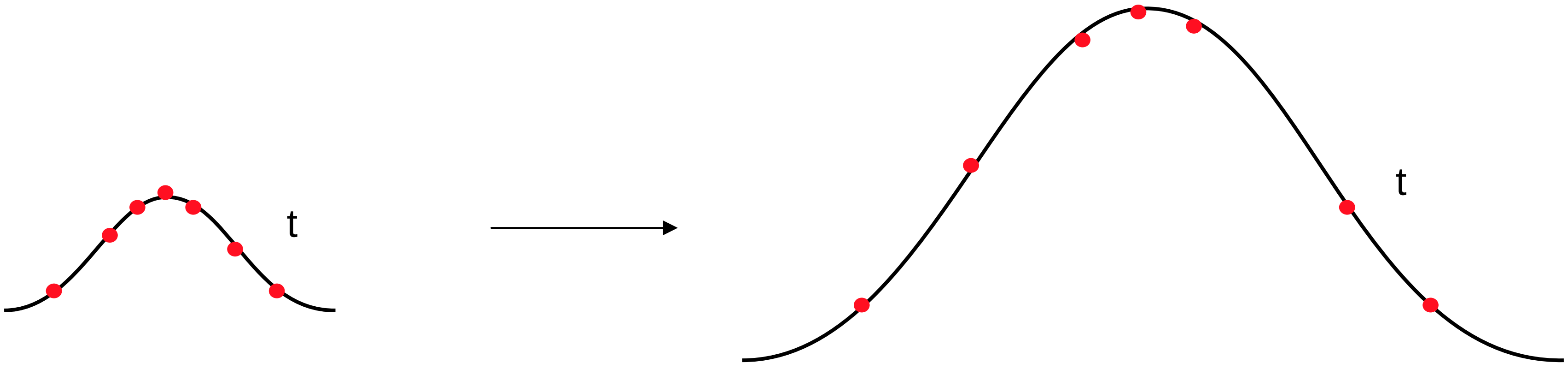
Is Anders Right?

With thanks to Roberto Buizza, Renate Hagedorn, Martin Leutbecher, Andy Lawrence, Lenny Smith

Probability Distribution of Truth



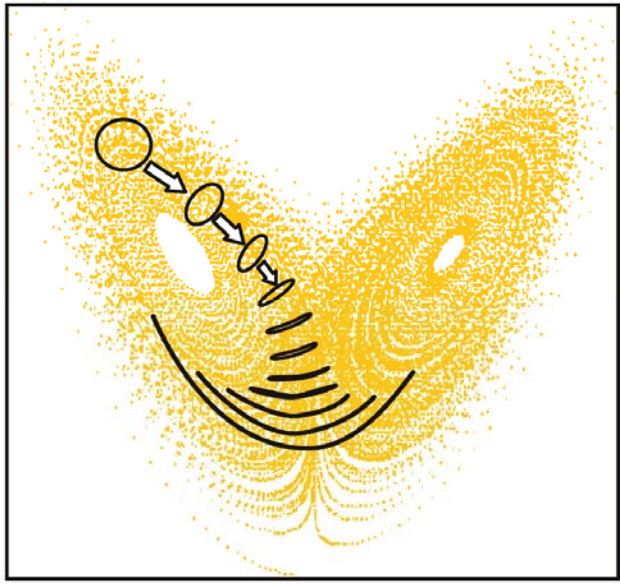
$\rho(X, t_0)$ $\qquad$ $\rho(X, t)$

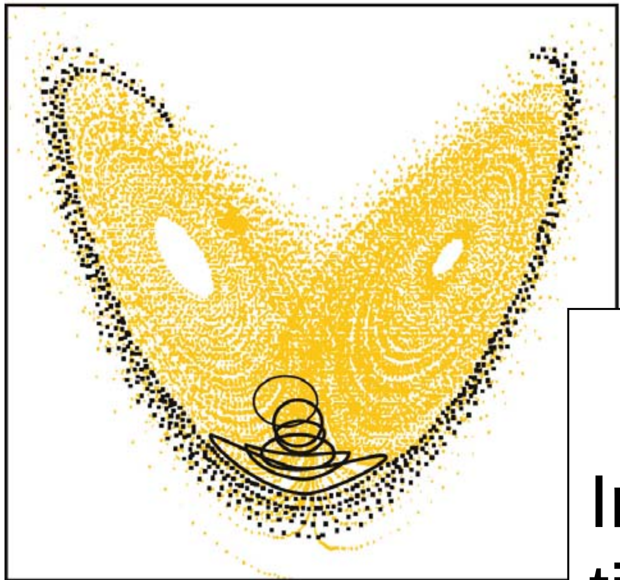
$$\rho(X,t) = \Lambda(t,t_0)\rho(X',t_0)$$

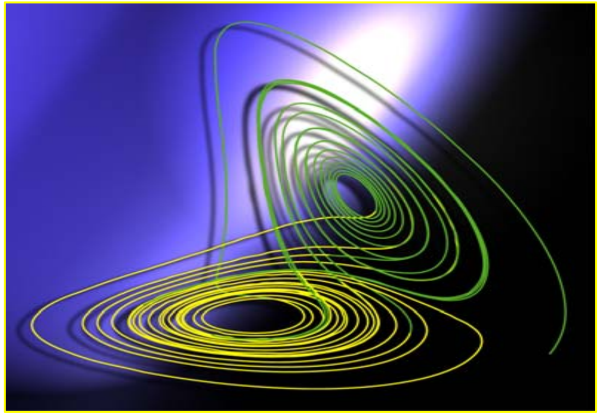
The **Liouville equation** (Ehrendorfer, 2006). In practice solved using ensemble prediction techniques

A perfect EPS is a random drawing from $\rho(X,t)$

ECMWF



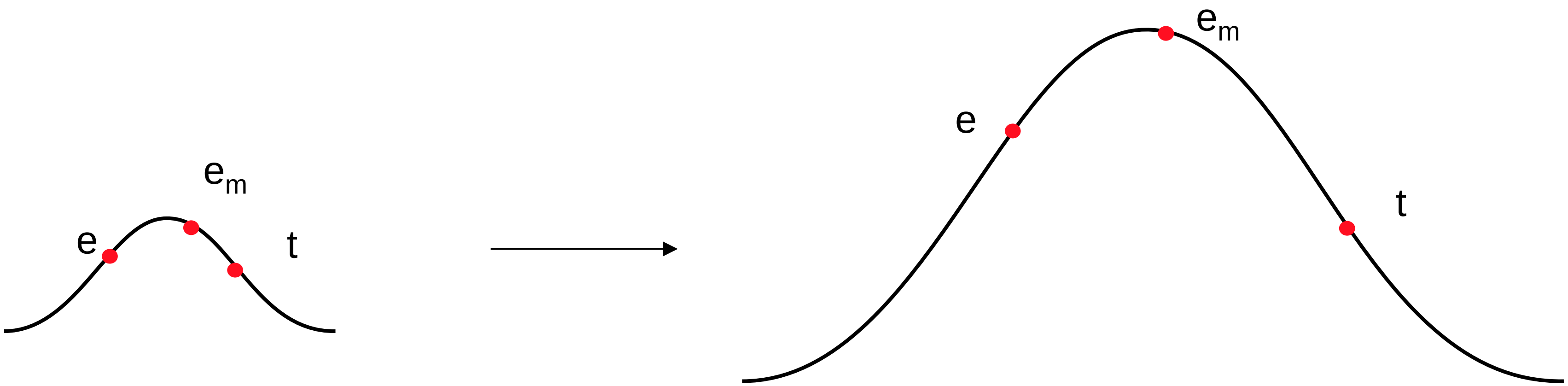




Scientific Basis for Ensemble Prediction

In a nonlinear dynamical system, the finite-time growth of initial uncertainties is flow dependent.

In a perfect EPS…



$\rho(X, t_0)$ $\rho(X, t)$

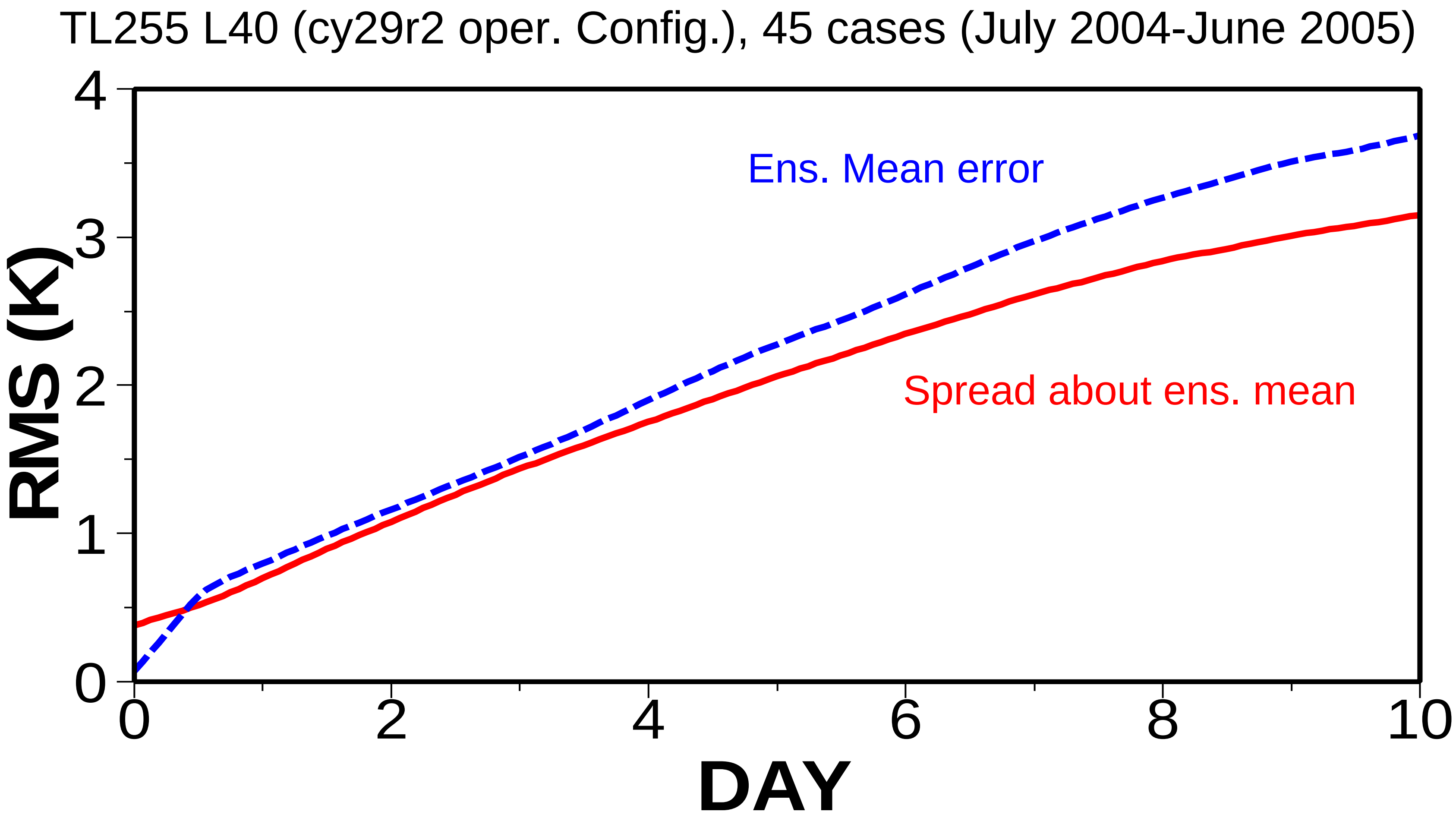
..“**e**” and “**t**” are drawings from the same underlying probability distribution. Therefore “**e**” and “**t**” have the same expectation values. In particular:

$$\Rightarrow \quad \|e - e_m\| = \|t - e_m\|$$

Implies the spread should match the skill for a good EPS!!

