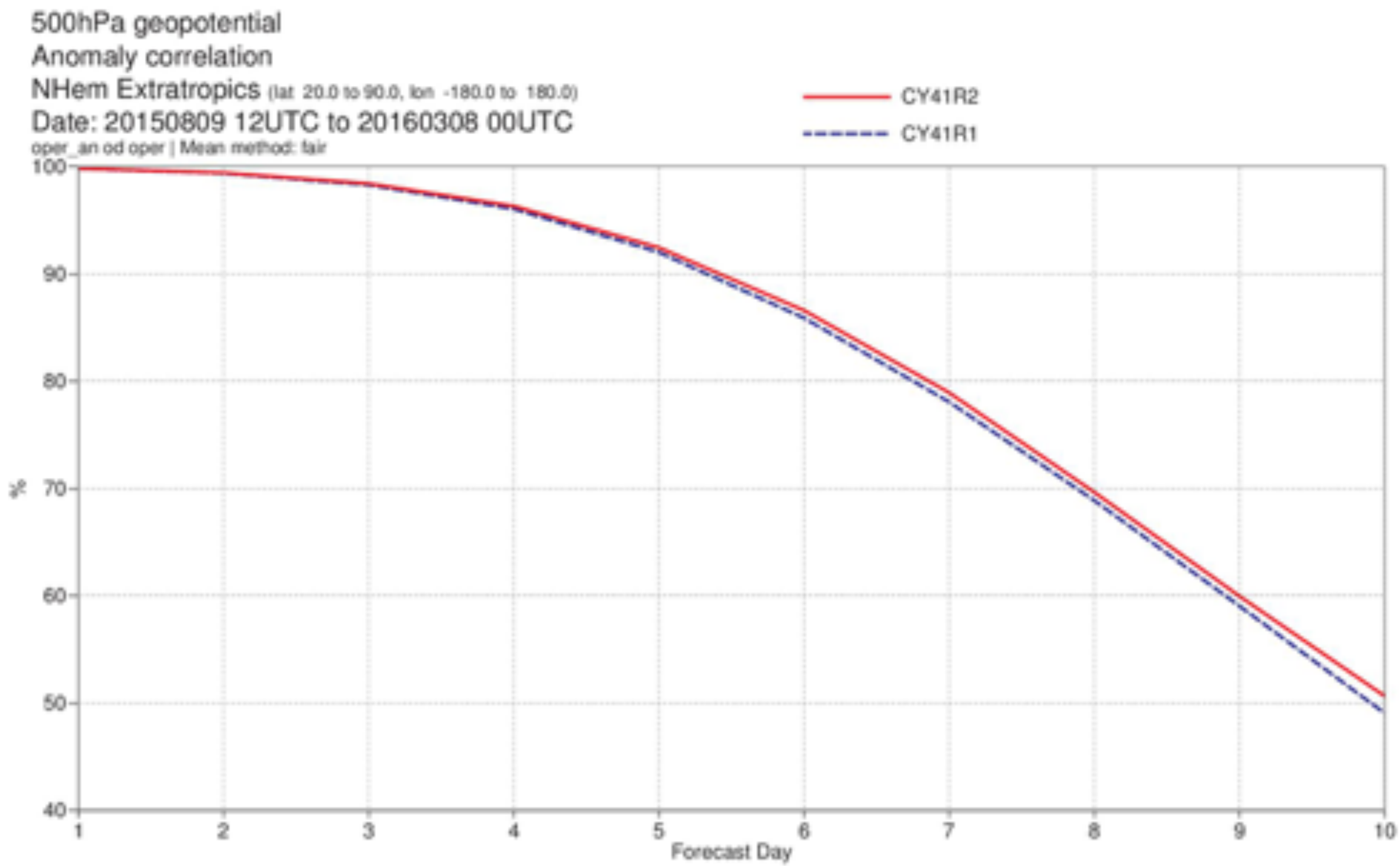
	Operational suite	Uncertainty sources		
		Obs	ICs	Model
HRES - 9km	T_{CO}1279 (~9 km) L137 (0-10d)	--	--	--
4DVAR- 9km	T_{CO}1279 (inner T_{CO}255/319/399) L137	--	--	--
EDA – 18km	25 members: T_{CO}639 (~18km) L137	δo	--	SPPT(1L)
ENS – 18km	51 members: T_{CO}639 (~18km) L91 (0-15d) T_{CO}319 (~36km) L91 (15-46d) - Ocean: NEMO ORCA100z42 *(to 025z75 in Q4)*	-- --	EDA^{25} & $SVs^{50*Nareas}$ $ORAS4^{5}$	SPPT(3L) & SKEB --
S4 – 80km	51 members: T_L255 (~80km) L91 *(to T_{CO}319L137 in 2017)* - Ocean: NEMO ORCA100z42 *(to 025z75 in 2017)*	-- --	SVs $ORAS4^{5}$	SPPT(3L) & SKEB --

T_{CO} – Cubic octahedral Gaussian reduced grid

T_L – Gaussian linear grid

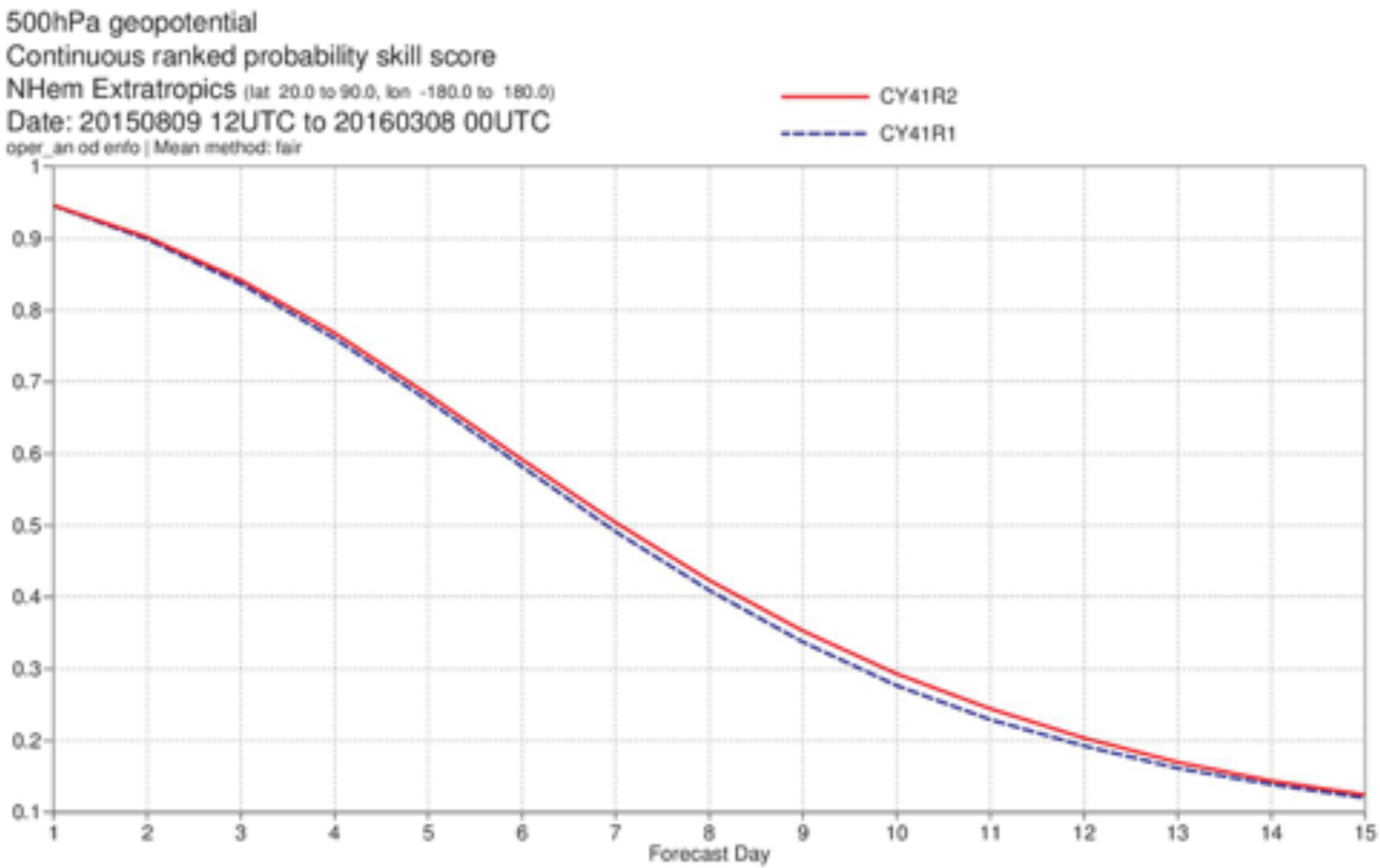
3. Impact of recent upgrade on single high-resolution fc

Results based on 7 months indicate predictability gains for single high-resolution forecasts in the medium-range of **about 3 hours** (up to 6 hours for surface variables).



3. Impact of recent upgrade on ensemble fcs

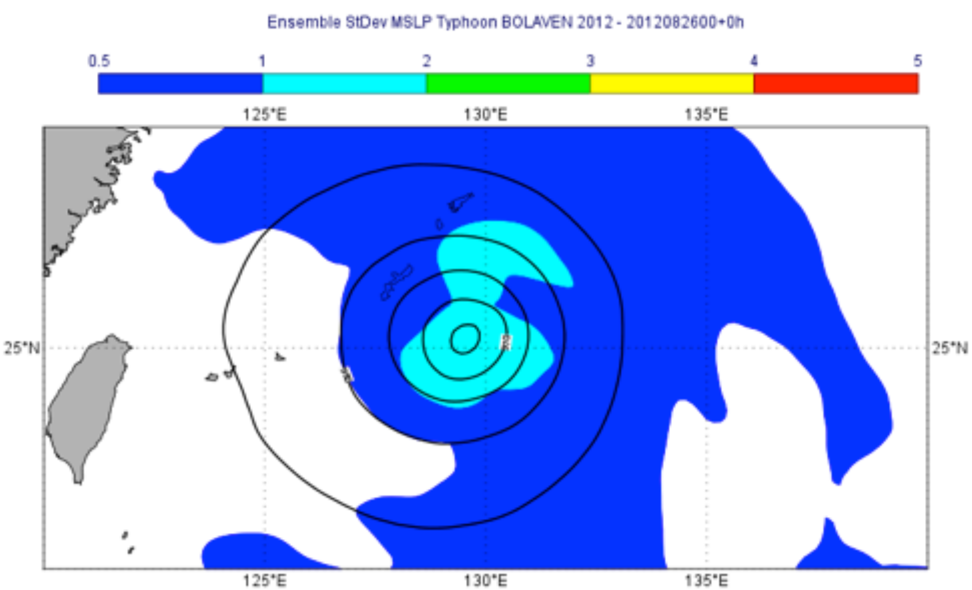
Results based on 7 months indicate predictability gains for the ensemble in the medium-range of **about 6 hours** (up to 12-18 hours for surface variables).



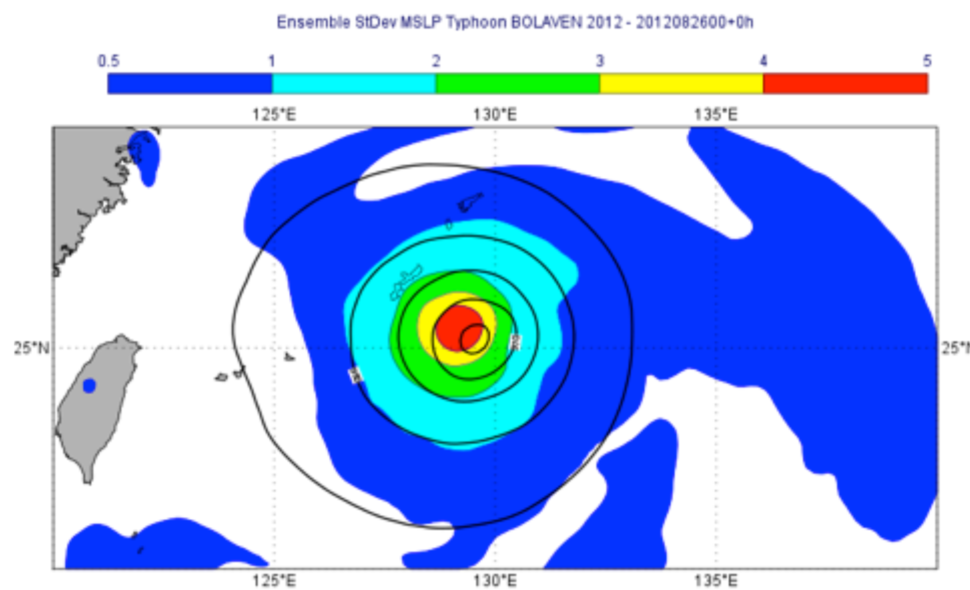
3. Impact of using a higher-resolution EDA

An increase in the resolution of the EDA outer loops from T_L399 to T_L639 improves 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