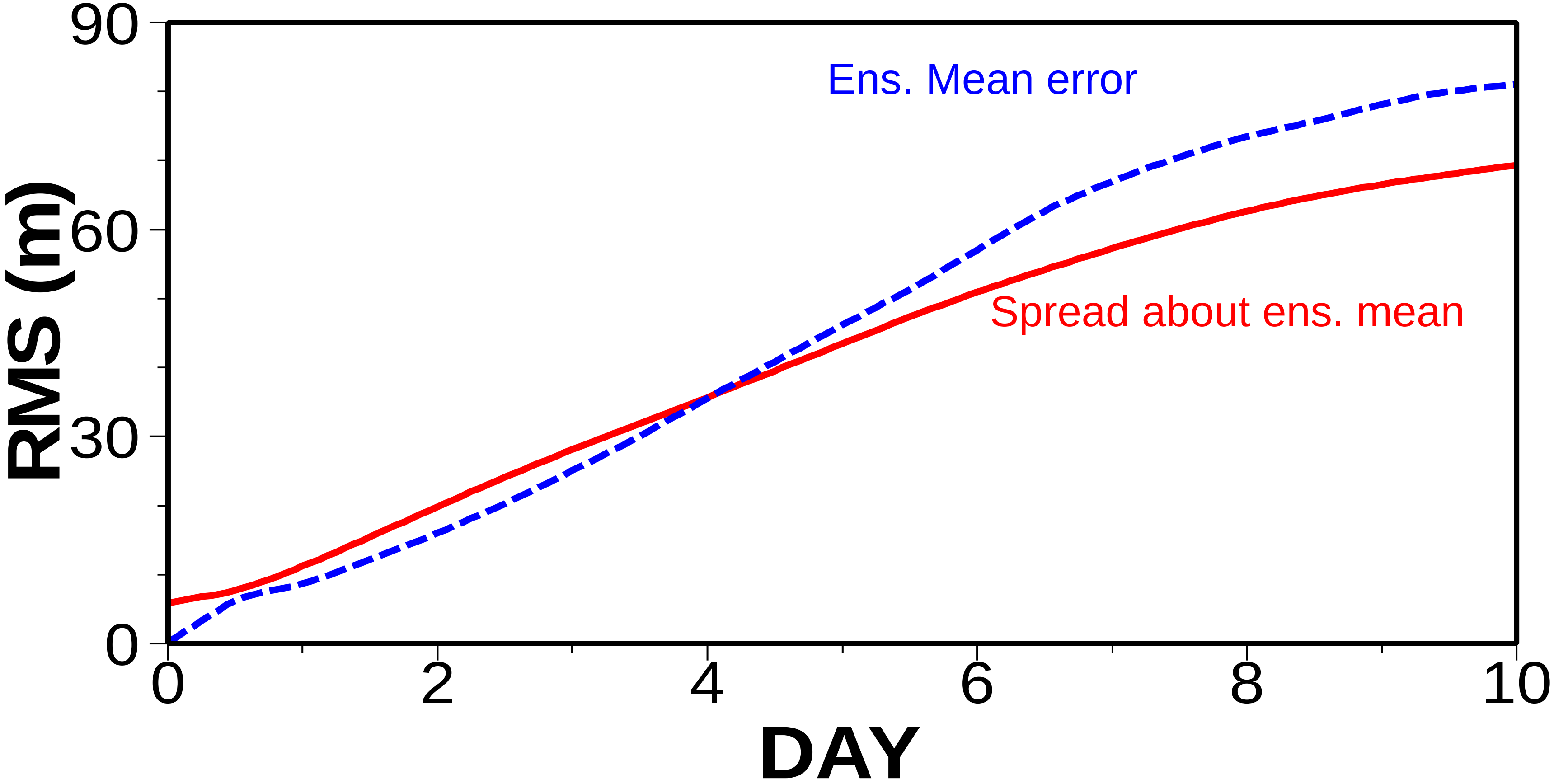
ECMWF

Spread and ens mean error, N-Hem T850



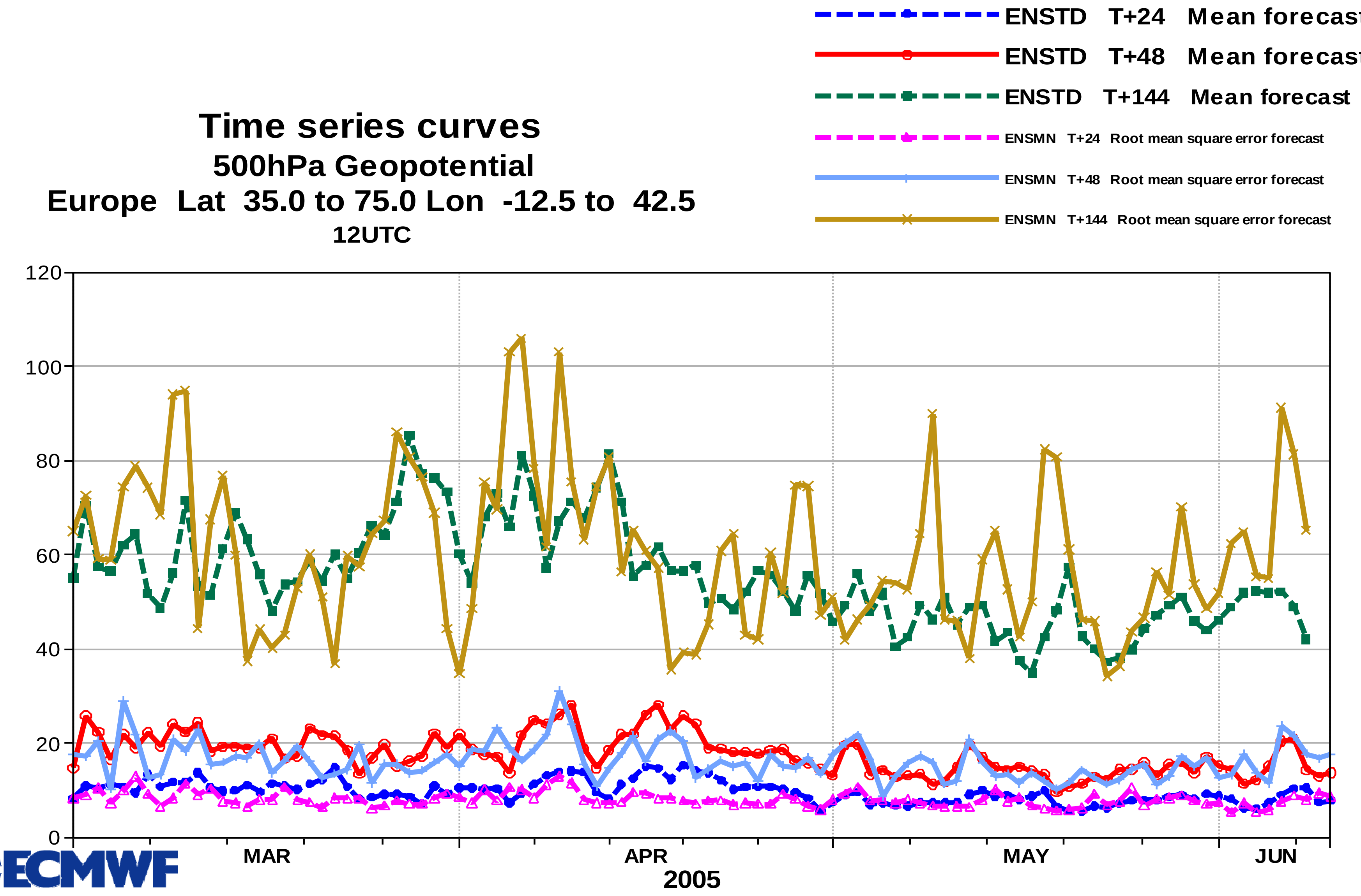
Spread and ens mean error, N-Hem Z500

TL255 L40 (cy29r2 oper. Config.), 45 cases (July 2004-June 2005)