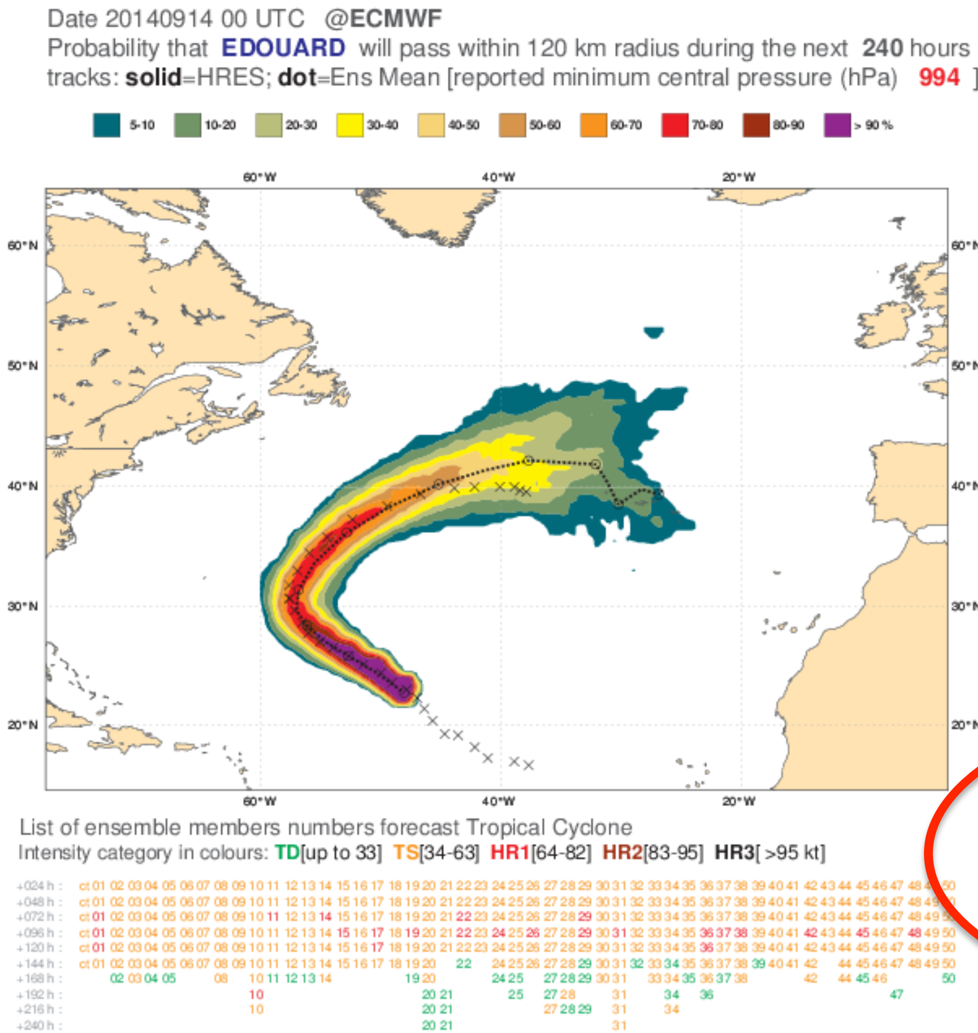
Ensemble with perturbations from 639 EDA

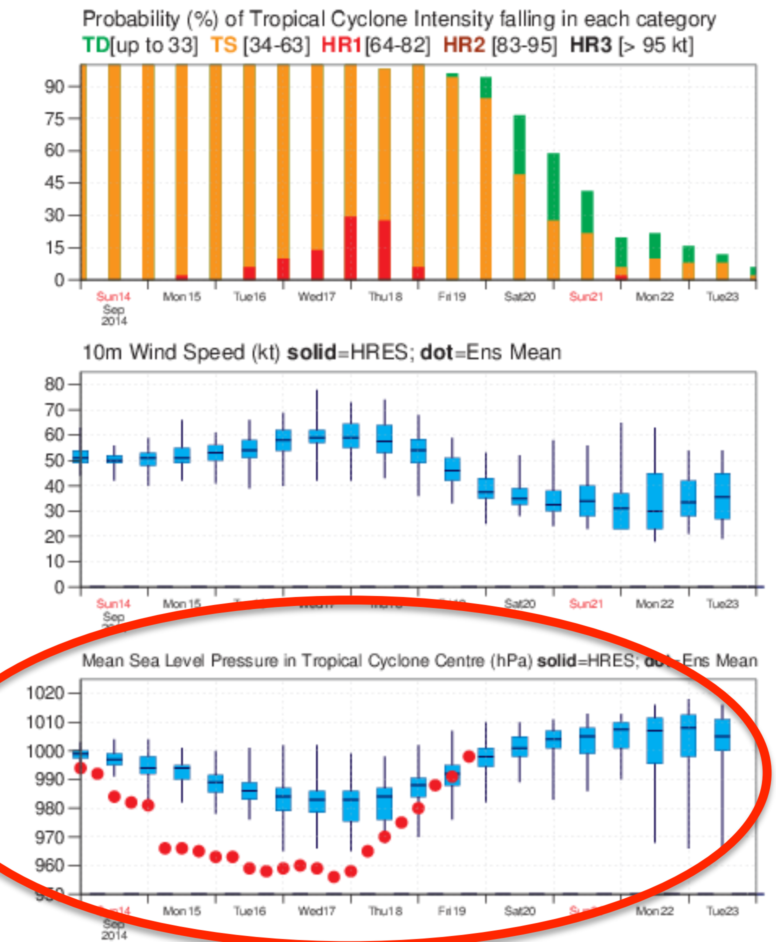


(From Simon Lang)

3. Combined impact of increasing EDA & ENS resolution

OPE: T_L1279 An, T_L399 EDA, T_L639 ENS

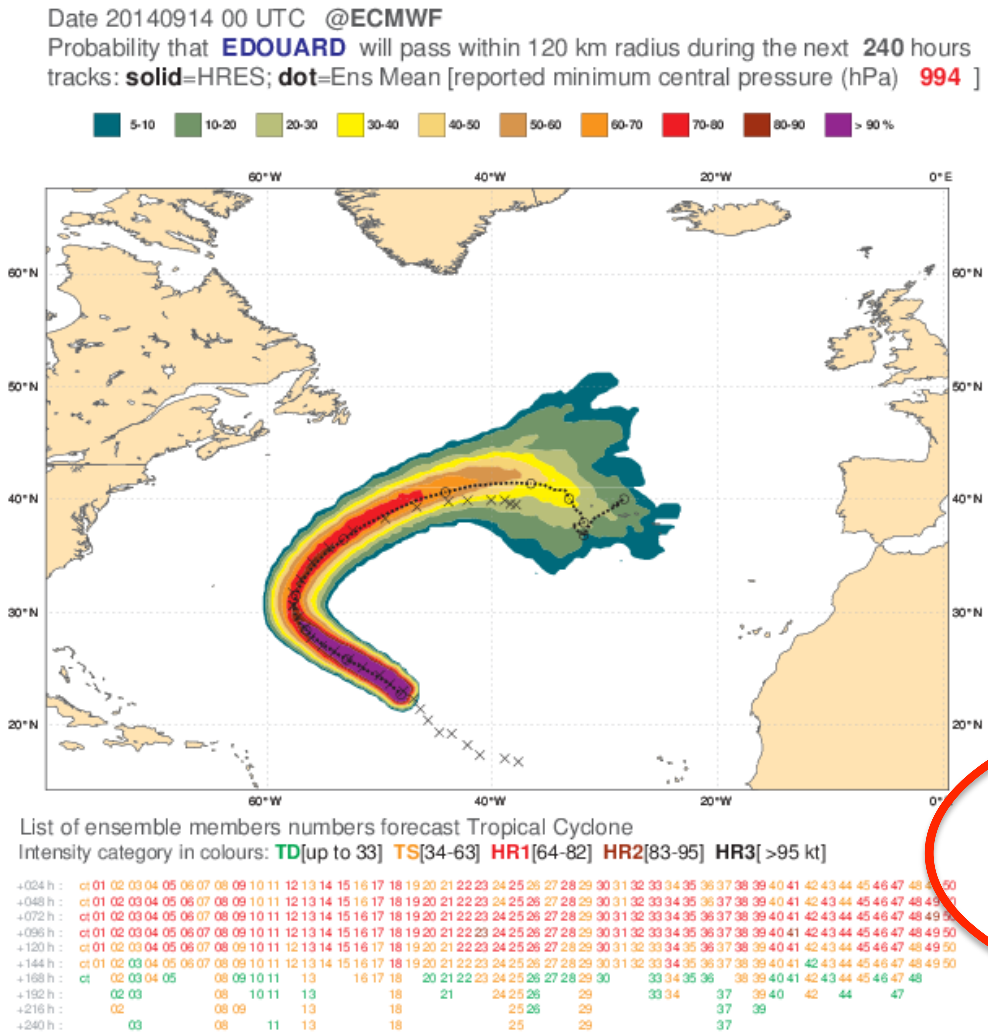




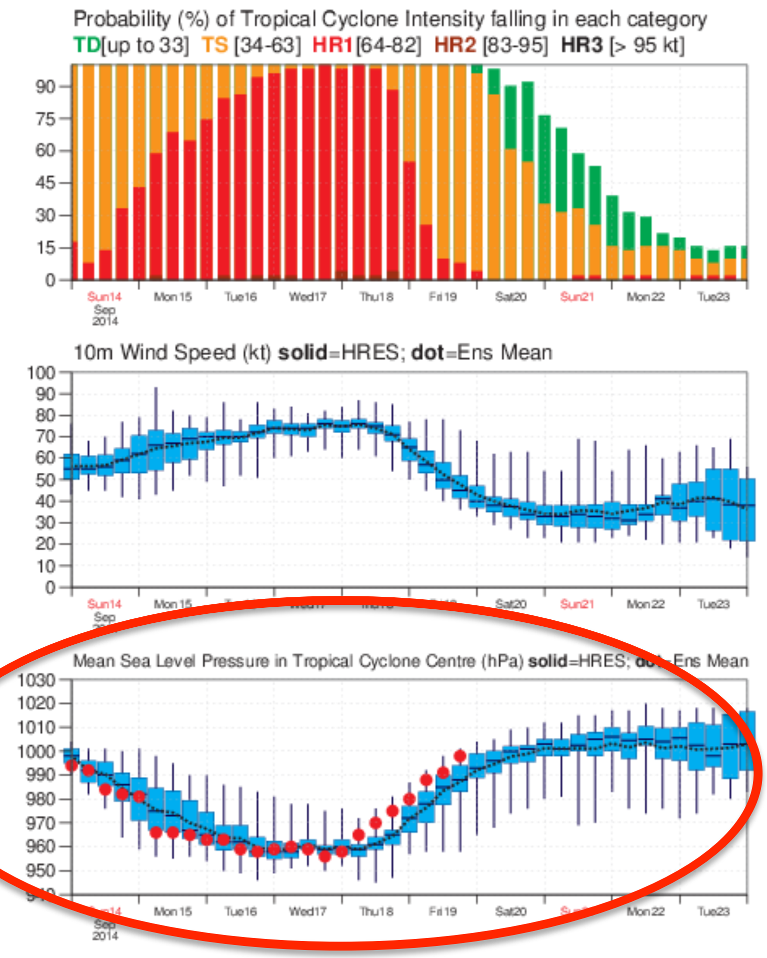
Too small initial spread and deepening not captured

(From Simon Lang)

3. Combined impact of increasing EDA & ENS resolution



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



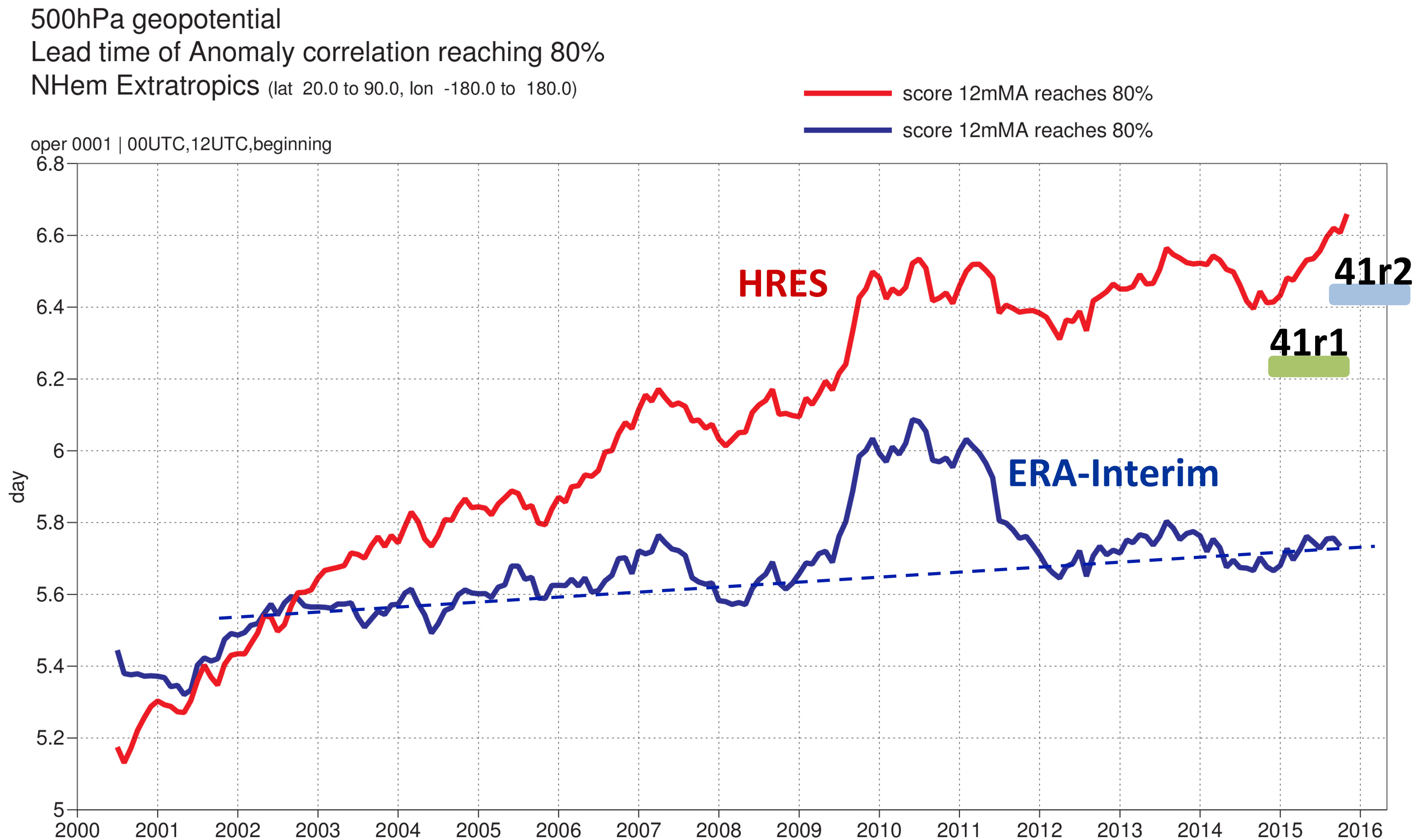
(From Simon Lang)

Deepening is captured and spread reflects uncertainty

Plan

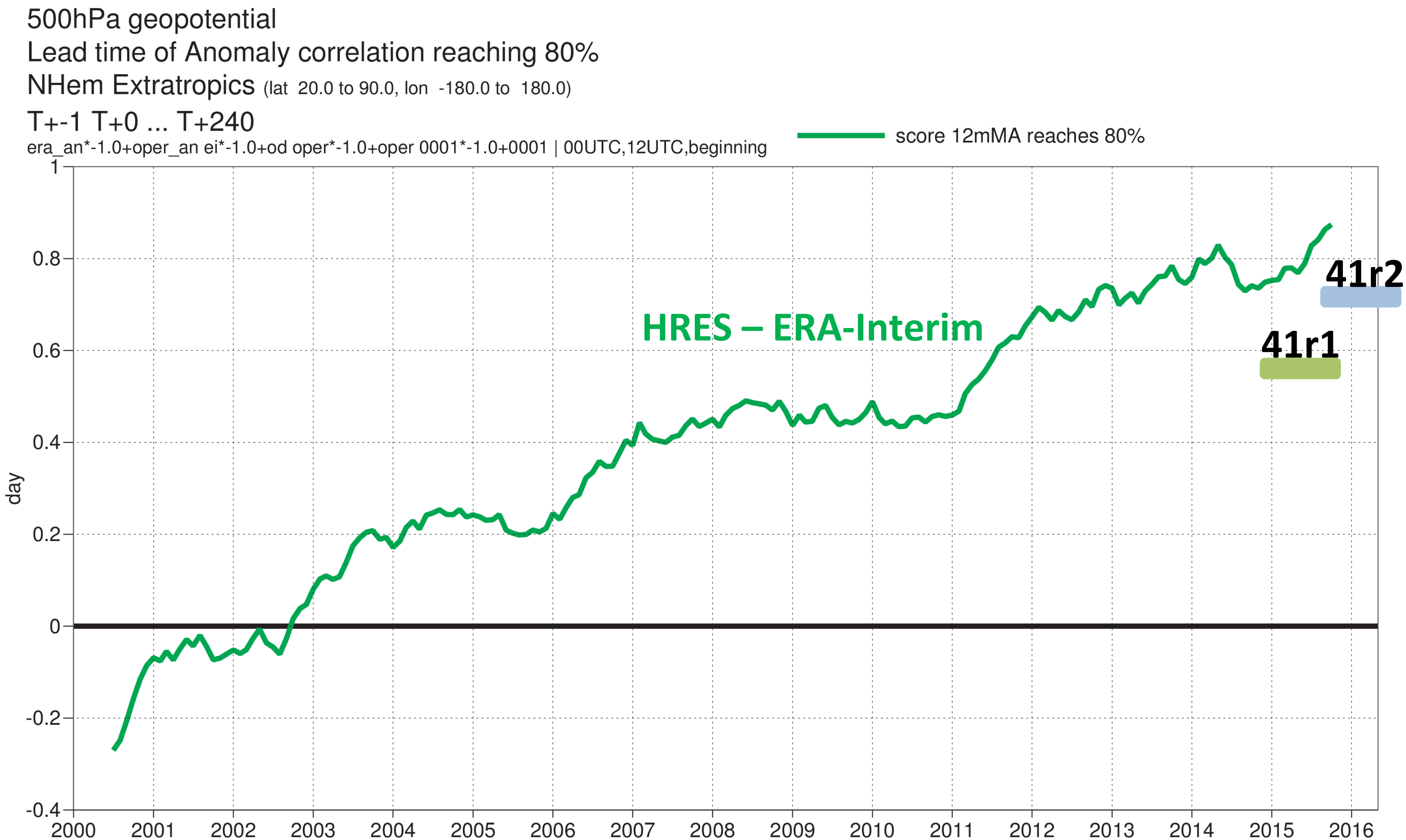
1. Why do we need ensembles?
2. The ECMWF ensembles
3. The resolution upgrade of March 2016
4. Few key performance indices
5. Prediction of an extreme precipitation event in Houston
6. Conclusions

4. HRES and ERAI headline score – Z500 ACC NH



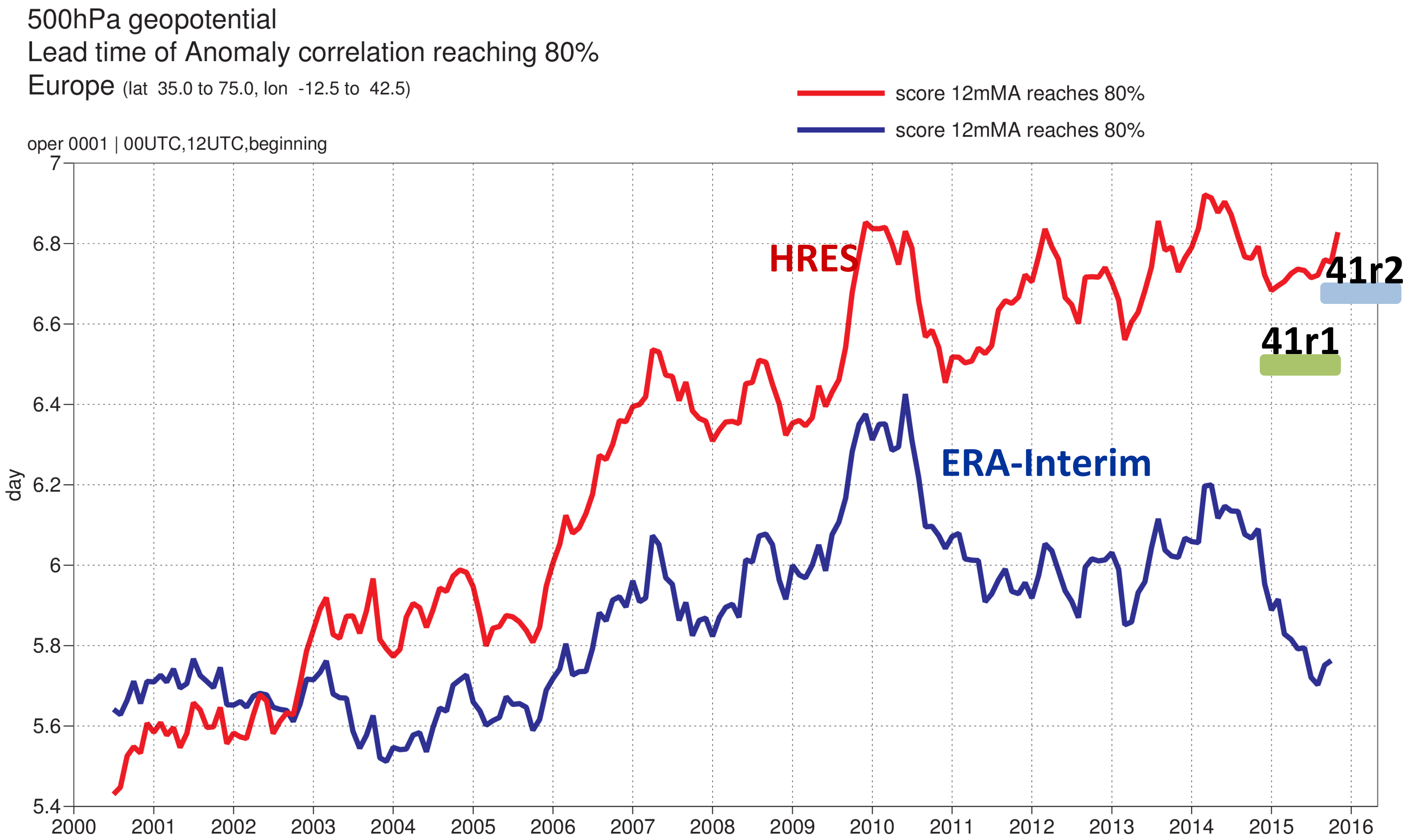
(From Thomas Haiden)

4. (HRES-ERAI) headline score – Z500 ACC NH



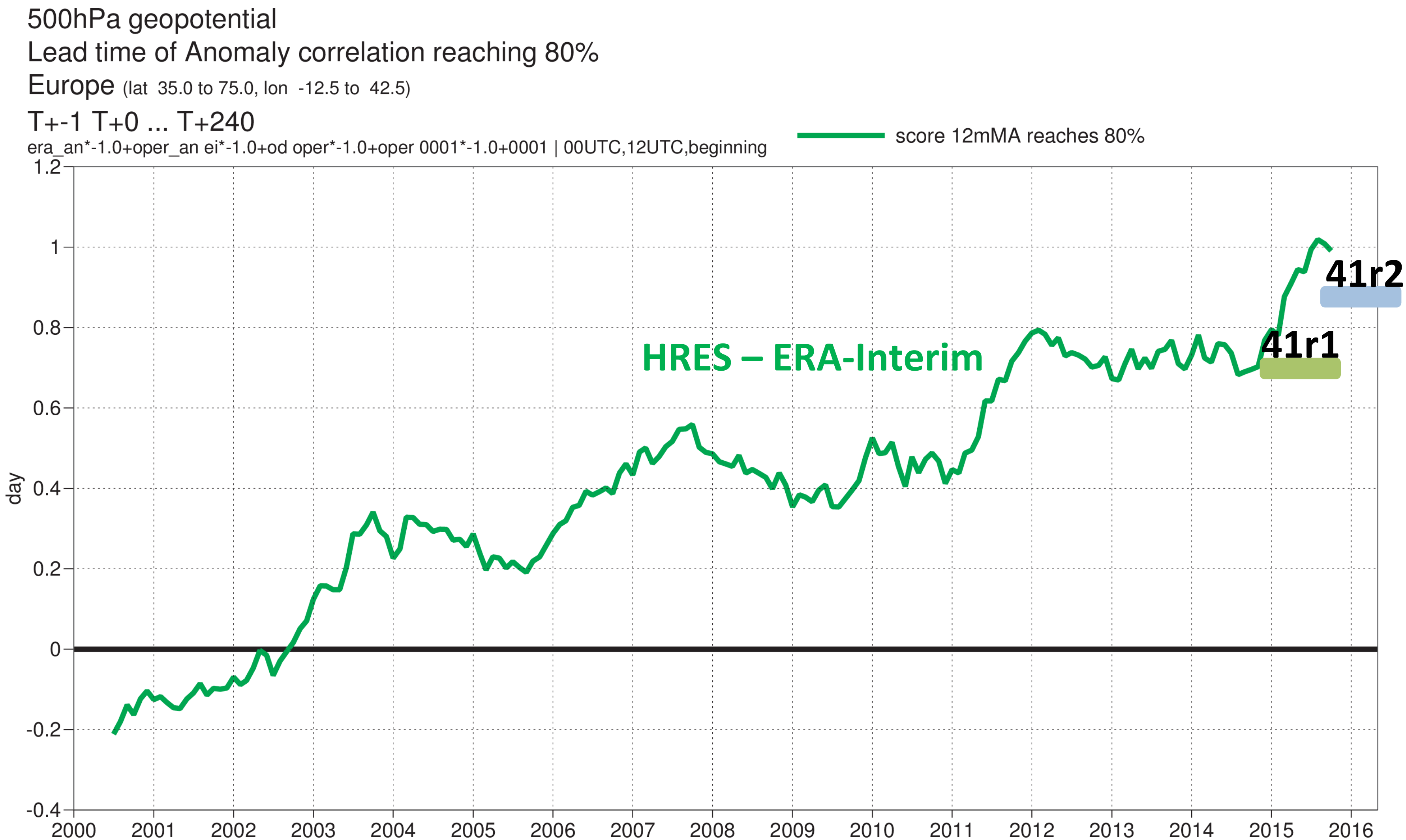
(From Thomas Haiden)