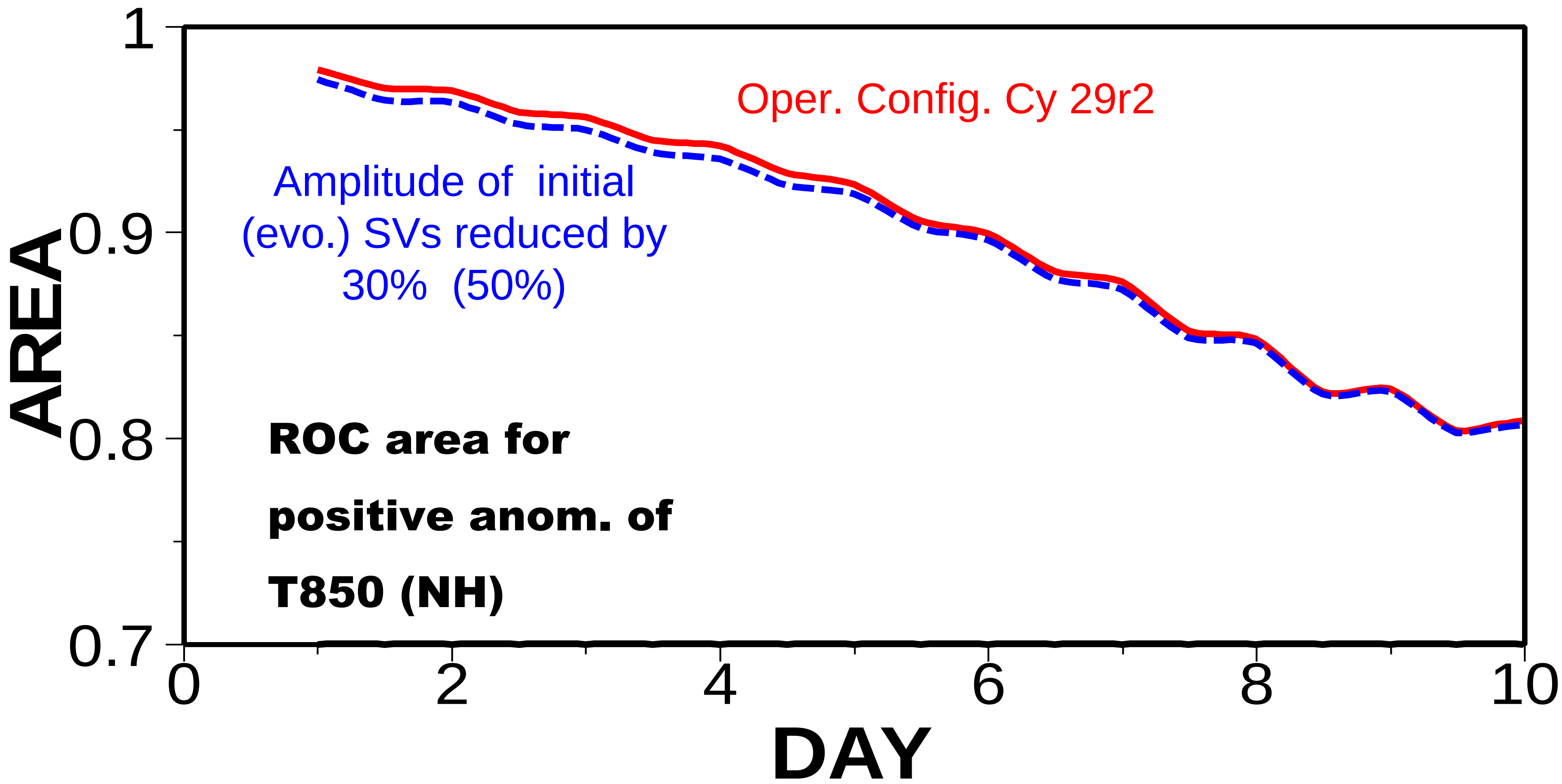


ECMWF

EPS spread/Error

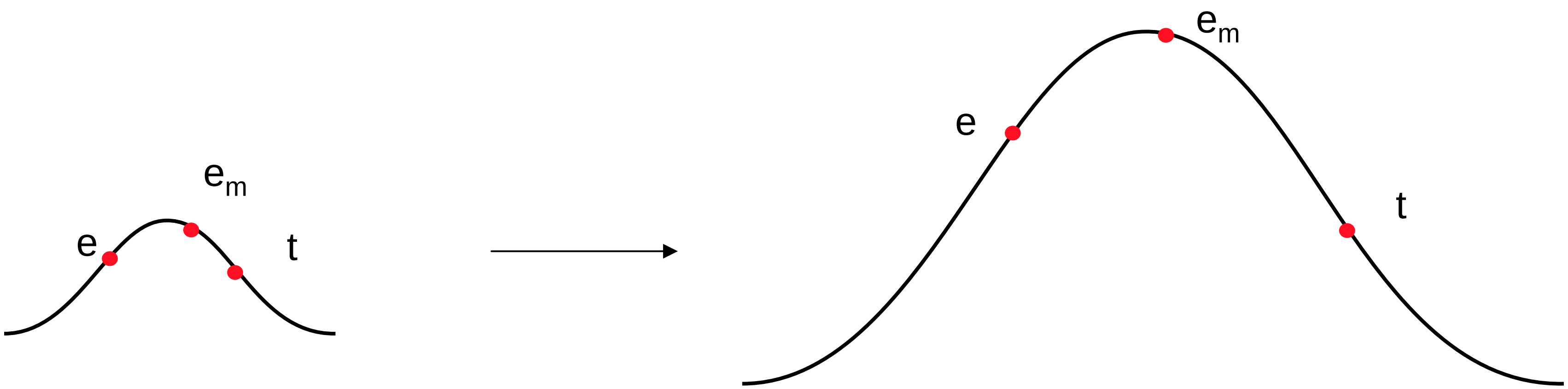


Reduce spread ⇒ Reduce Skill



29 cases in April/May 2005, both experiments cycle 29r2

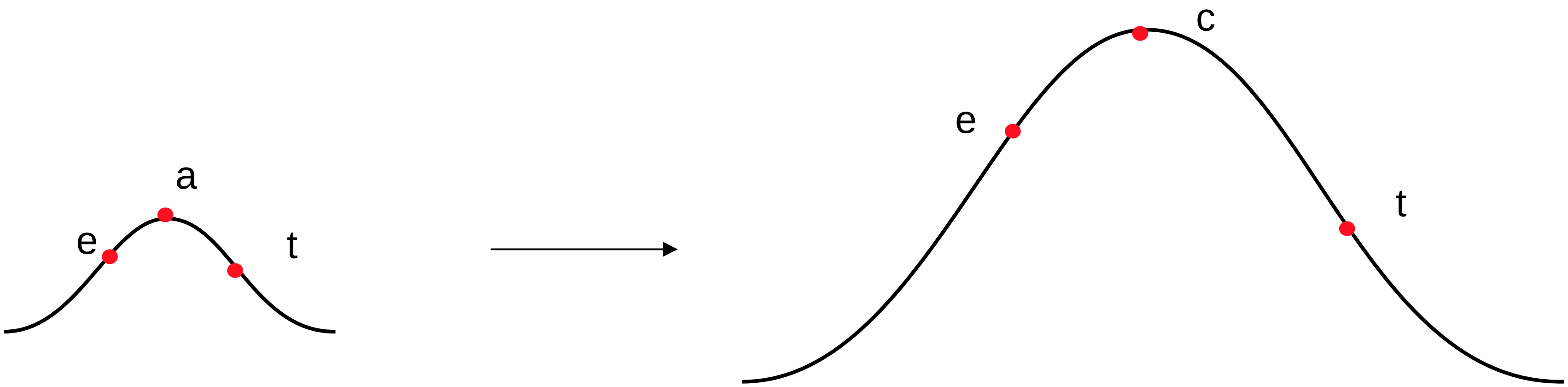
A Perfect EPS



$\rho(X, t_0)$ $\rho(X, t)$

$$\Rightarrow \quad \|e - t\| = \sqrt{2}\,\|e_m - t\|$$

In the short range..



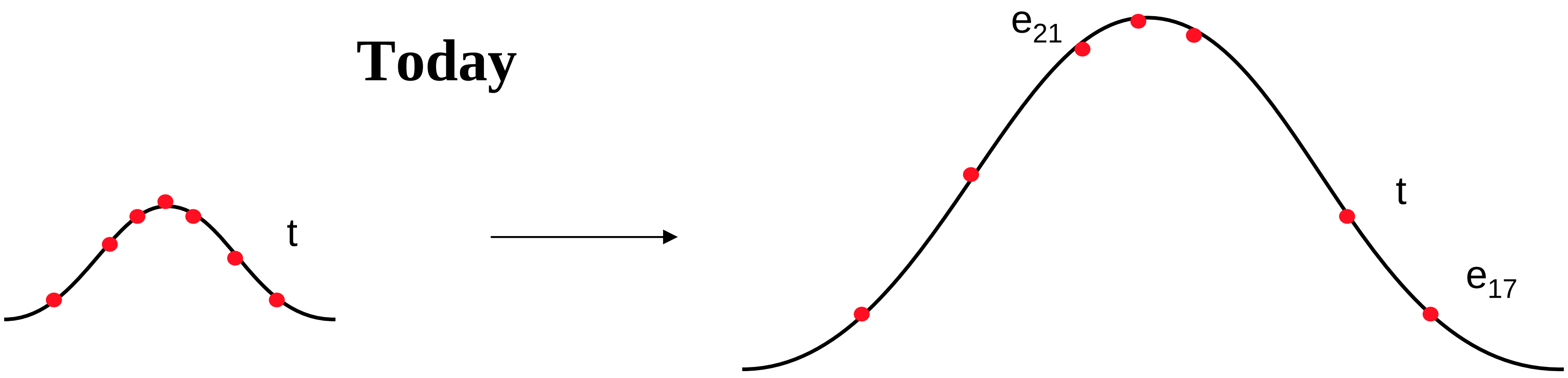
$\rho(X, t_0)$ $\rho(X, t)$

$$\|e - t\| = \sqrt{2}\,\|e_m - t\|$$

$$\Rightarrow \quad \|e - t\| = \sqrt{2}\,\|c - t\|$$

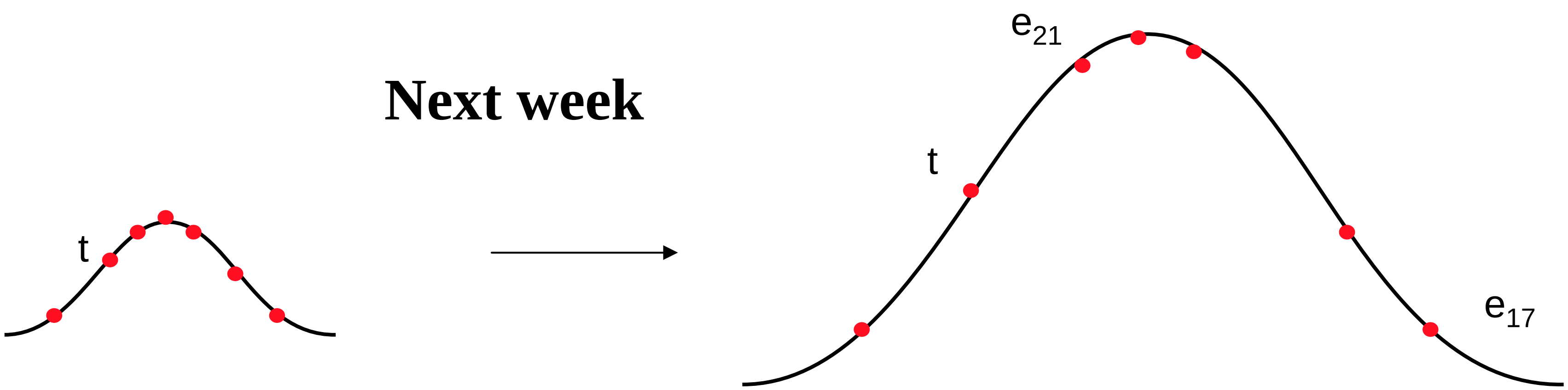
Probability Distribution of Truth

Today



$\rho(X, t_0)$ $\qquad$ $\rho(X, t)$

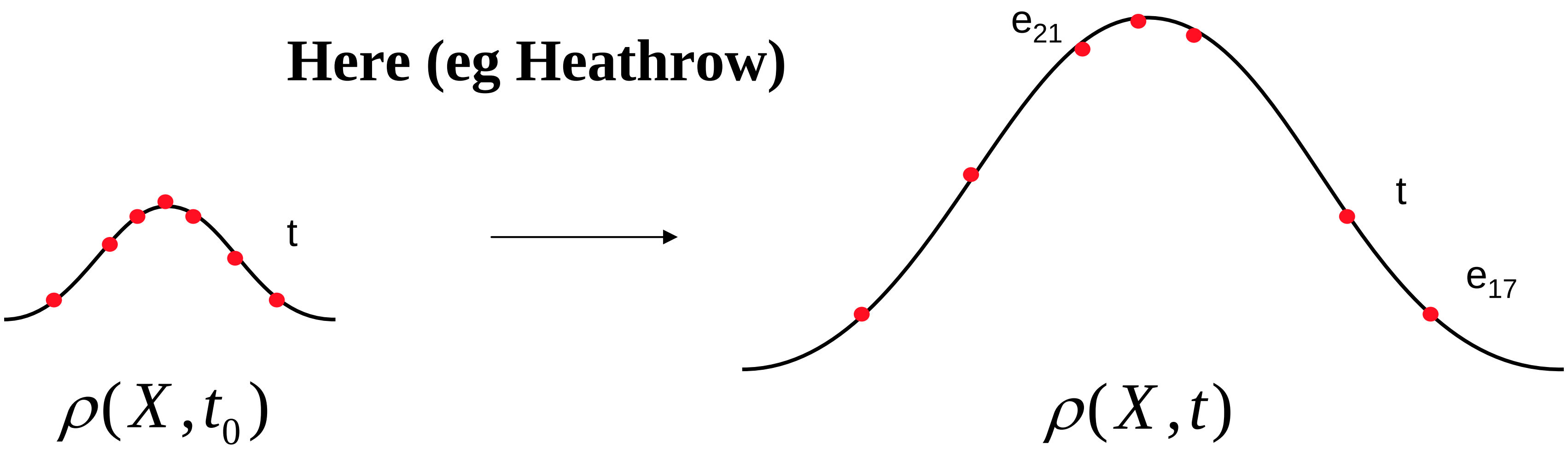
Next week



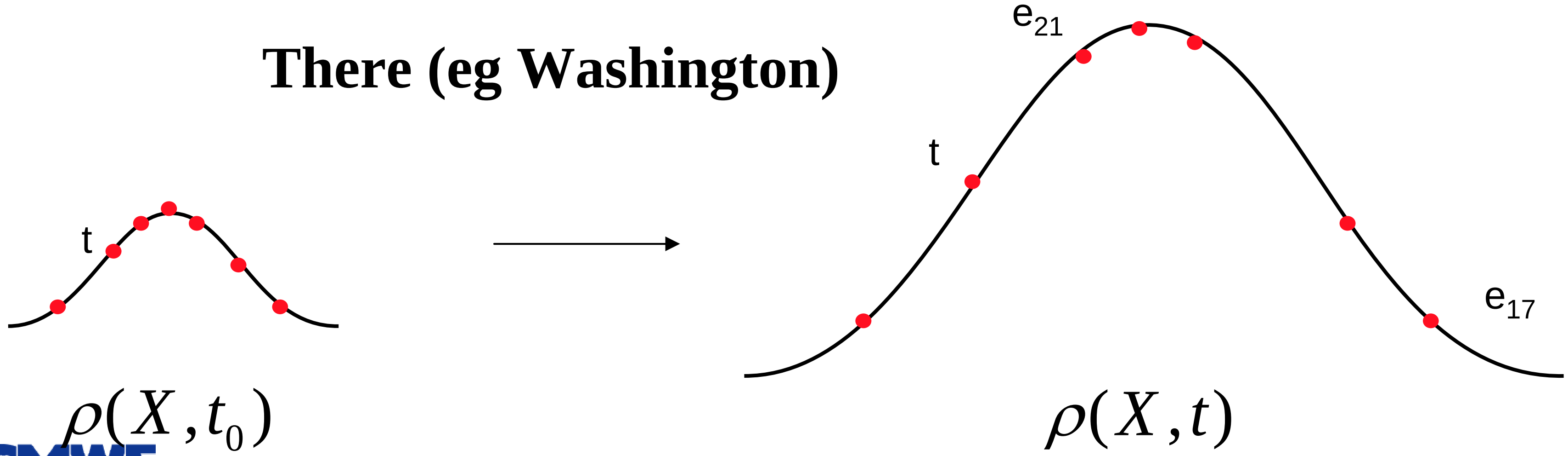
$\rho(X, t_0)$ $\qquad$ $\rho(X, t)$

Probability Distribution of Truth

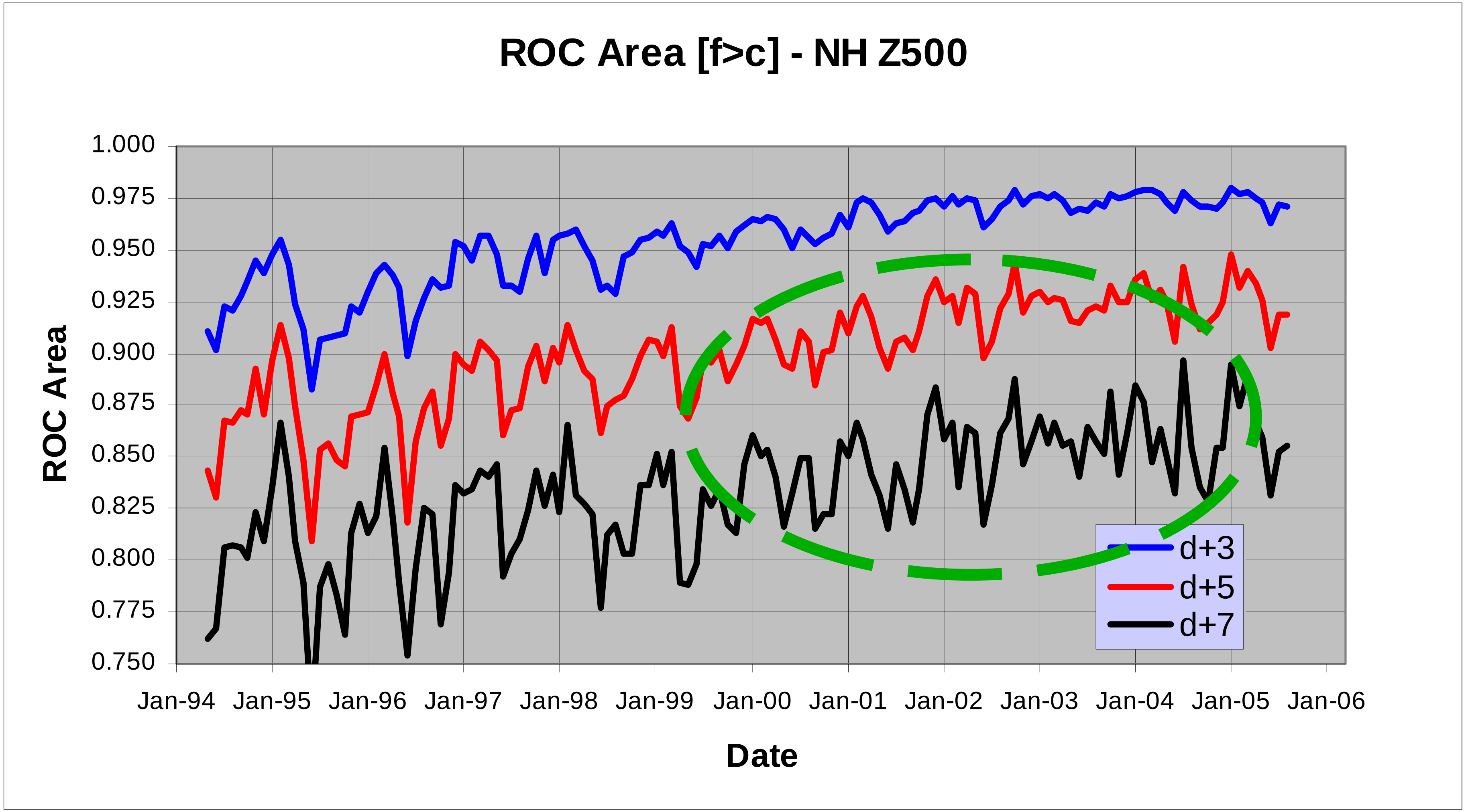
Here (eg Heathrow)



There (eg Washington)