4. HRES and ERAI headline score – Z500 ACC Europe



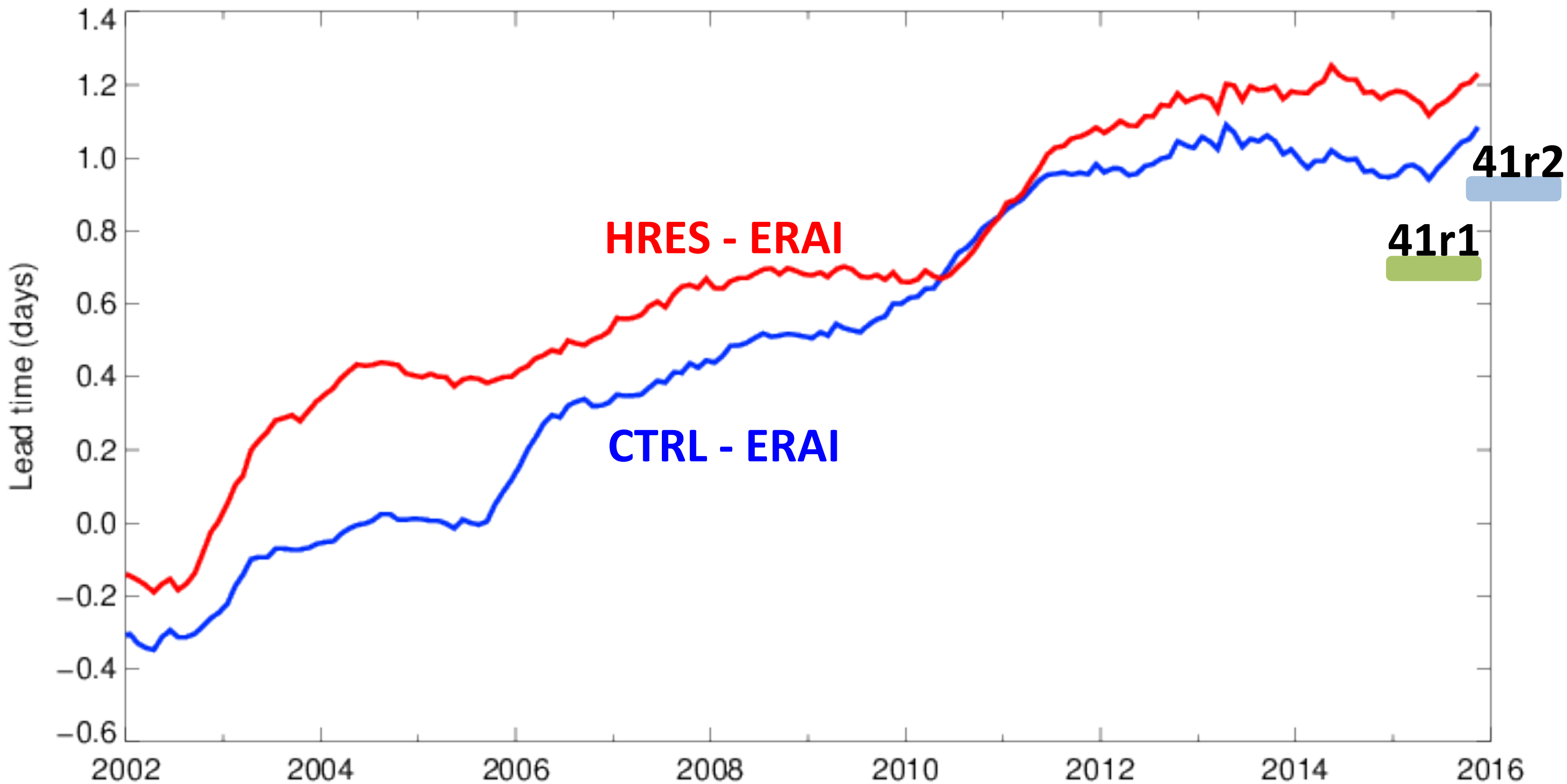
(From Thomas Haiden)

4. (HRES-ERAI) headline score – Z500 ACC Europe



(From Thomas Haiden)

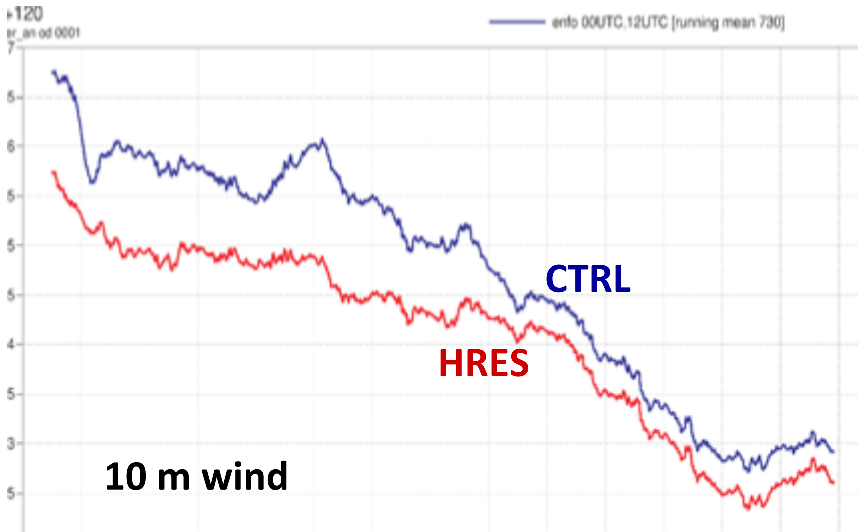
4. HRES, ENS-CON, ERAI – 24-h precipitation NH

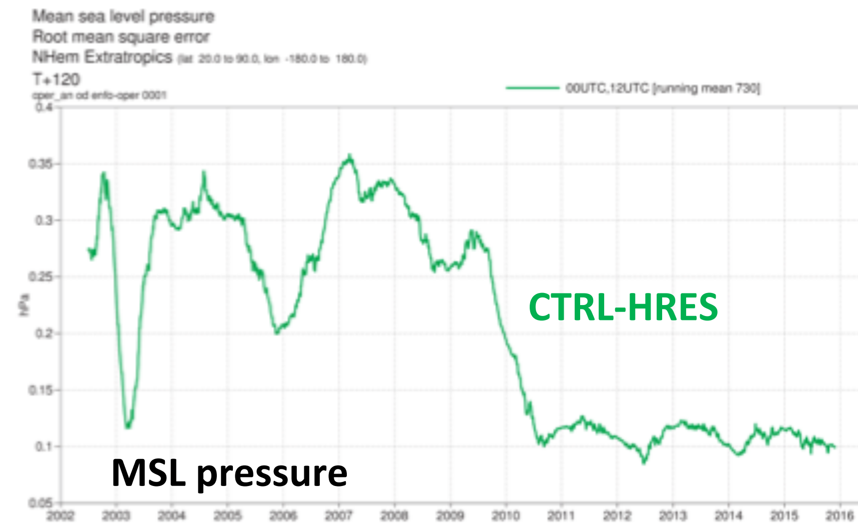


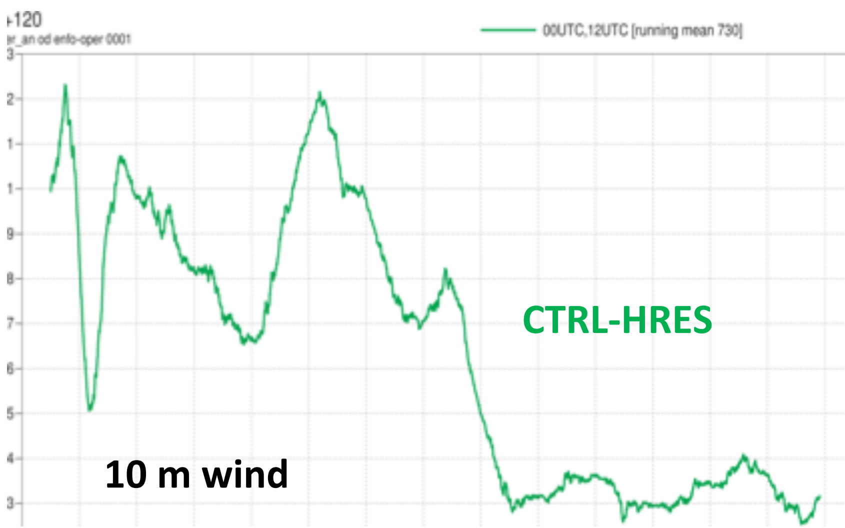
Difference HRES-CTRL: reduced with 41r1, but no strong signal from 41r2 (yet).

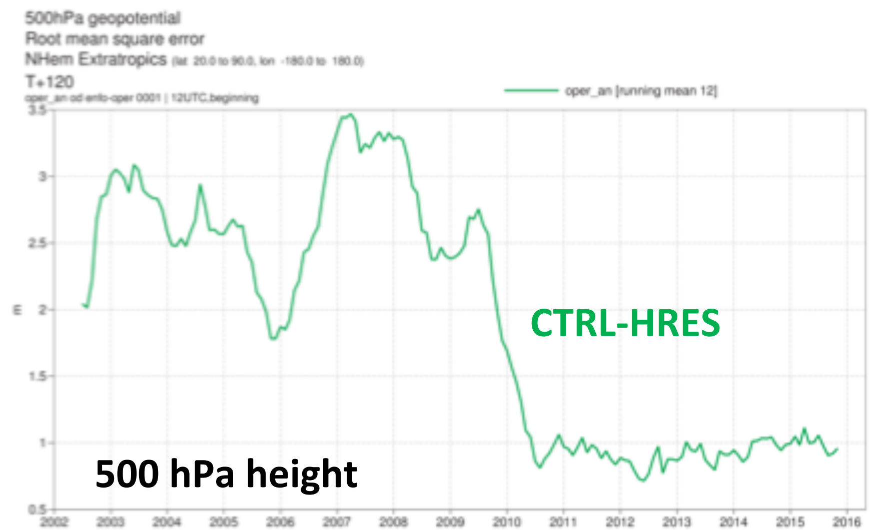
(From Thomas Haiden)

4. HRES and ENS-CON – NH





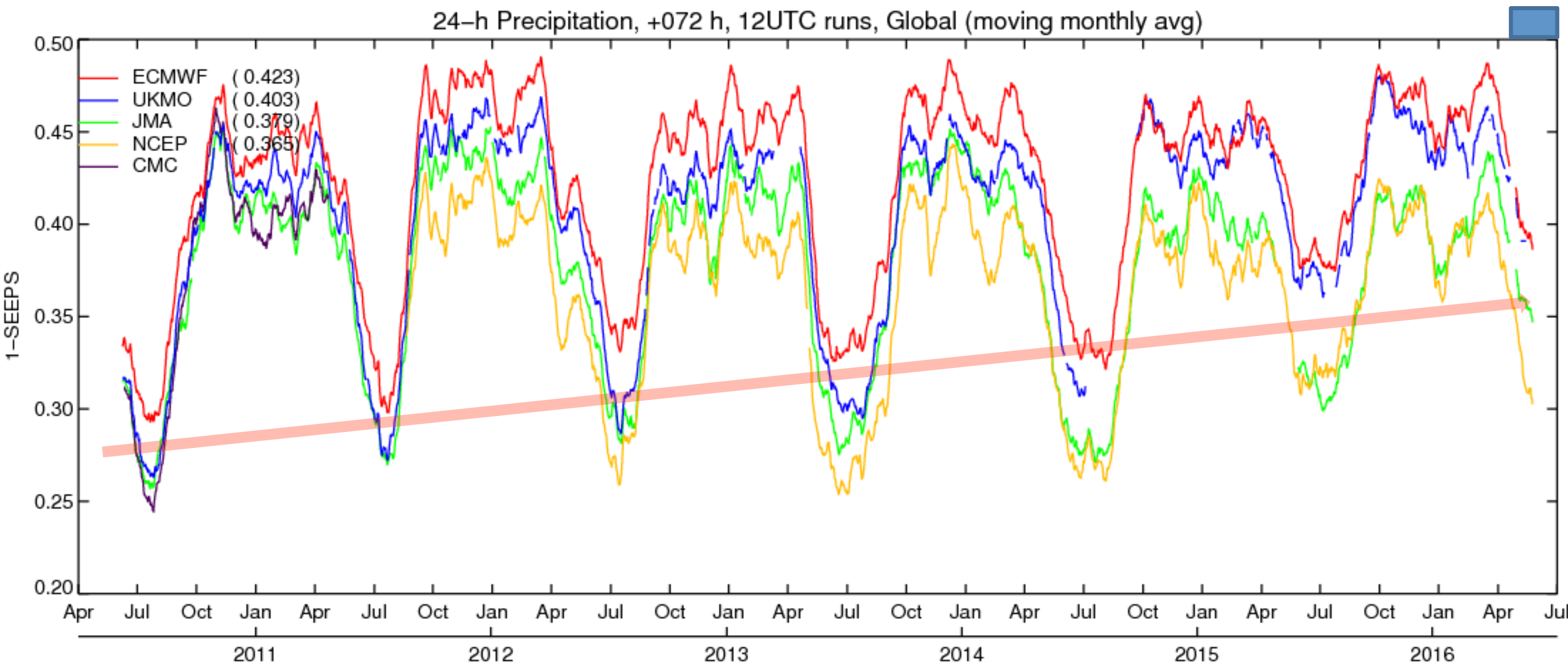




(From Thomas Haiden)

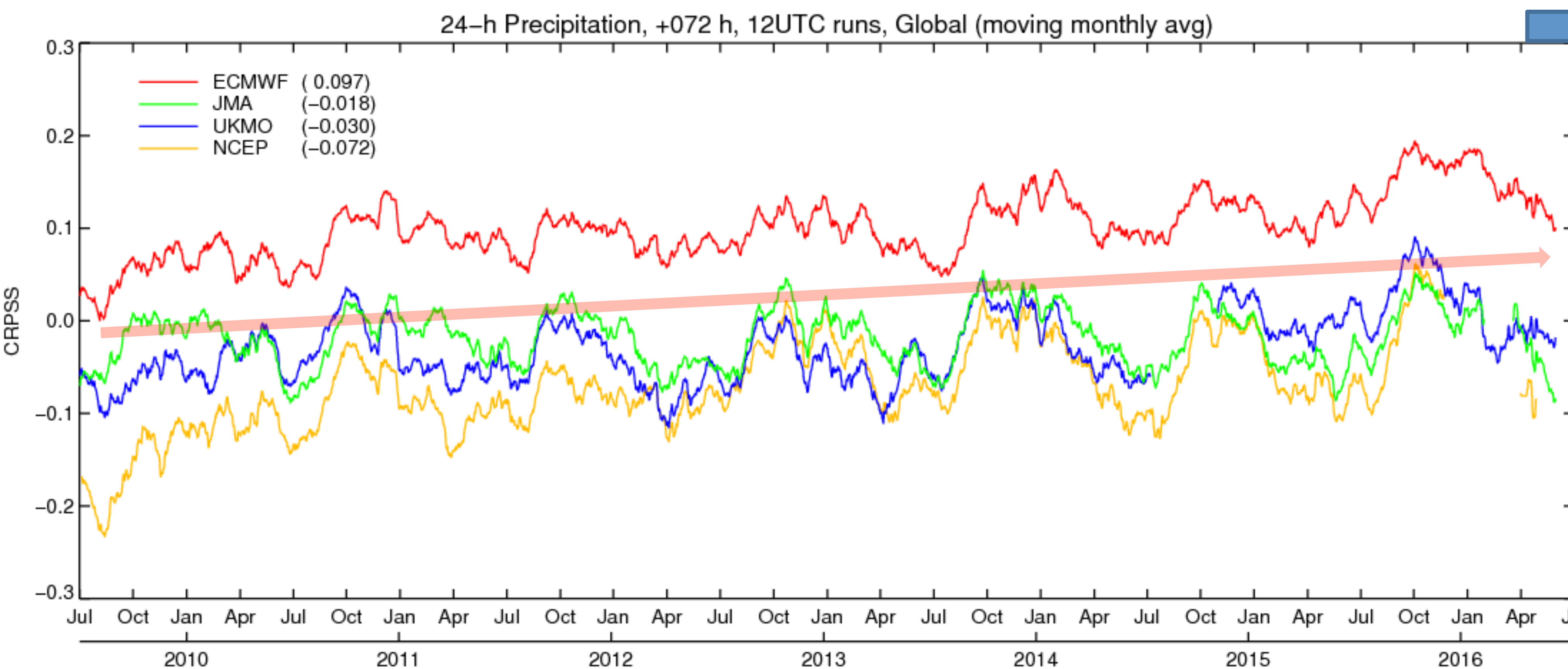
4. High-resolution single precipitation fc: t+72h

The accuracy of single high-resolution 72-h precipitation forecast (verified against observations), measured by the SEEPS score, for ECMWF and UKMO has been slowly improving.