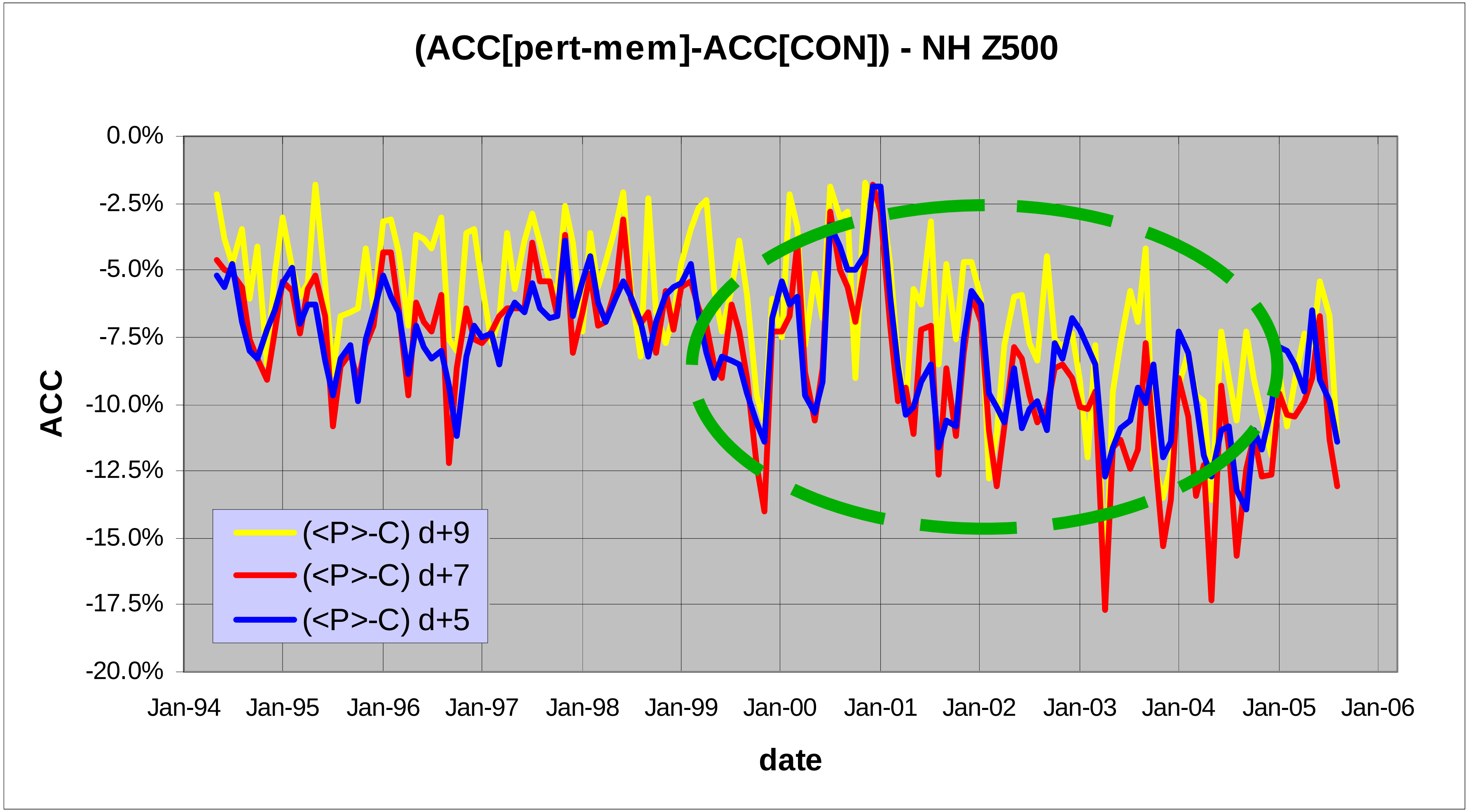


NH: ROC Area for (f>c) - d+3, d+5 and d+7

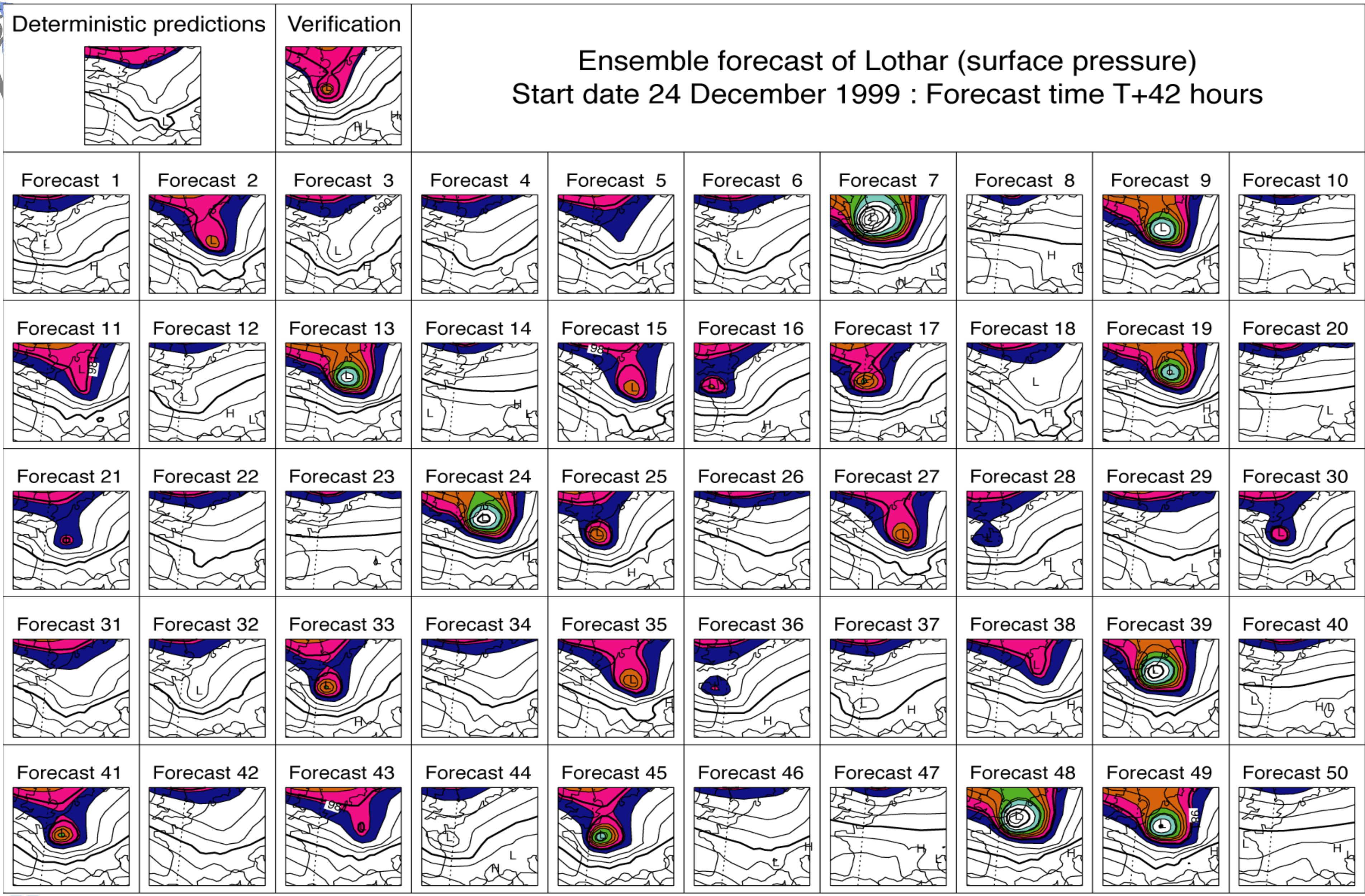


ECMWF

NH: diff averaged pert-members and CON - d+5, d+7 and d+9



ECMWF

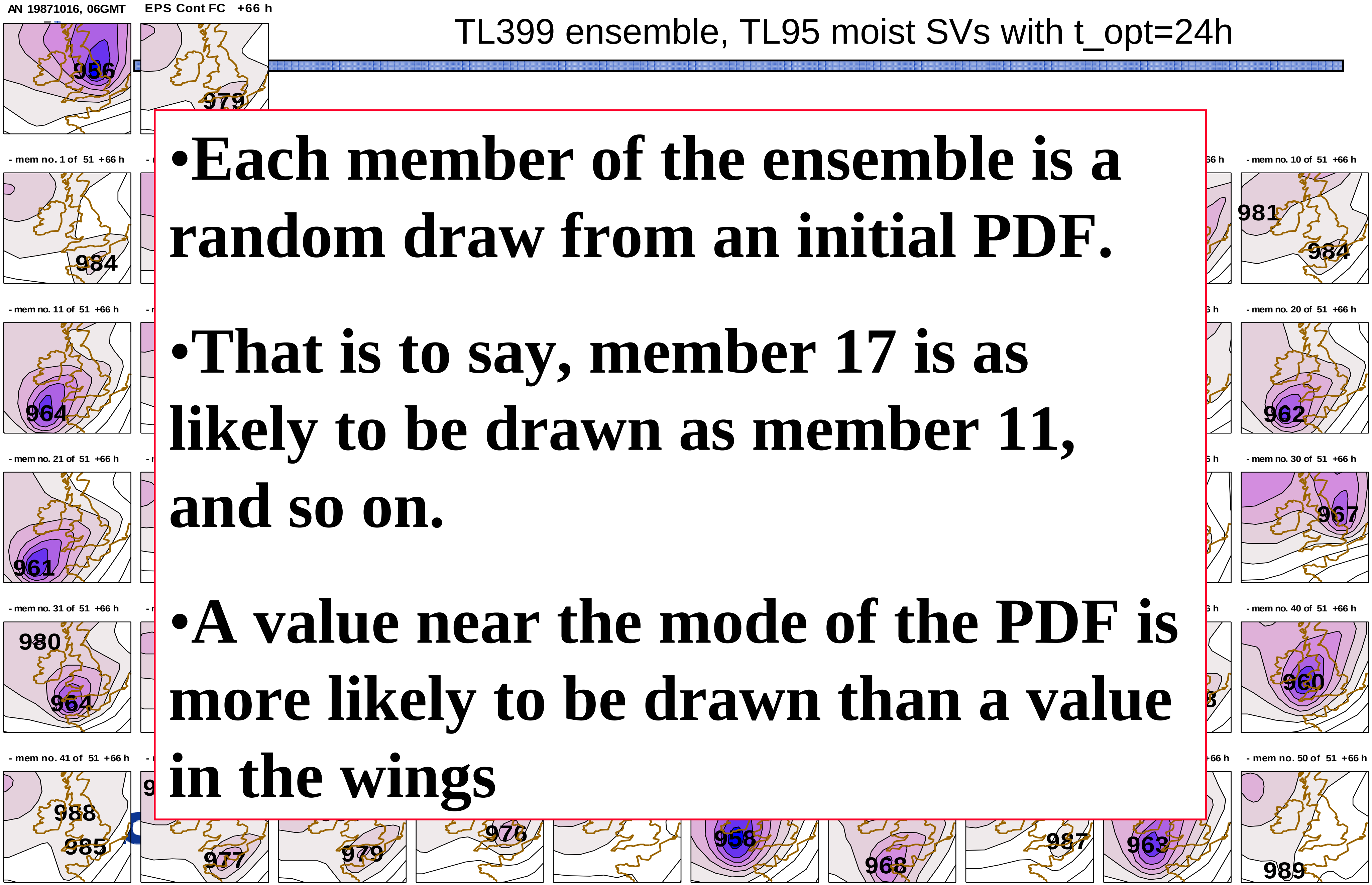


Deterministic predictions
Verification
Ensemble forecast of Lothar (surface pressure)
Start date 24 December 1999 : Forecast time T+42 hours
Forecast 1
Forecast 2
Forecast 3
Forecast 4
Forecast 5
Forecast 6
Forecast 7
Forecast 8
Forecast 9
Forecast 10
Forecast 11
Forecast 12
Forecast 13
Forecast 14
Forecast 15
Forecast 16
Forecast 17
Forecast 18
Forecast 19
Forecast 20
Forecast 21
Forecast 22
Forecast 23
Forecast 24
Forecast 25
Forecast 26
Forecast 27
Forecast 28
Forecast 29
Forecast 30
Forecast 31
Forecast 32
Forecast 33
Forecast 34
Forecast 35
Forecast 36
Forecast 37
Forecast 38
Forecast 39
Forecast 40
Forecast 41
Forecast 42
Forecast 43
Forecast 44
Forecast 45
Forecast 46
Forecast 47
Forecast 48
Forecast 49
Forecast 50

ECMWF

Mean sea level pressure (5hPa) VT: 16 Oct 1987, 6 UTC

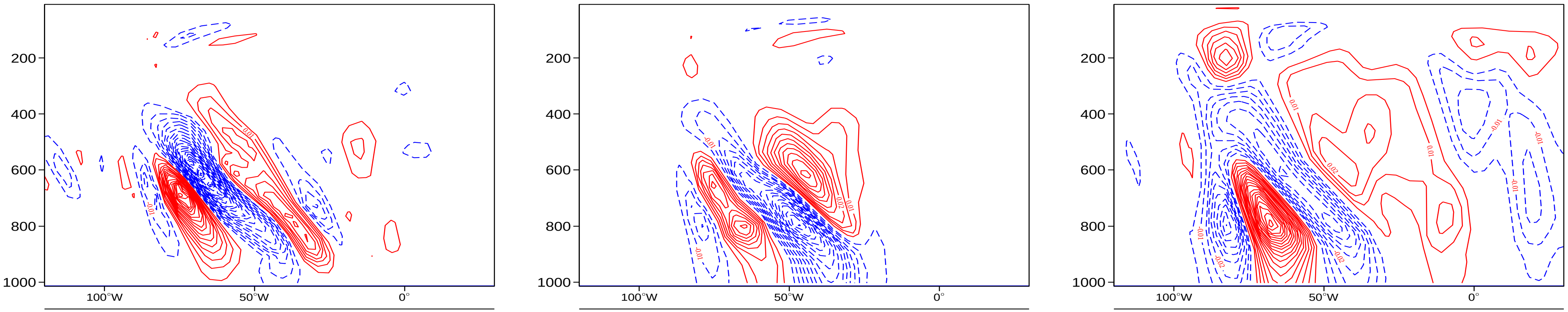
TL399 ensemble, TL95 moist SVs with t_opt=24h



- **Each member of the ensemble is a random draw from an initial PDF.**
- **That is to say, member 17 is as likely to be drawn as member 11, and so on.**
- **A value near the mode of the PDF is more likely to be drawn than a value in the wings**

EPS perturbations

- Leading singular vector temperature cross-section (along 50° N)
- for 23rd December 2003



- TESV

'Full' Hessian SV

$\nabla\nabla J = J_b + J_o$

'Partial' Hessian SV

$\nabla\nabla J = J_b$

The most accurate calculation possible of the initial perturbation which at D+2 has optimal projection on the leading eigenvector of the forecast error covariance matrix (cf Ehrendorfer and Tribbia, 1997)

Weather Roulette

Collaboration with L.Smith, LSE

- London-Heathrow, 2m temperature
- 2002: training data for dressing
- 2003: test data
- odds: set by dressed T511 forecast
- bets: placed by best member dressed EPS
- start capital: £1 (re-invest all money, unlimited stakes)

odds(bin) = 1 / prob_hr(bin)
bets(bin) = prob_eps(bin) * capital(t-1)

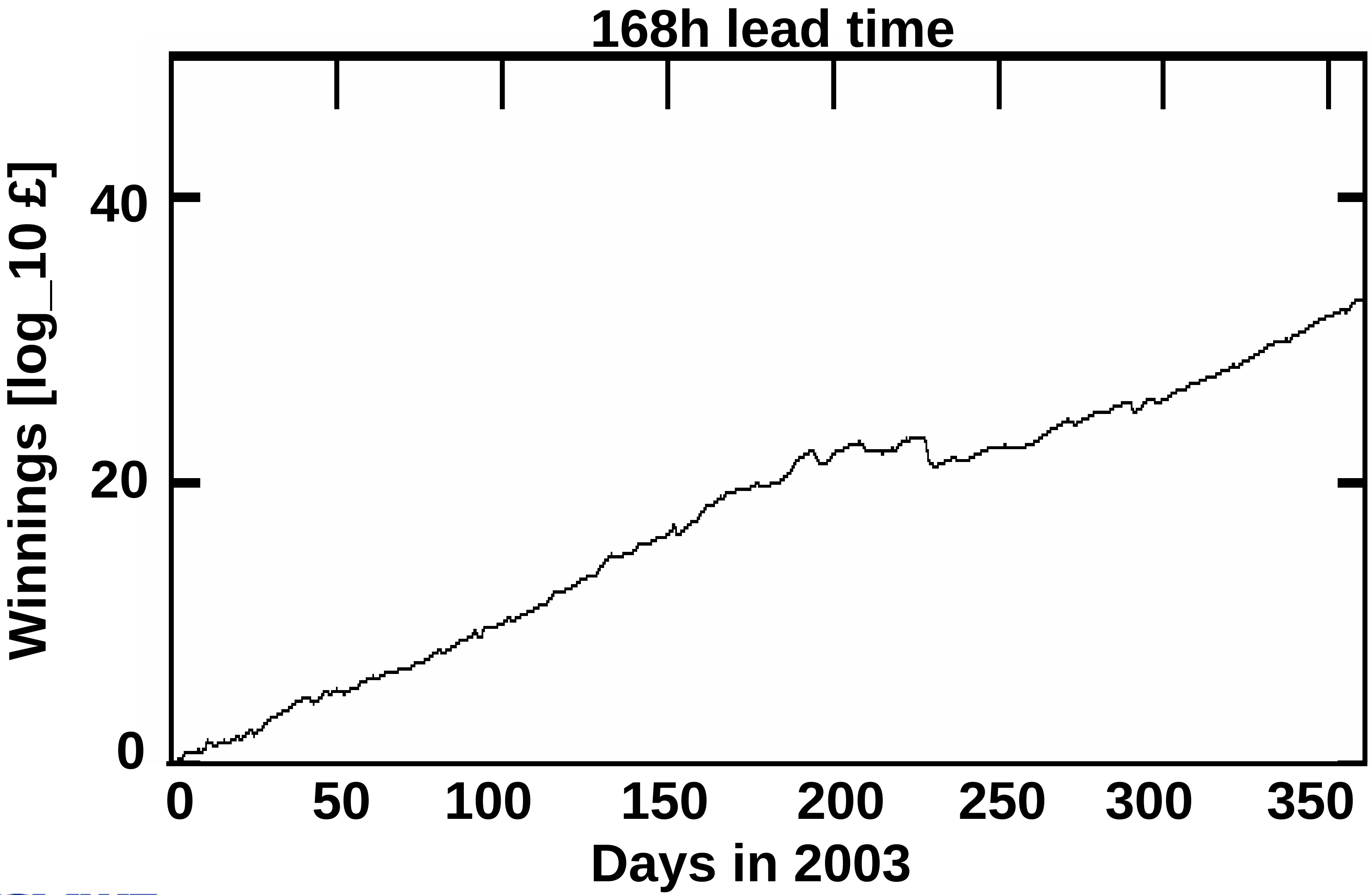
<u>Daily winnings:</u>

win(t) = odds(bin_v) * bets(bin_v) – capital(t-1)
= (prob_eps(bin_v)/prob_hr(bin_v) – 1) * capital(t-1)

TEMPERATURE (°C)

25	26	27	28	29
20	21	22	23	24
15	16	17	18	19
10	11	12	13	14
5	6	7	8	9
0	1	2	3	4
−5	−4	−3	−2	−1