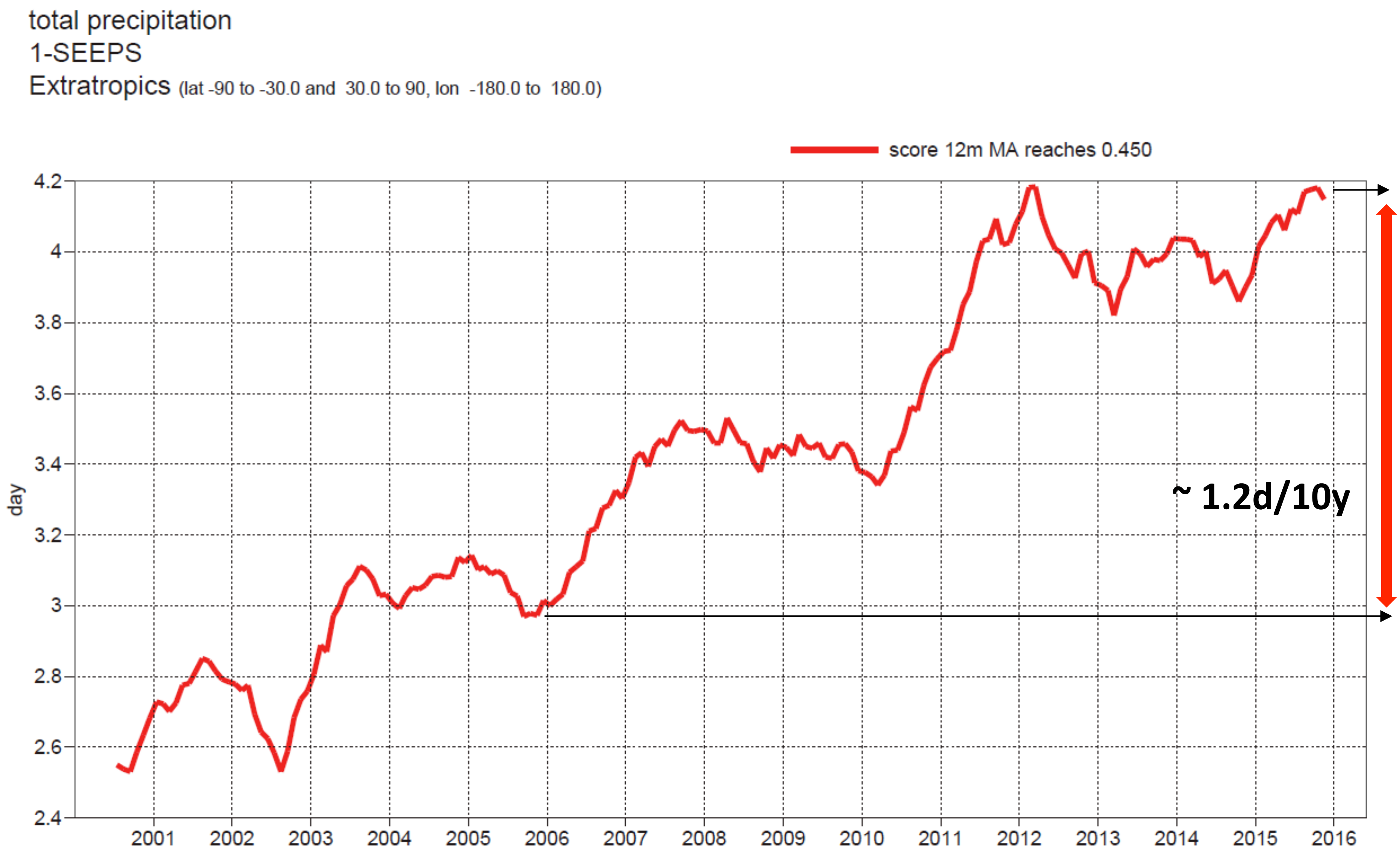


4. Impact of recent upgrade on single high-resolution fc

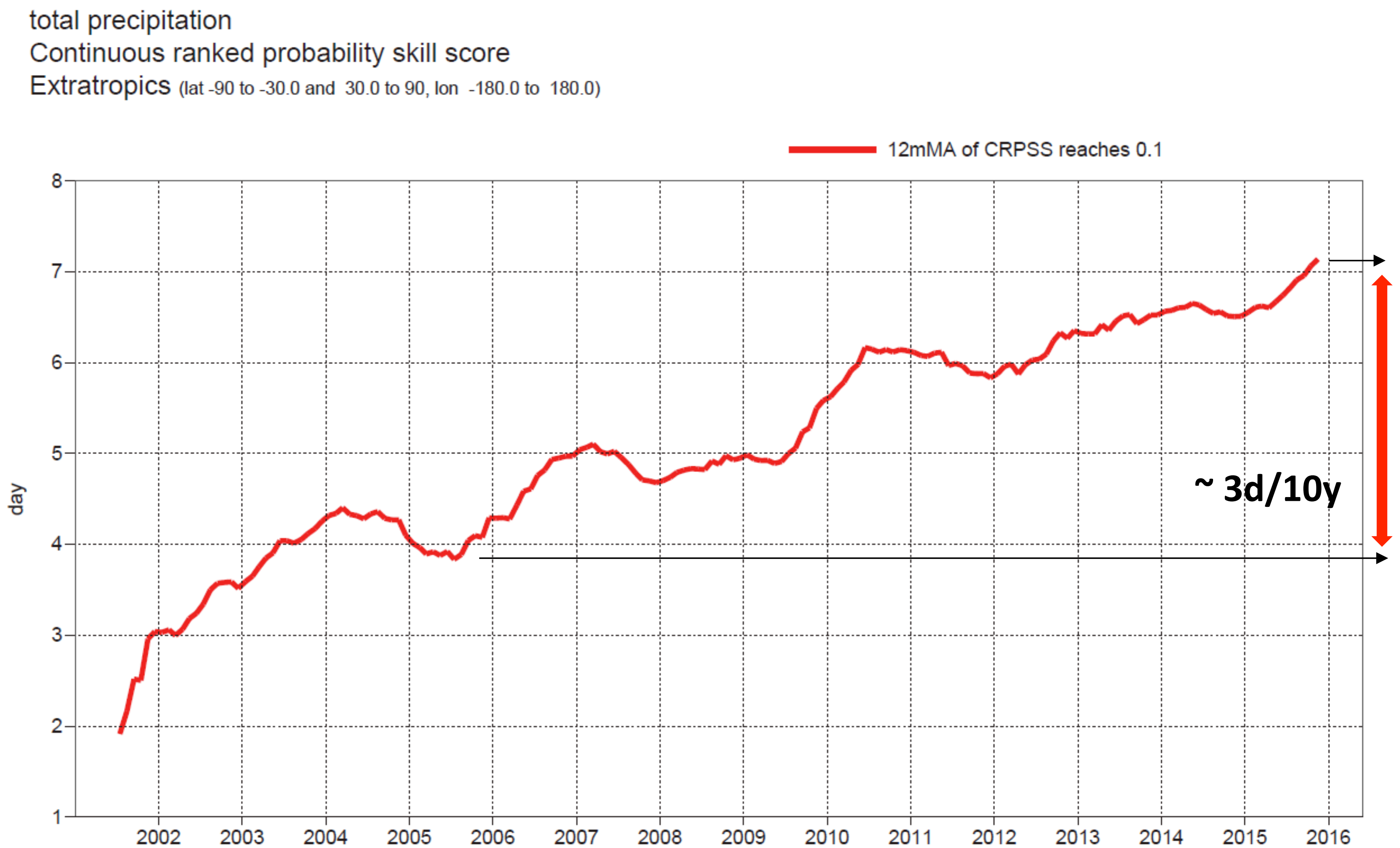
The accuracy of 72-h probabilistic precipitation forecast (verified against observations), measured by the CRPSS, has also been slowly improving.



4. ENS CRPSS – 24-h precipitation NH

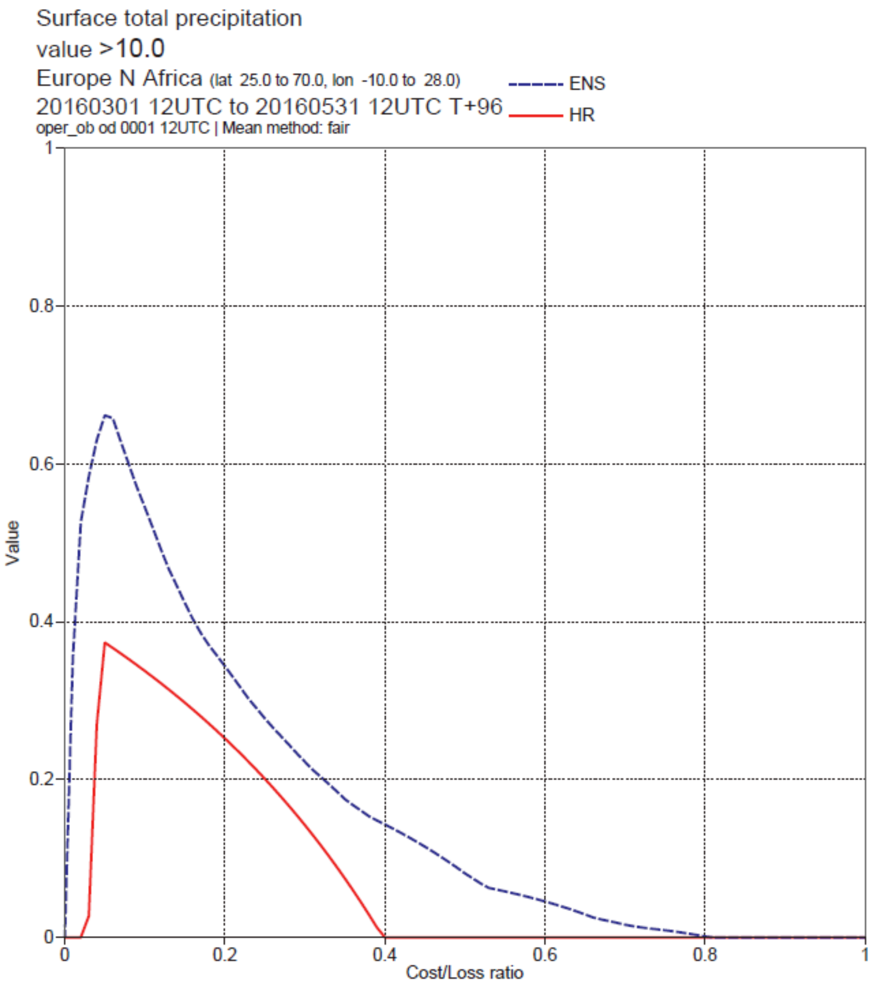


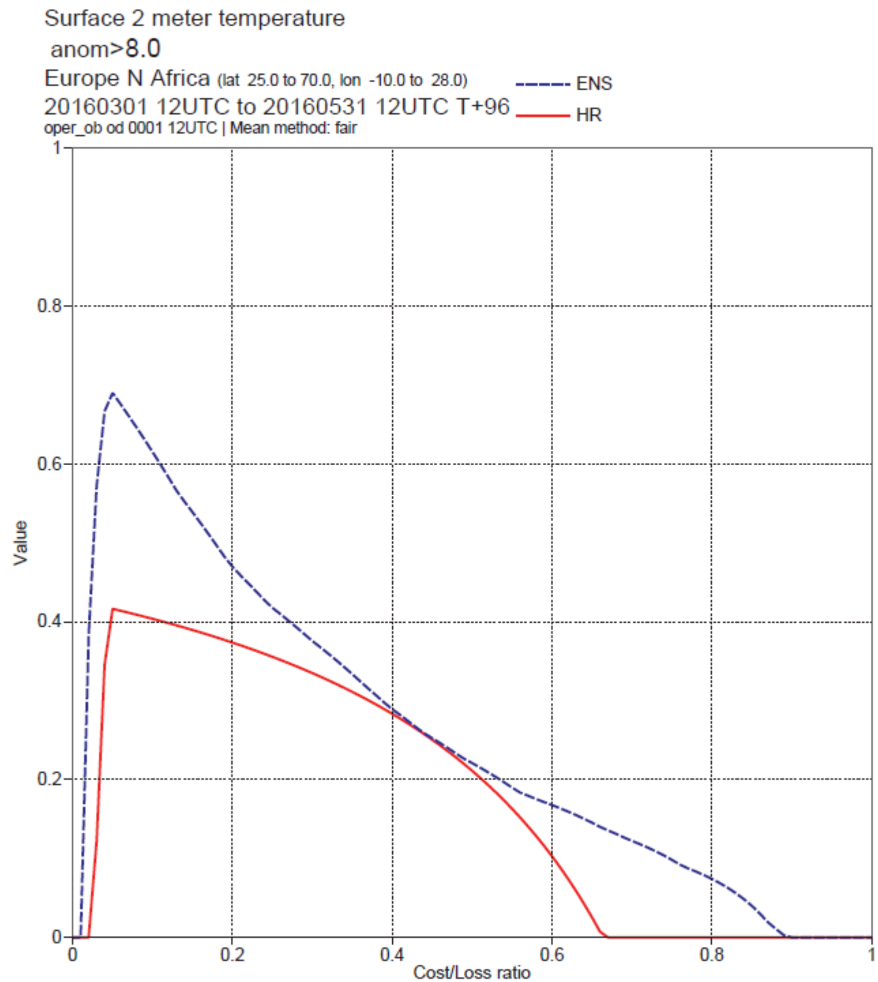
4. ENS CRPSS – 24-h precipitation NH



4. The Potential Economic Value: a cost-loss based metric

A cost-loss model can be used to assess the Potential Economic Value of single and probabilistic forecasts. Results indicate that these latter are more valuable.

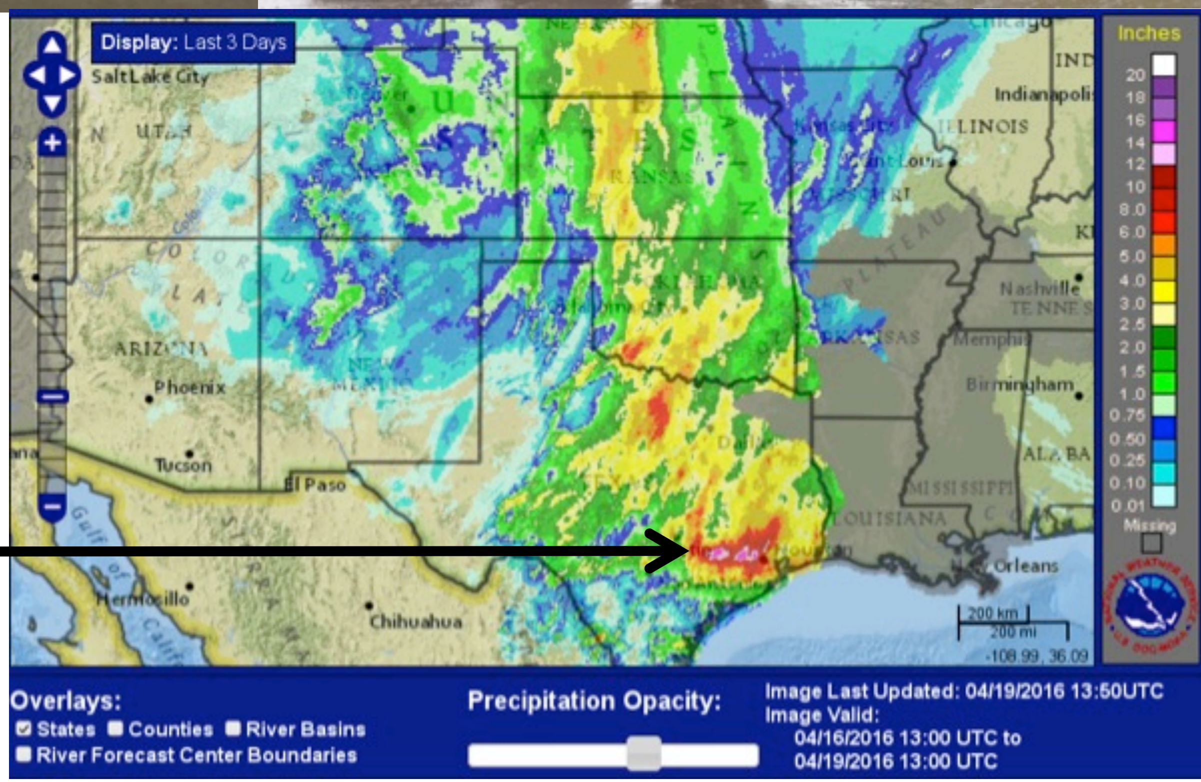




Plan

1. Why do we need ensembles?
2. The ECMWF ensembles
3. The resolution upgrade of March 2016
4. Few key performance indices
5. Prediction of an extreme precipitation event in Houston
6. Conclusions

6. Houston, Texas – April 2016: 16@12 to 19@12

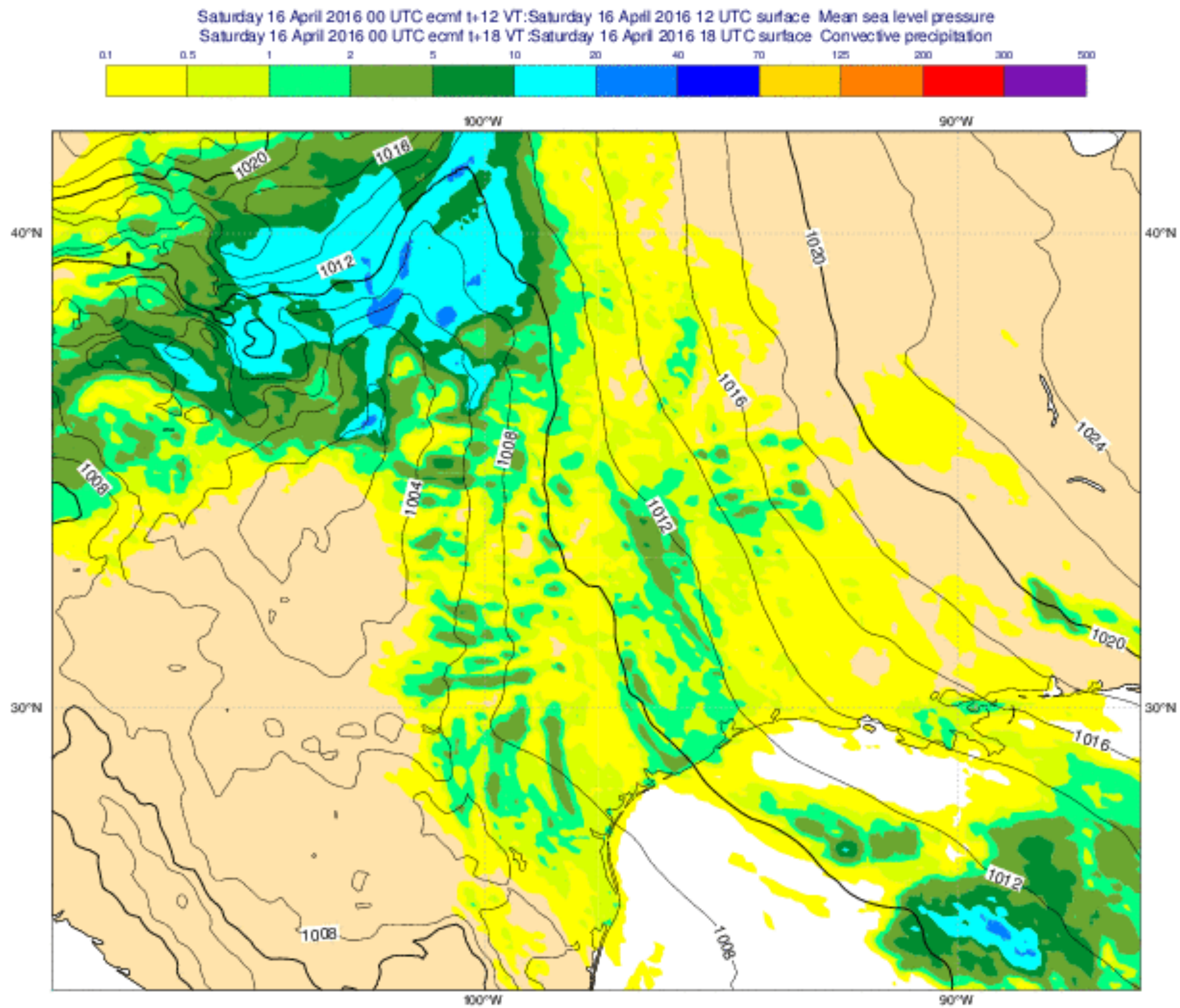


Up to 400-450 mm in mainly just over a day

(From Ervin Tsoter)

6. Houston, Texas, 16-19 April 2016

A very slow moving system bringing a continuous supply of warm and moist air from the Gulf.

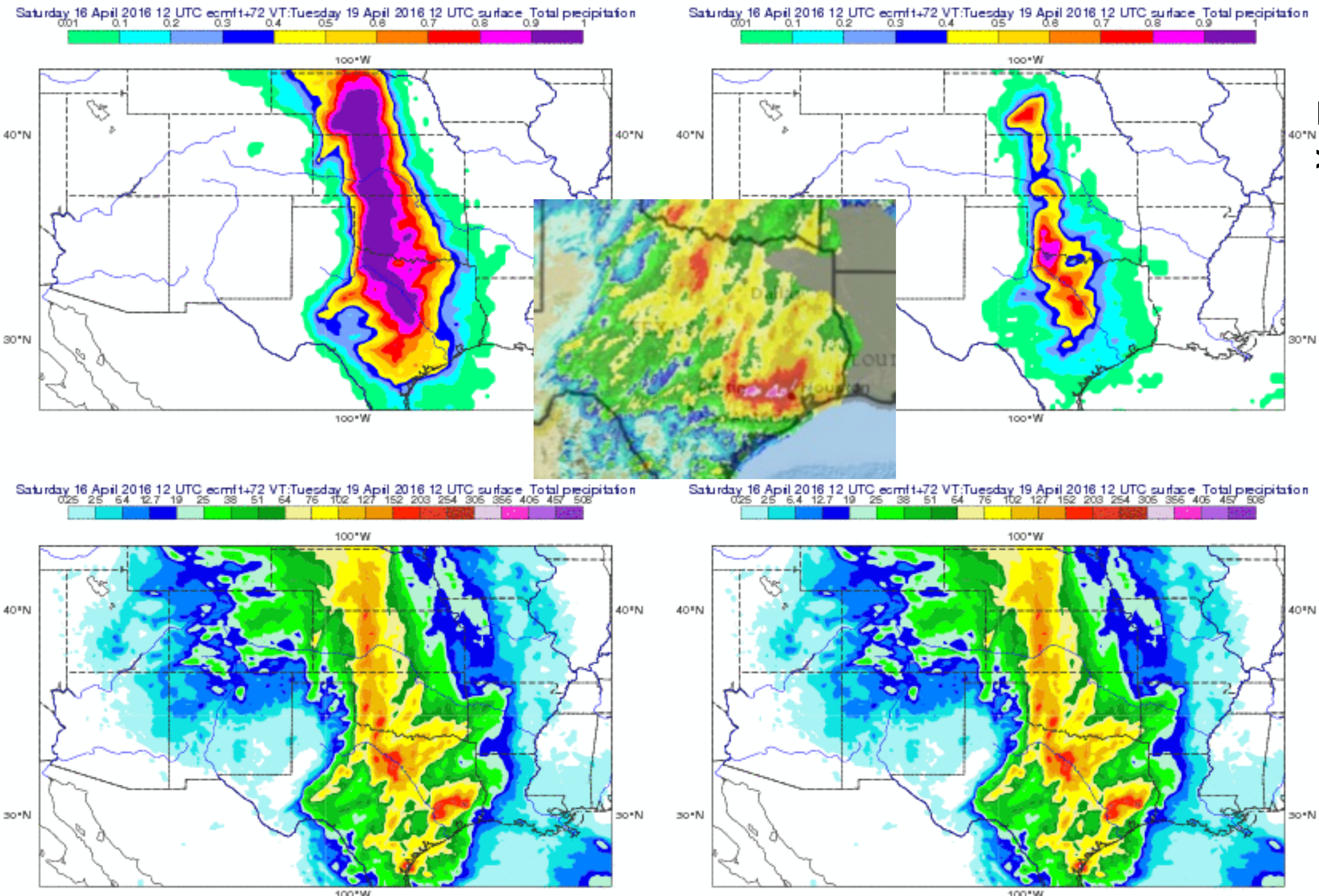


(From Ervin Tsoter)