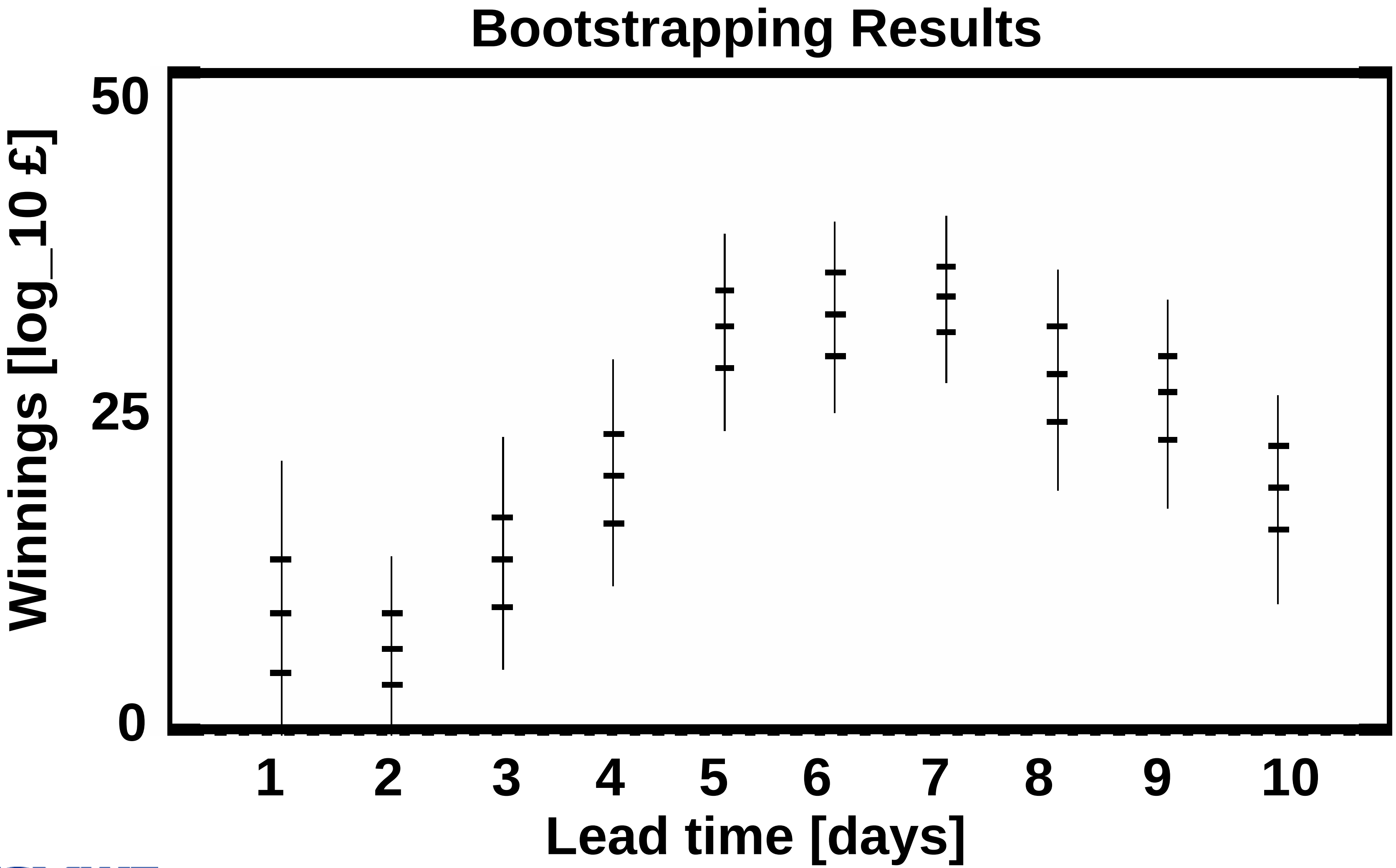
Weather Roulette



ECMWF

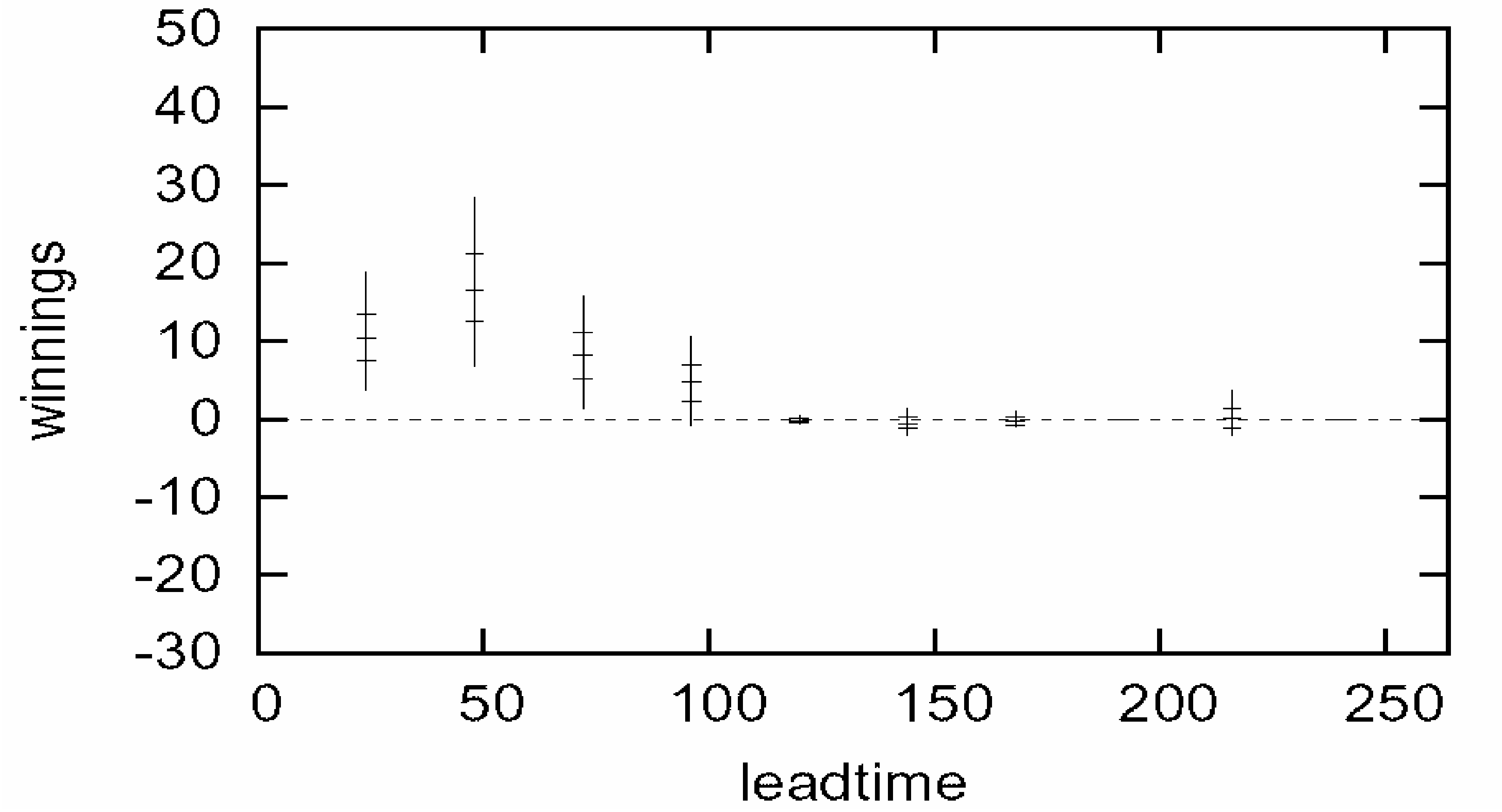
Weather Roulette

Collaboration with L.Smith, LSE



ECMWF

Bootstrapping: optimal blend of HR&EPS vs. EPS



Lagged deterministic ensembles at SMHI – a good idea?

- Small ensemble sizes (poor probabilistic resolution)
- Control cannot be recovered from perturbed ensemble
- The effective perturbations are not independent (ie not orthogonal). Fewer phase-space directions spanned than equivalent size EPS.

ECMWF

Ens. Mean error and covariances of forecast errors

Data: N-Hem extra-tropics, Z500, DJF04/05, daily, forecast step 120h

errors normalised with stdev of control fc error

EPS

	cf	pf 1	pf 2	pf 3	pf 4	pf 5
cf	1.00	0.78	0.79	0.80	0.79	0.78
pf 1	0.78	1.79	0.86	0.90	0.90	0.92
pf 2	0.79	0.86	1.90	0.90	0.97	0.91
pf 3	0.80	0.90	0.90	1.87	0.95	0.91
pf 4	0.79	0.90	0.97	0.95	1.94	0.96
pf 5	0.78	0.92	0.91	0.91	0.96	1.81

lagged cf ens

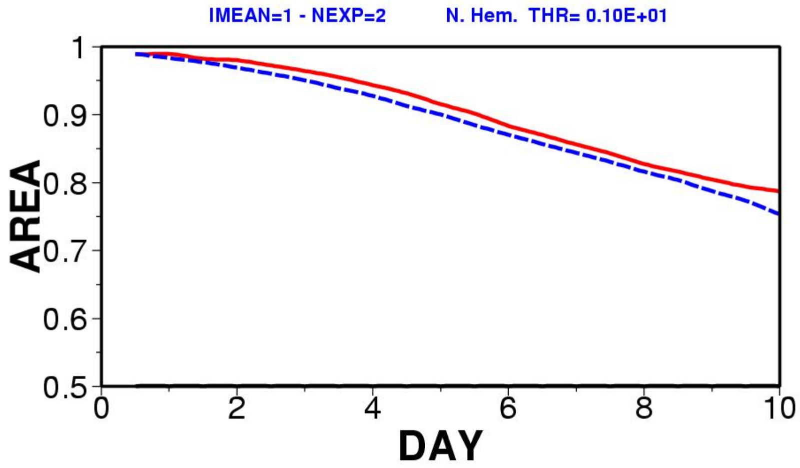
0h=cf	12h	24h	36h	48h	60h	
1.00	0.86	0.78	0.72	0.67	0.65	0h
0.86	1.28	1.08	0.97	0.89	0.83	12h
0.78	1.08	1.61	1.33	1.18	1.07	24h
0.72	0.97	1.33	1.99	1.59	1.40	36h
0.67	0.89	1.18	1.59	2.39	1.84	48h
0.65	0.83	1.07	1.40	1.84	2.79	60h

RMS error of the ensemble mean = $[1/n^2 \text{ x sum of matrix entries}]^{1/2}$

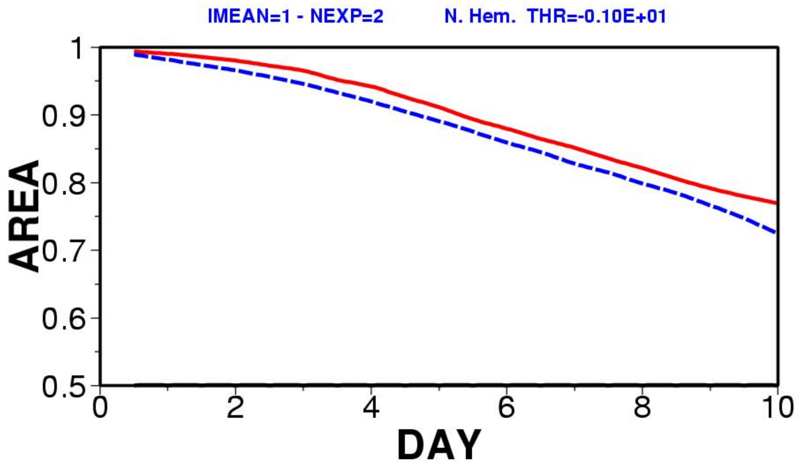
# member	2	3	4	5	6
lagged cf	1.00	1.02	1.04	1.07	1.09
EPS	1.04	1.03	1.02	1.02	1.01

ROCA π[(f-c)>σ], π[(f-c)<-σ], EPS&HHL – Z500 NH, win04/05

Top – Area under the Relative Operating Characteristics (ROCA) for the probabilistic prediction of a positive anomaly larger than 1 climatological standard deviation: the EPS (red line) has a higher ROCA than the HHL (blue line).



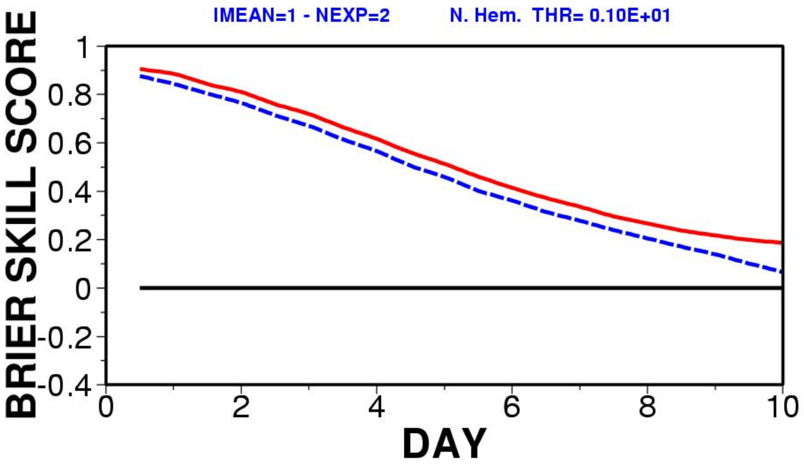
Bottom – Area under the Relative Operating Characteristics (ROCA) for the probabilistic prediction of a negative anomaly smaller than 1 climatological standard deviation: the EPS (red line) has a higher ROCA than the HHL (blue line).



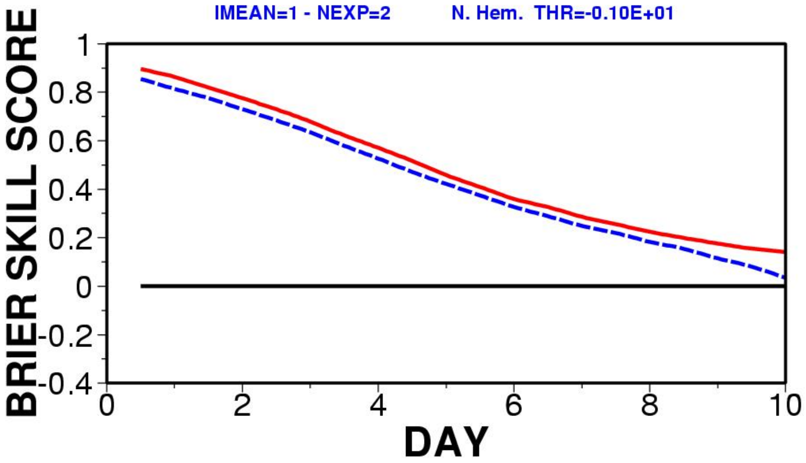
ECMWF

BSS π[(f-c)>σ], π[(f-c)<-σ], EPS&HHL – Z500 NH, win04/05

Top – Brier skill score (BSS) for the probabilistic prediction of a positive anomaly larger than 1 climatological standard deviation: the EPS (red line) has a higher ROCA than the HHL (blue line).



Bottom – Brier skill score (BSS) for the probabilistic prediction of a negative anomaly smaller than 1 climatological standard deviation: the EPS (red line) has a higher ROCA than the HHL (blue line).

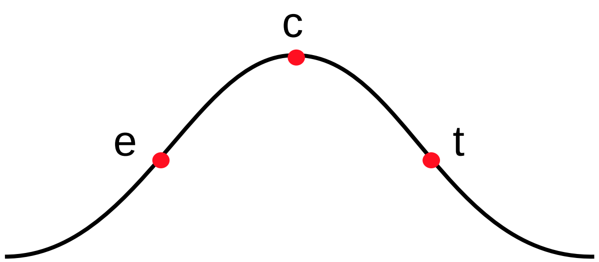


ECMWF

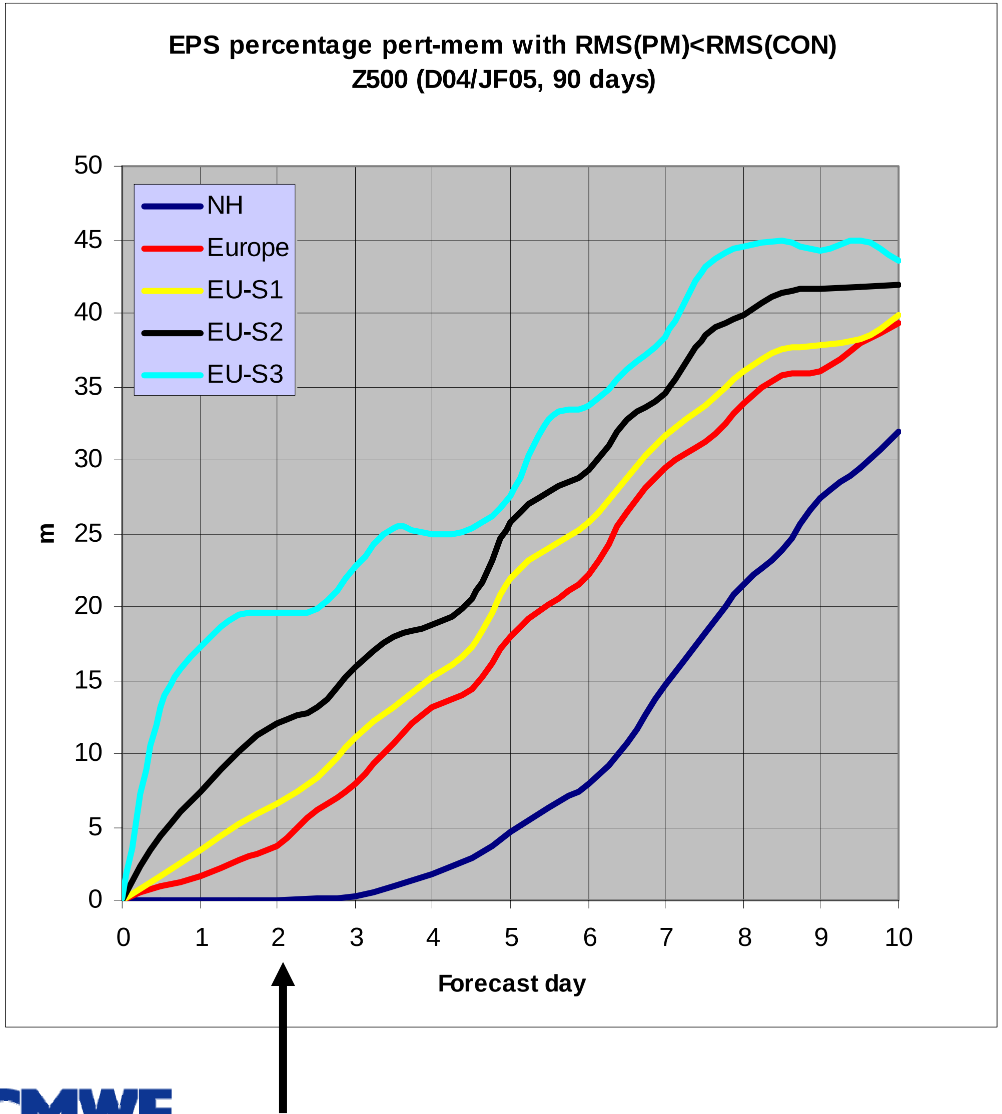
How many members better than control – perfect EPS

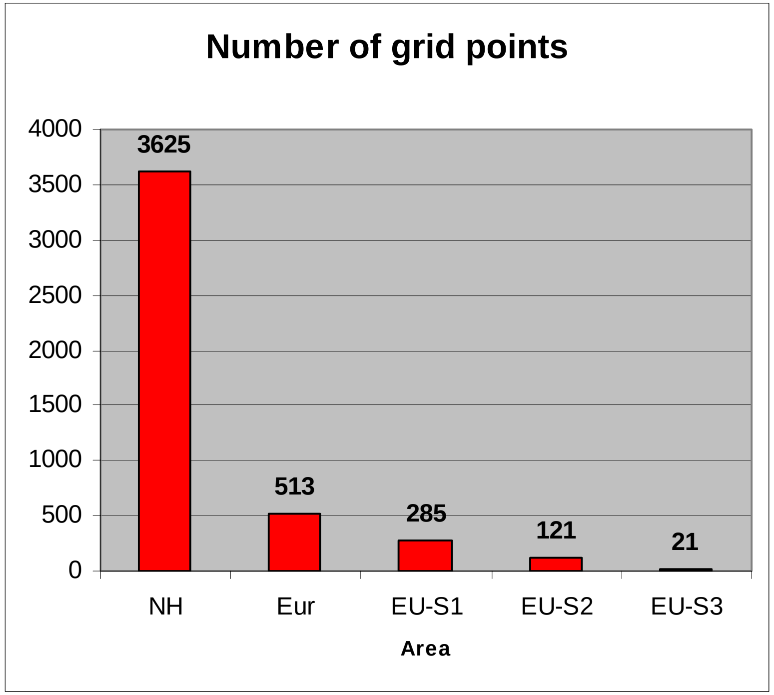
Assume an "ensemble" given by an isotropic Gaussian distribution about a control in *n* dimensions. Further assume a perfect ensemble scenario, i.e. the error of the control is also given by this Gaussian distribution. What is the probability *ρ* of a perturbed member (a draw from the Gaussian) to be closer (in the Euclidean norm) to the true state than the control?

n	1	2	3	4	5	10	20	50	100
ρ_n	0.35	0.28	0.22	0.18	0.16	0.07	0.02	4×10^{-4}	1×10^{-6}



% of pert-mem better than con for different areas – Z500



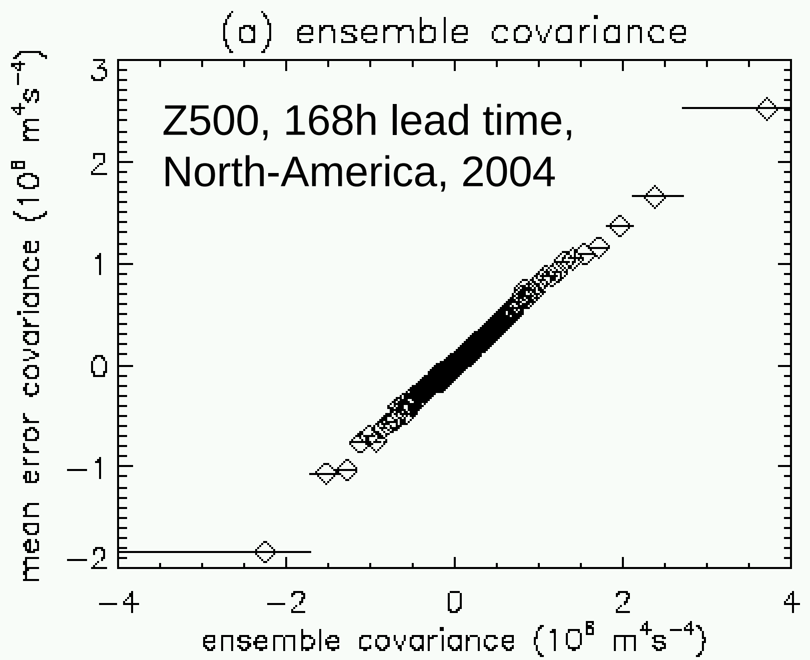


Assume an "ensemble" given by an isotropic Gaussian distribution about a control in *n* dimensions. Further assume a perfect ensemble scenario, i.e. the error of the control is also given by this Gaussian distribution. What is the probability *ρ* of a perturbed member (a draw from the Gaussian) to be closer (in the Euclidean norm) to the true state than the control?

n	1	2	3	4	5	10	20	50	100
ρ_n	0.35	0.28	0.22	0.18	0.16	0.07	0.02	4×10^{-4}	1×10^{-6}

Sub-Europe (n = 3) **Europe** (n = 10) **NH** (n = 100)

Predicting spatial error covariance



Forecast error cov., explained by linear fit

forecast variable	lead time	ensemble covariance	mean past covariance	both predictors
500 mb GP	96h	6.9 ± 0.5	2.0 ± 0.2	7.2 ± 0.5
500 mb GP	168h	16.5 ± 0.5	3.8 ± 0.2	16.5 ± 0.5
500 mb GP	240h	18.9 ± 0.6	5.6 ± 0.2	19.2 ± 0.6
2 m TMP	96h	5.4 ± 0.9	4.8 ± 0.4	9.1 ± 0.8
2 m TMP	168h	9.6 ± 0.5	3.9 ± 0.3	11.1 ± 0.5
2 m TMP	240h	16.7 ± 0.7	5.1 ± 0.3	17.7 ± 0.7

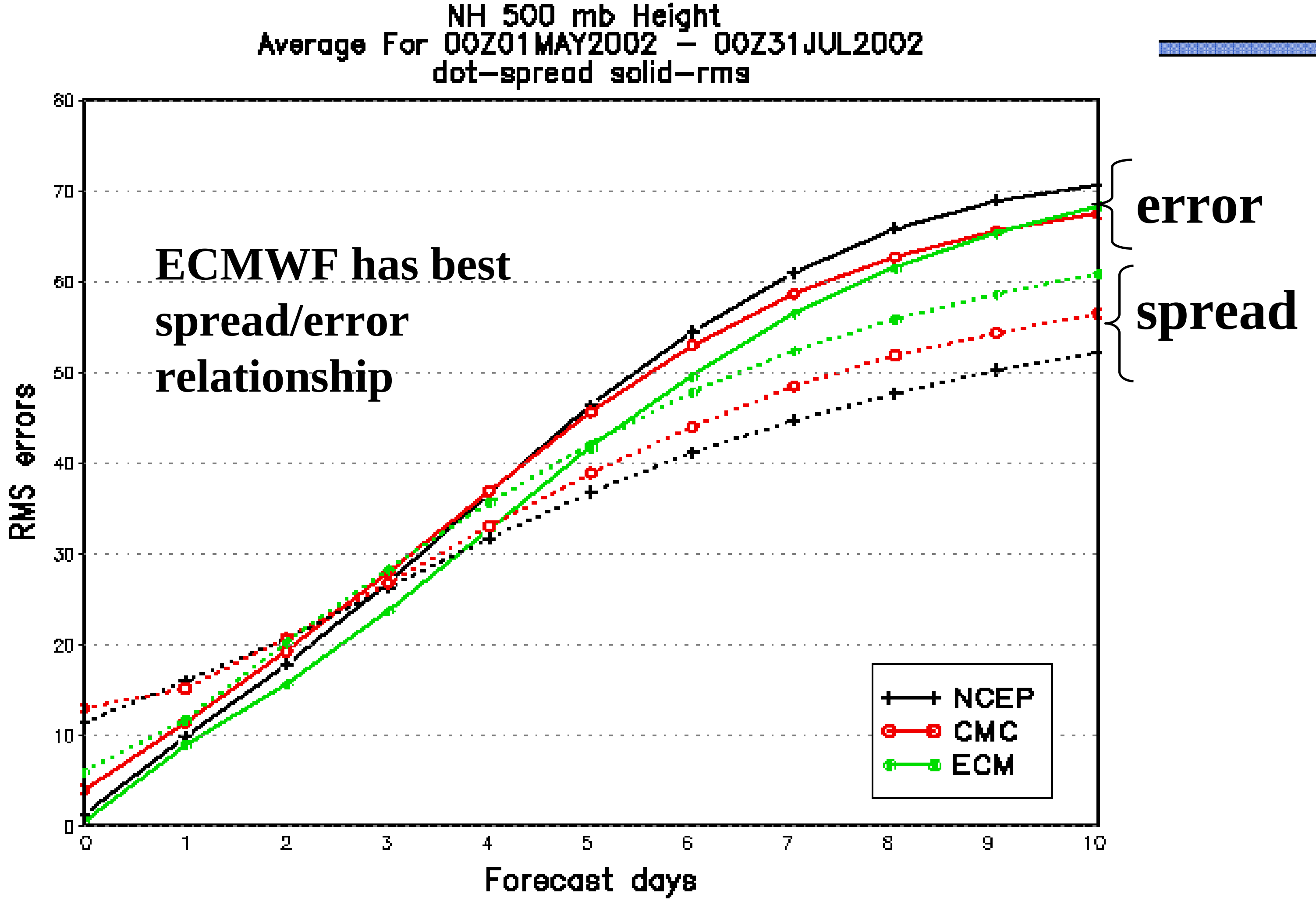
Roulston, 2005

ECMWF

Conclusions

- Spread must balance skill for a good EPS. Reducing spread reduces probabilistic skill.
- ECMWF has the best balance of current operational systems (Buizza et al, 2005) but is not perfect. Representation of model uncertainty still a factor.
- Stamp maps show equally-likely random drawings from initial PDF.
- Singular vectors using full 4DVAR analysis error covariance matrix are similar to energy-metric singular vectors, therefore the latter are consistent with analysis error statistics
- Lagged ensemble will under-perform against EPS because of poor ensemble size and correlation between effective ensemble perturbations.