6. +3d to +1d fcs of TP between 16@12 to 19@12

ENSprob > 60mm

ENSprob > 100mm

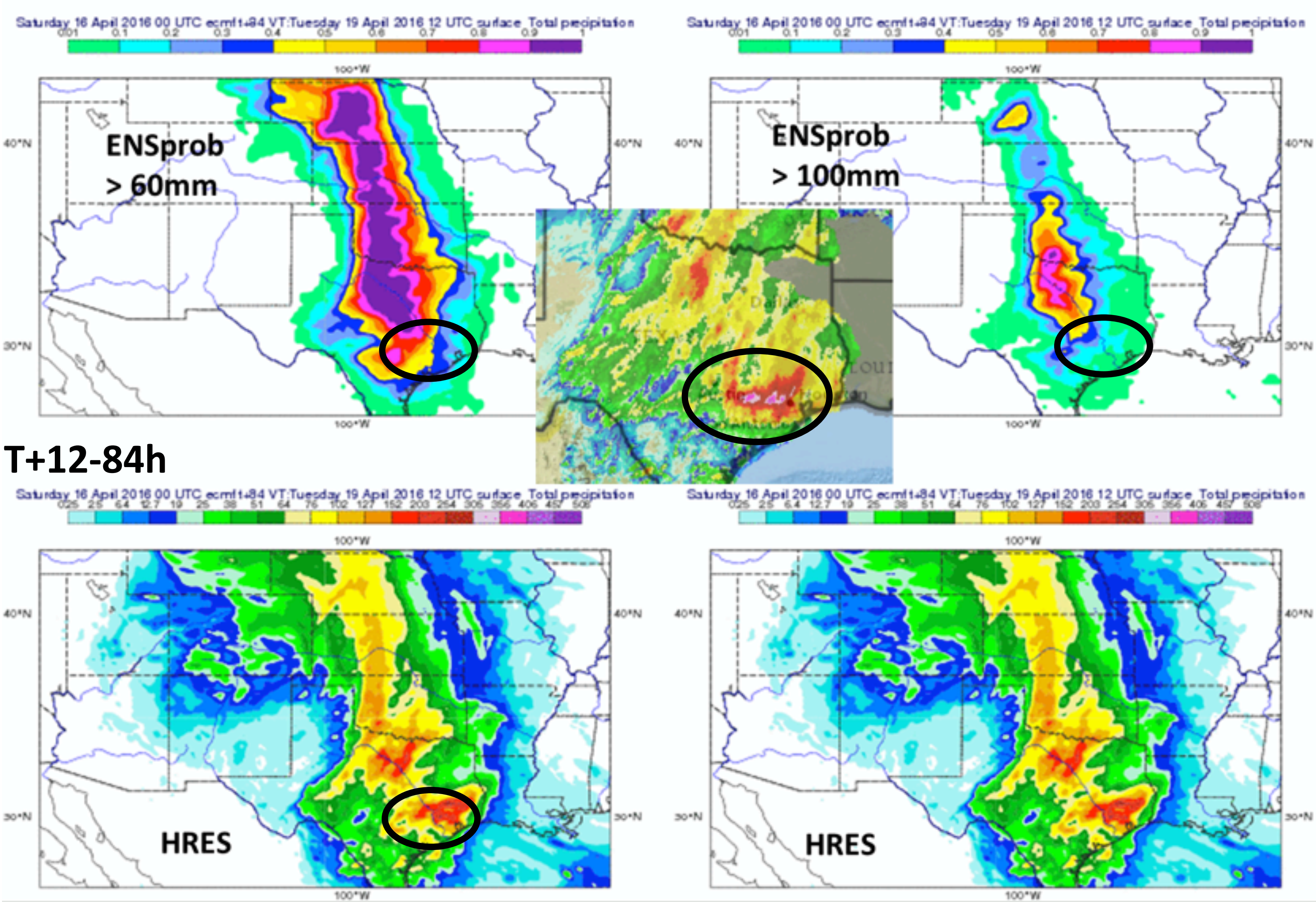


HRES

HRES

(From Ervin Tsoter)

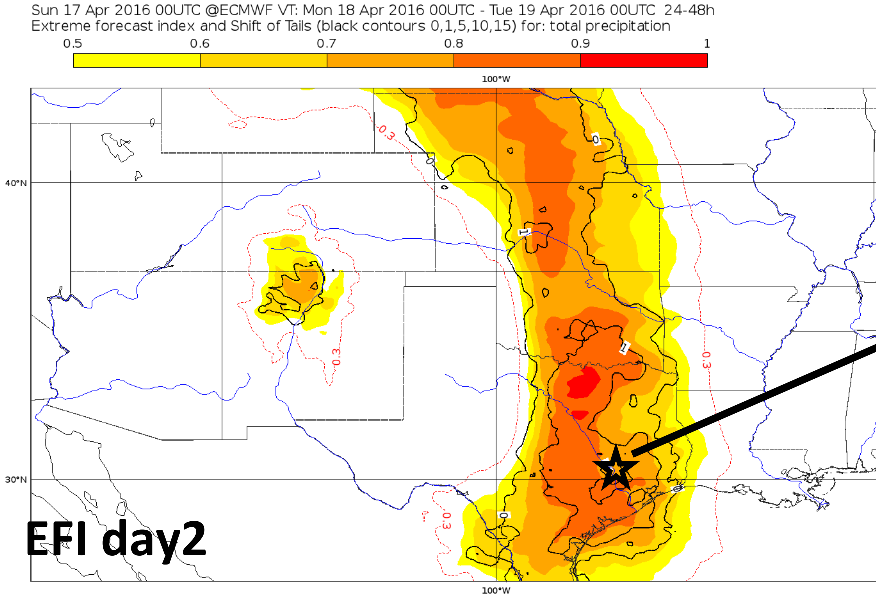
6. +3.5d fcs of TP between 16@12 to 19@12

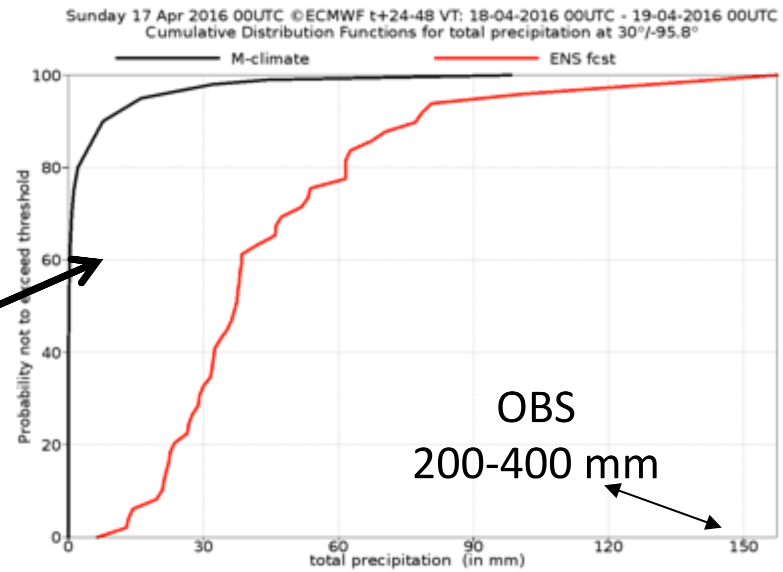


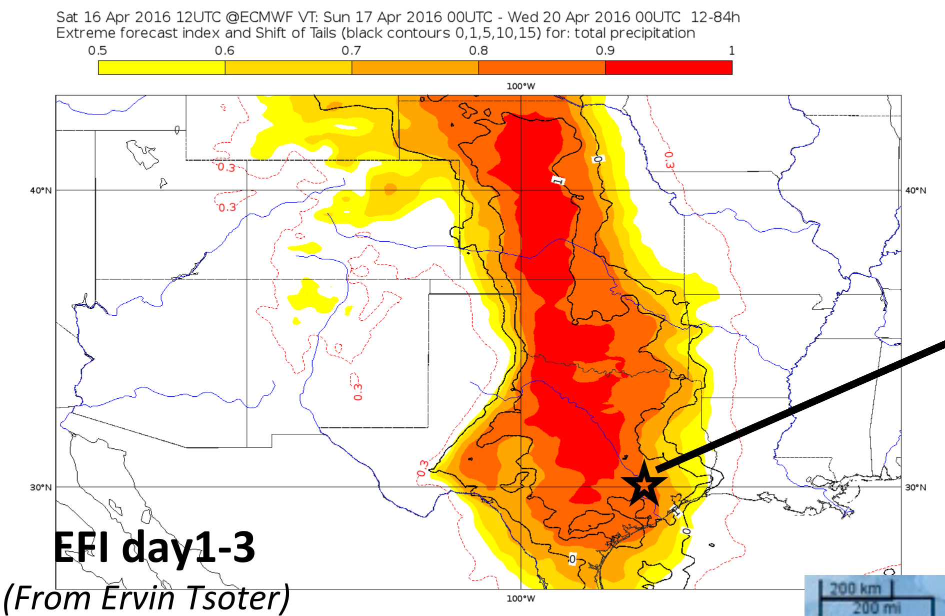
16th 00z T+12-84h

(From Ervin Tsoter)

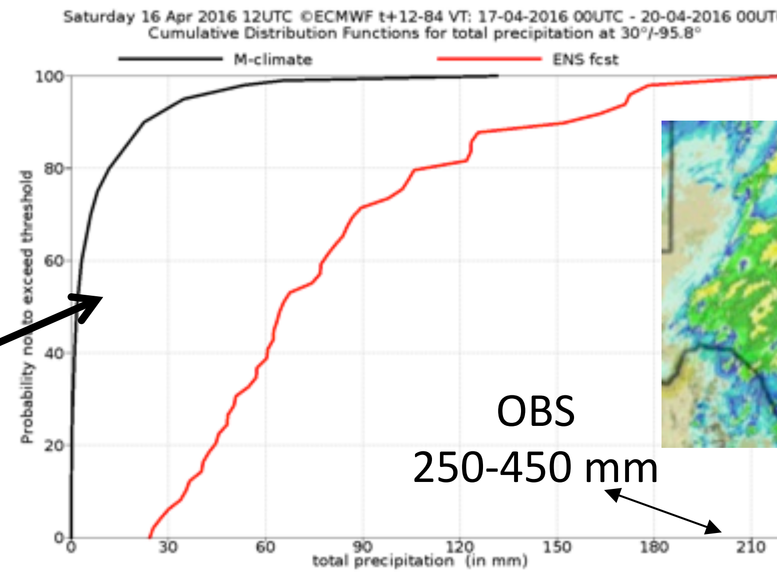
6. +2d and +3d fcs of TP between 16@12 to 19@12

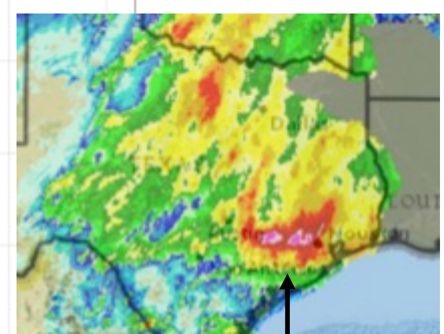






(From Ervin Tsoter)





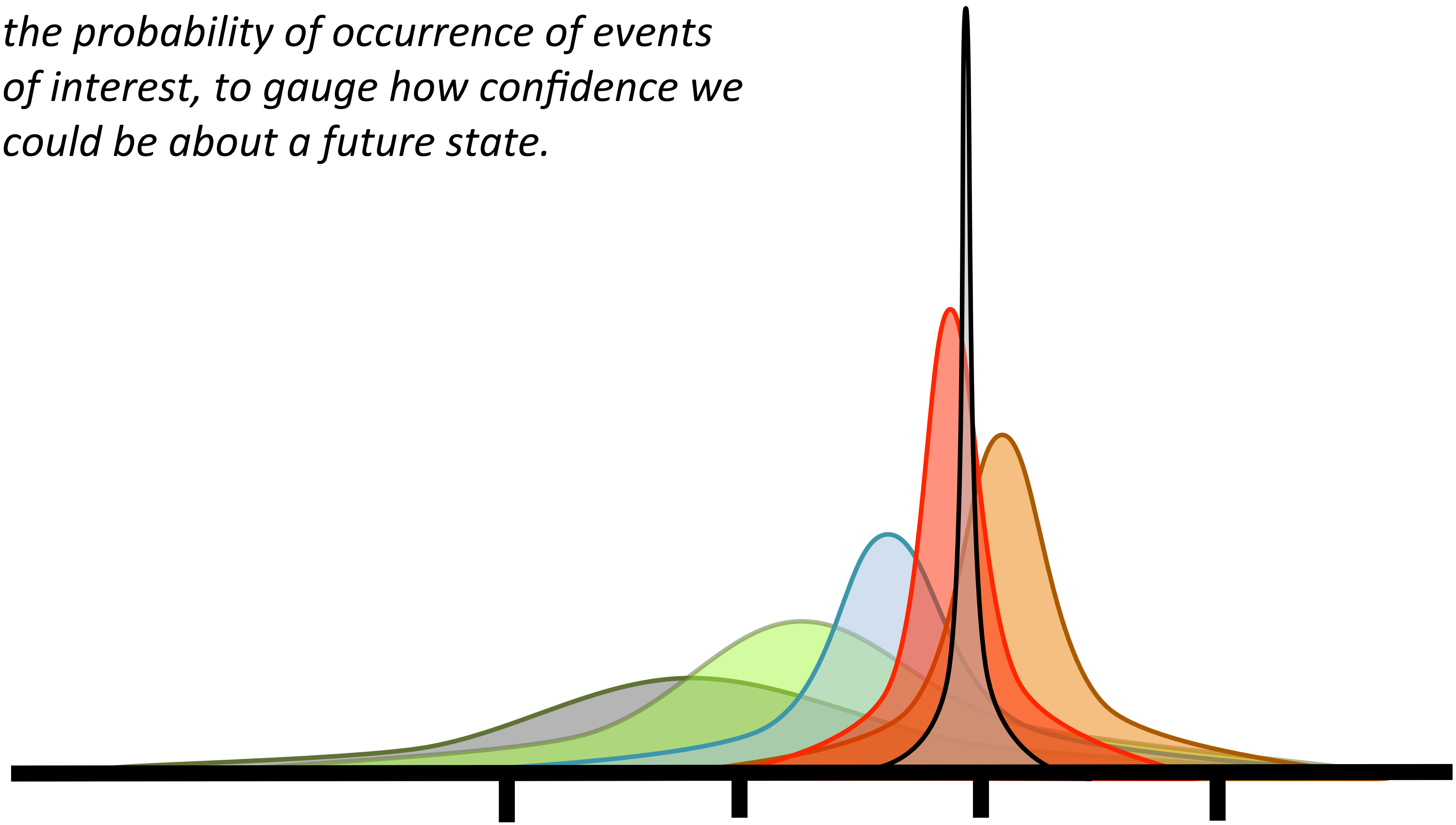
Red >150
Purple >300

Plan

1. Why do we need ensembles?
2. The ECMWF ensembles
3. The resolution upgrade of March 2016
4. Few key performance indices
5. Prediction of an extreme precipitation event in Houston
6. Conclusions

6. Ensembles are the cornerstone of our future

Because they can be used to estimate the probability of occurrence of events of interest, to gauge how confidence we could be about a future state.