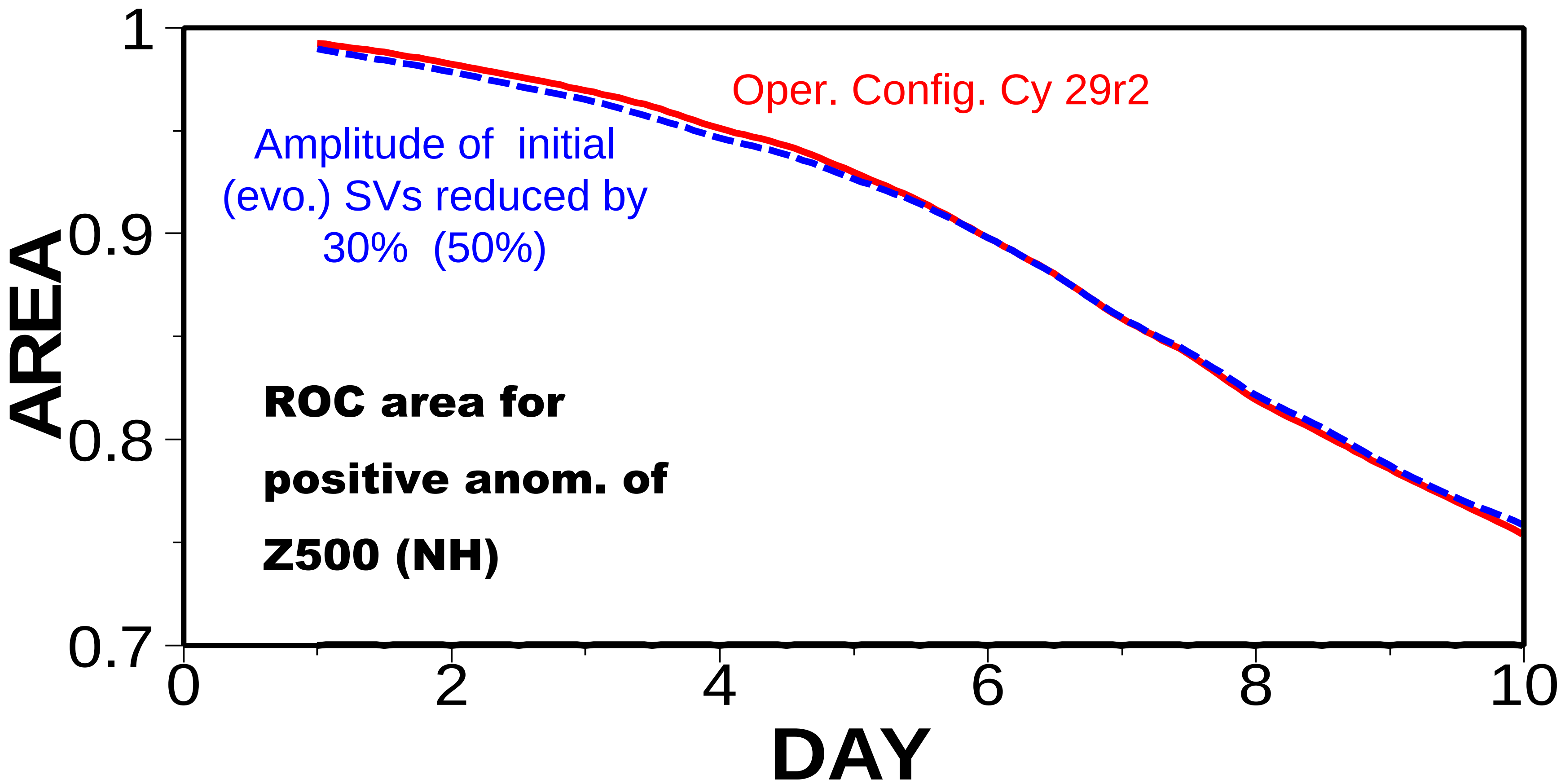
Spread-Error for Three Operational Ensemble Forecast Systems



May-June-July 2002 average RMS error of the ensemble-mean (solid lines) and ensemble standard deviation (dotted lines) of the EC-EPS (green lines), the MSC-EPS (red lines) and the NCEP-EPS (black lines). Values refer to the 500 hPa geopotential height over the northern hemisphere latitudinal band 20º-80ºN. Buizza et al (2005)

ECMWF

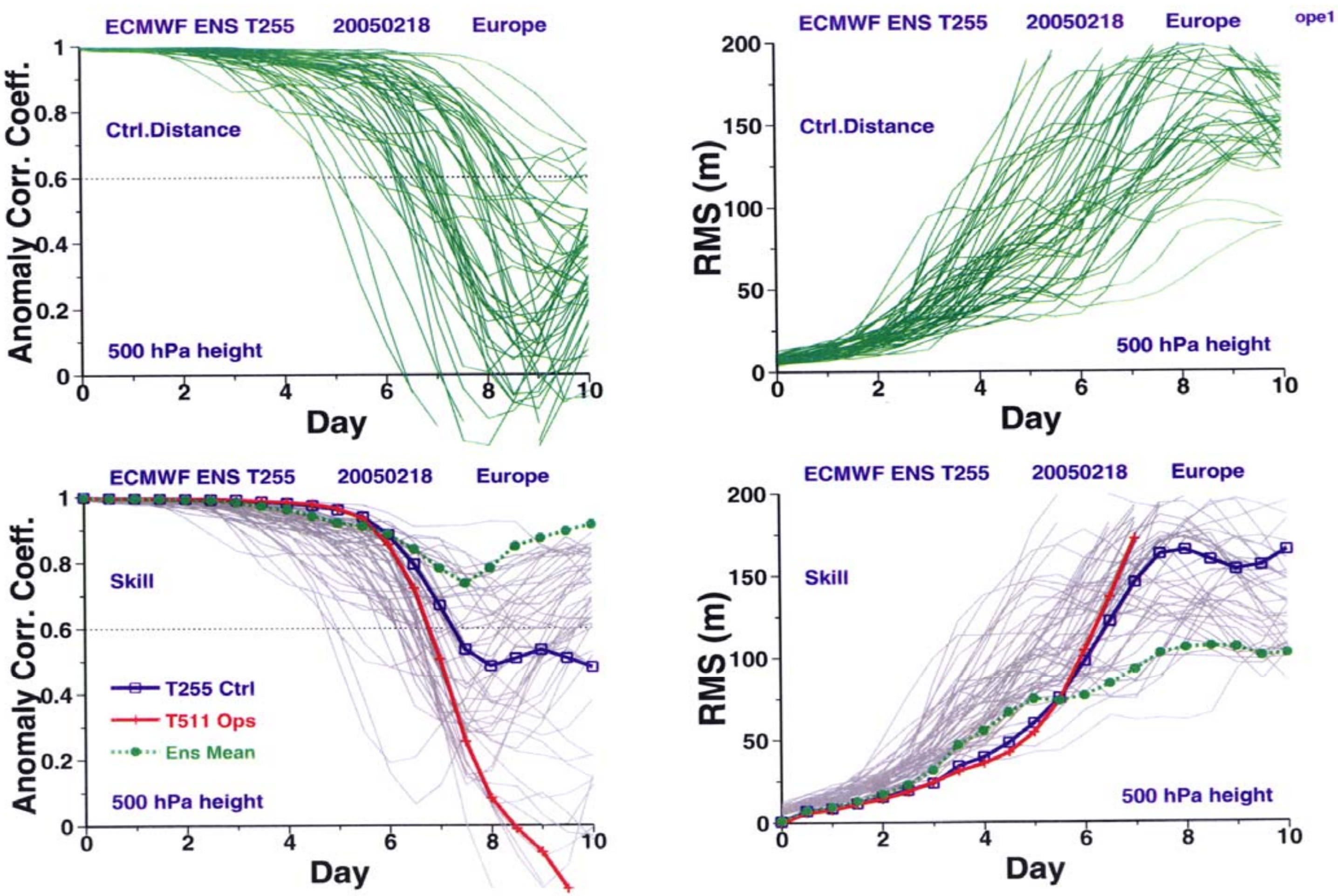
Reduce spread ⇒ Reduce Skill



29 cases in April/May 2005, both experiments cycle 29r2

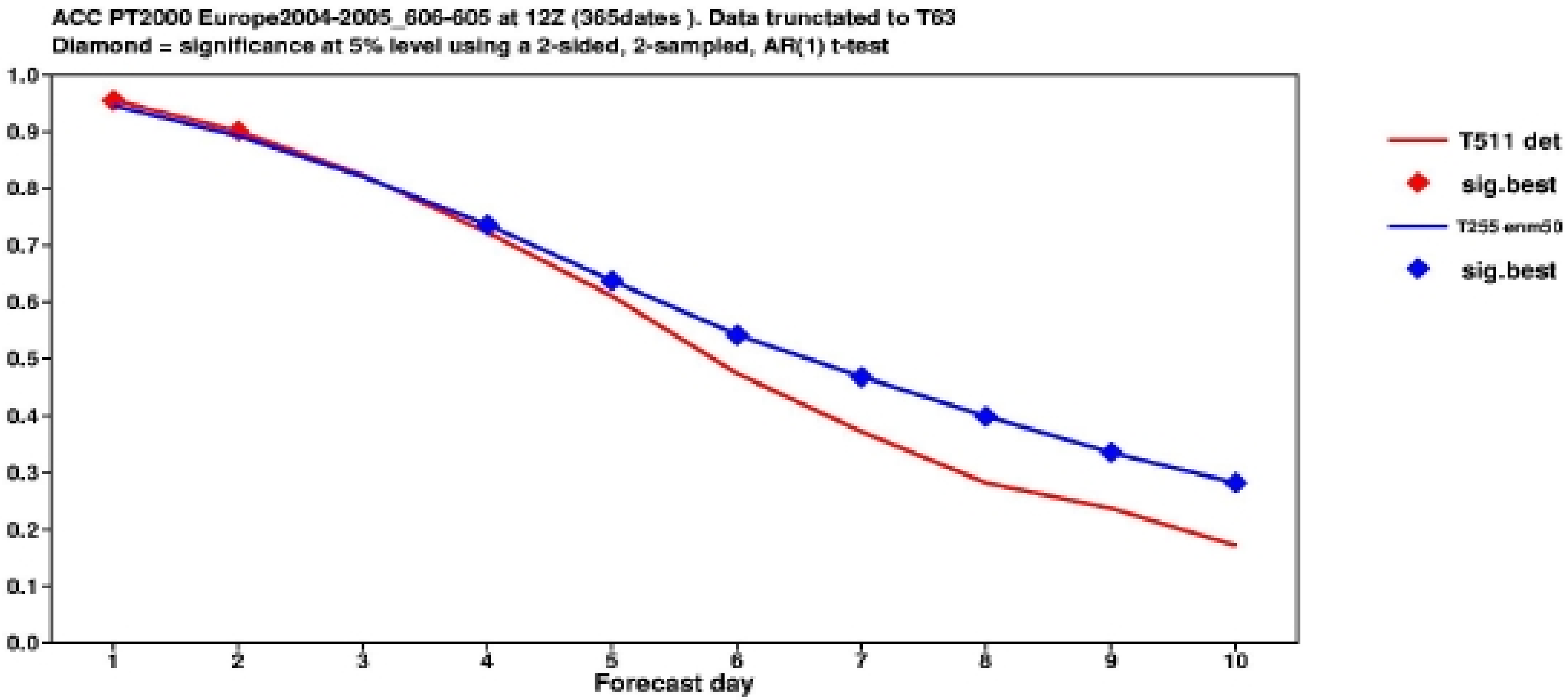
ECMWF

How many members should be better than the control on average?



- **We should expect mean rms error of perturbed EPS members to be up to 40% worse than the control – this is part of the required spread/skill balance**
- **Counting the number of perturbed members better than the control is not a useful diagnostic of EPS performance – it is a function of the number of degrees of freedom in the underlying flow.**
- **What is the right way to compare the EPS vs deterministic forecasts (eg in assessing what fraction of the operational computational resource should be devoted to the EPS compared with the high-res deterministic)?**

Ensemble Mean



EPS competitive with or better than the T511 throughout the range, in terms of θ on PV=2 (where nonlinear filtering of unpredictable scales by EPS begins early in the forecast range)