6. 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

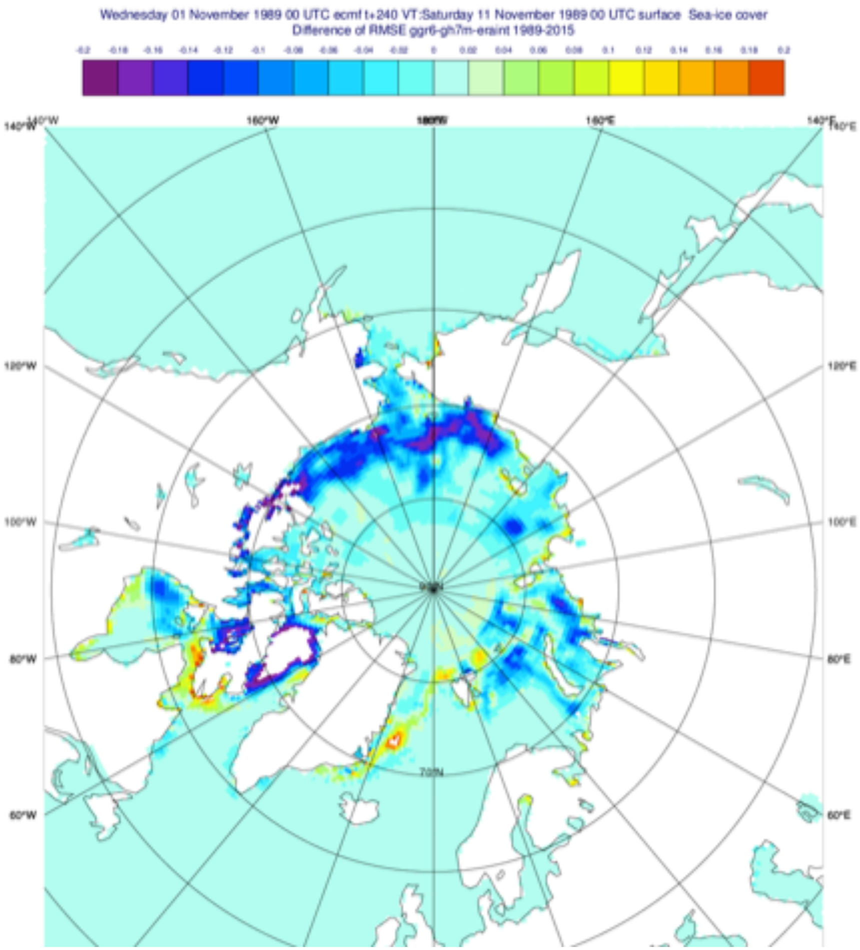
2. **Initial Conditions** estimation: integrate further the analysis and forecast ensembles (EDA/ORA and ENS) and re-assess the potential benefit of starting ENS directly from EDA analyses; assess the impact of using a more strongly coupled DA

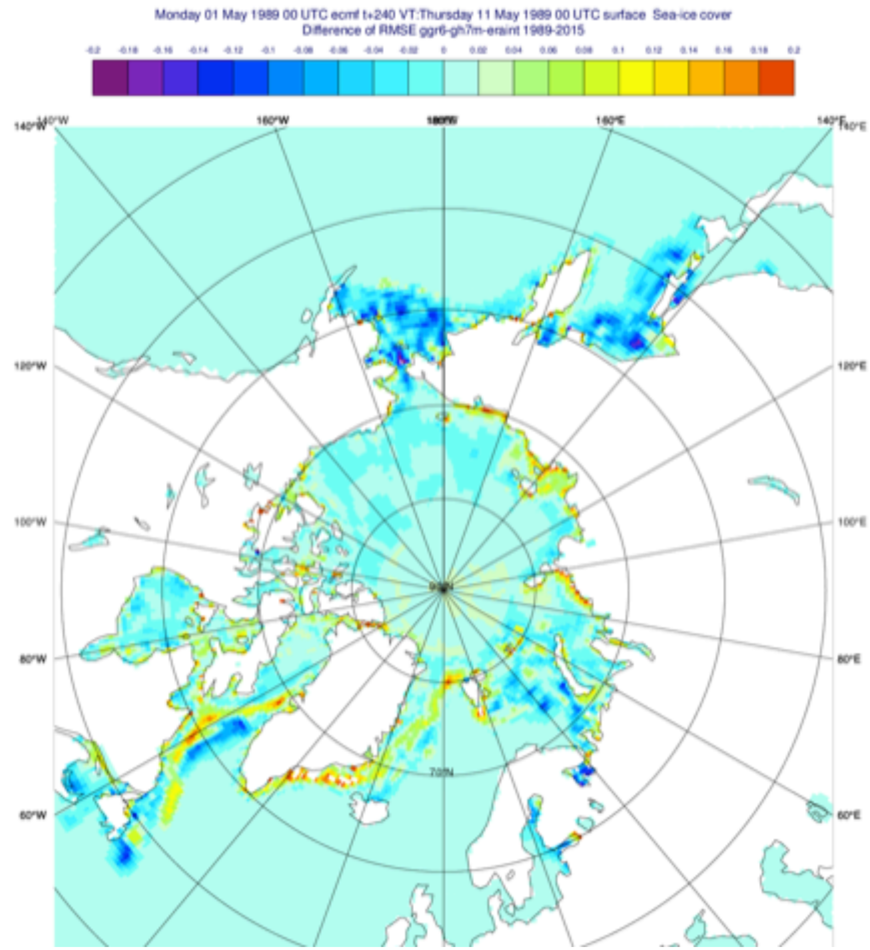
3. **Predictability:** identify sources of predictability, and ways to extract predictable signals for the ensemble PDF

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

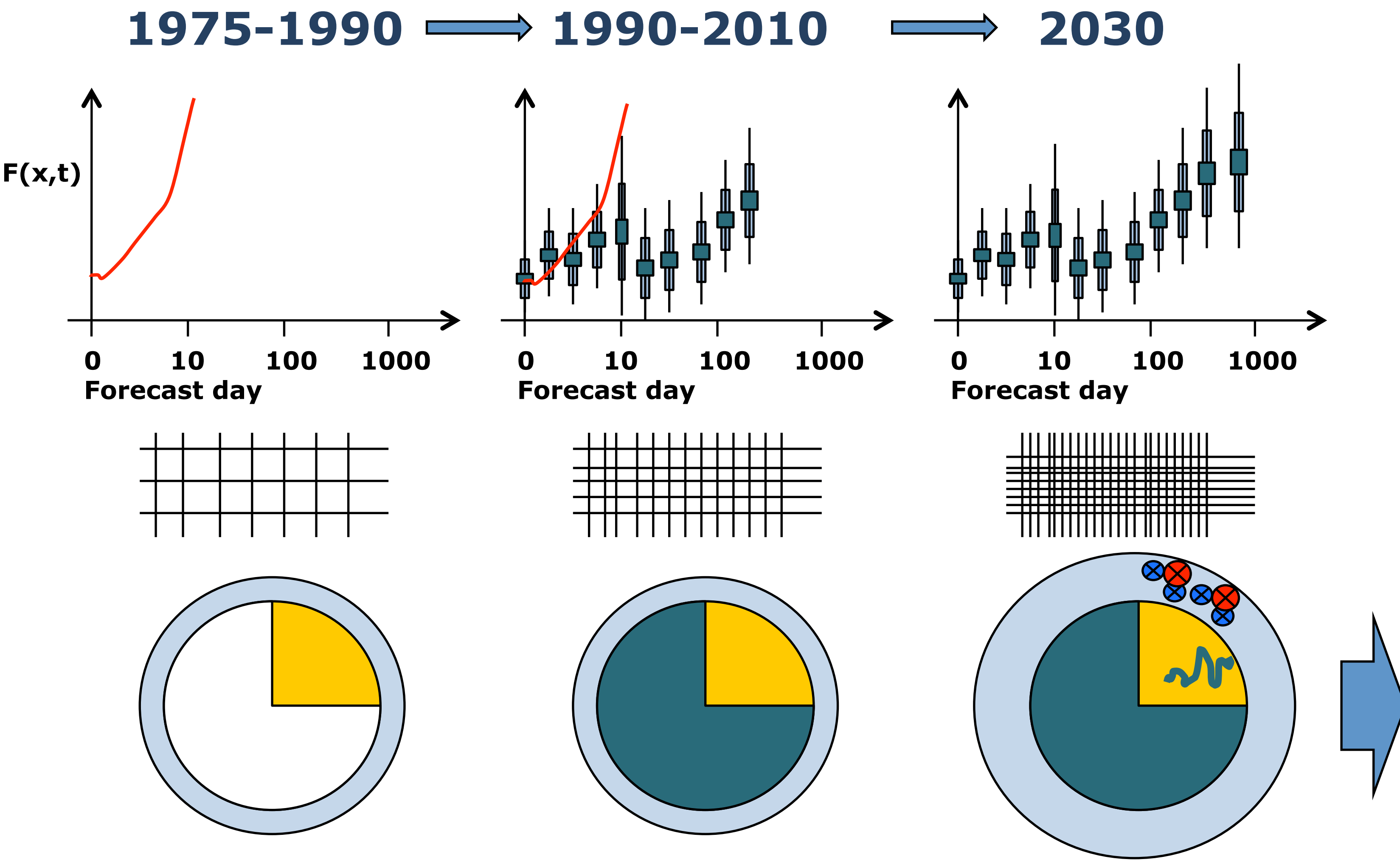
6. Q4-16: ENS with higher-res ocean (0.25°-z75) & sea-ice

The next model cycle with include a higher-resolution ocean (0.25° degrees with 75 layers) with dynamical sea-ice in ENS, with ICs from ORAS5 ensemble of analyses.



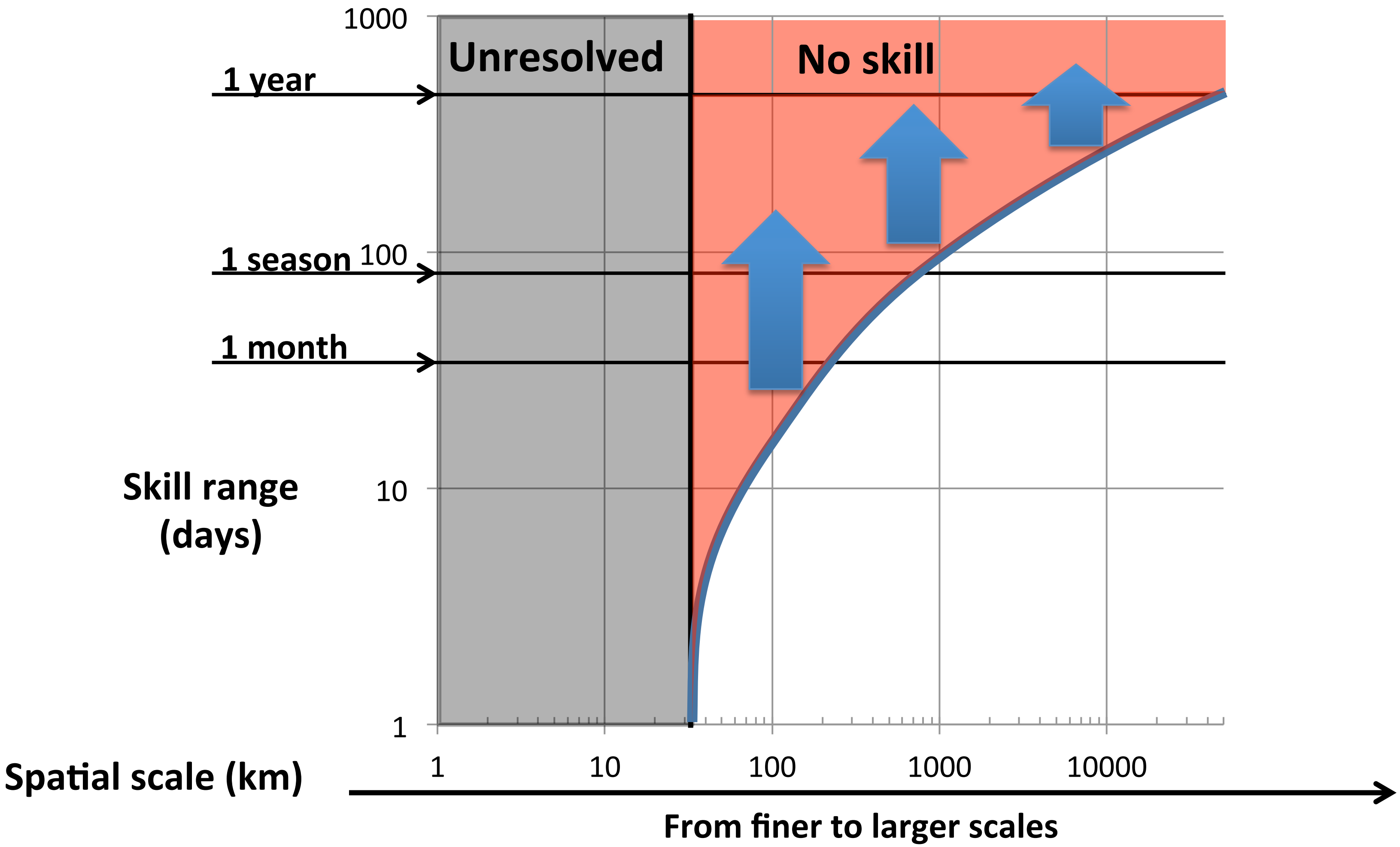


6. 10y strategy: ensemble of Earth-system model and DA



We aim to include all relevant Earth-system processes in our analyses and forecast ensembles.

6. So that we can further extend the fc skill horizon



Thank you very much for your attention …

ECMWF Annual Seminar 5-8 Sept 2016: Earth-system …