Verification of Daily CFS forecasts

Huug van den Dool & Suranjana Saha

- CFS was designed as 'seasonal' prediction system
- Hindcasts 1981-2005, 15 'members' per month
- Here we look at CFS(T62L64) as NWP (never mind the delayed ocean analysis)
- 25 years of forecasts (4500 forecasts out to 9 months) by a 'constant' T62L64 model !!!!!.

→ website http://cfs.ncep.noaa.gov/

- *Reference: S. Saha, S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Pena, S. Lord, G. White, W. Ebisuzaki, P. Peng, P. Xie , 2006 : The NCEP Climate Forecast System. Journal of Climate, Vol. 19, No. 15, pages 3483.3517*
- Today: No ensemble averages, hardly time averages
- One could compare to similarly large data sets, such as: CDAS & CDC's MRF & ECMWF 'monthly' system

Acknowledgement: Cathy Thiaw (reruns)
and on CDAS info Bob Kistler/Fanglin Yang/Pete Caplan

Issues

- Day-5 scores
- Climatology of scores
- Bias Correction (mean, sd)
- NH/SH/TR; Z500, PSI200, CHI200
- wk3, wk4, MJO
- Lorenzian Error growth equation
- The perfect forecast
- Scores as function of a) space, b) EOF mode
- Steps to improve signal to noise ratio
- FUTURE (CFSRR)

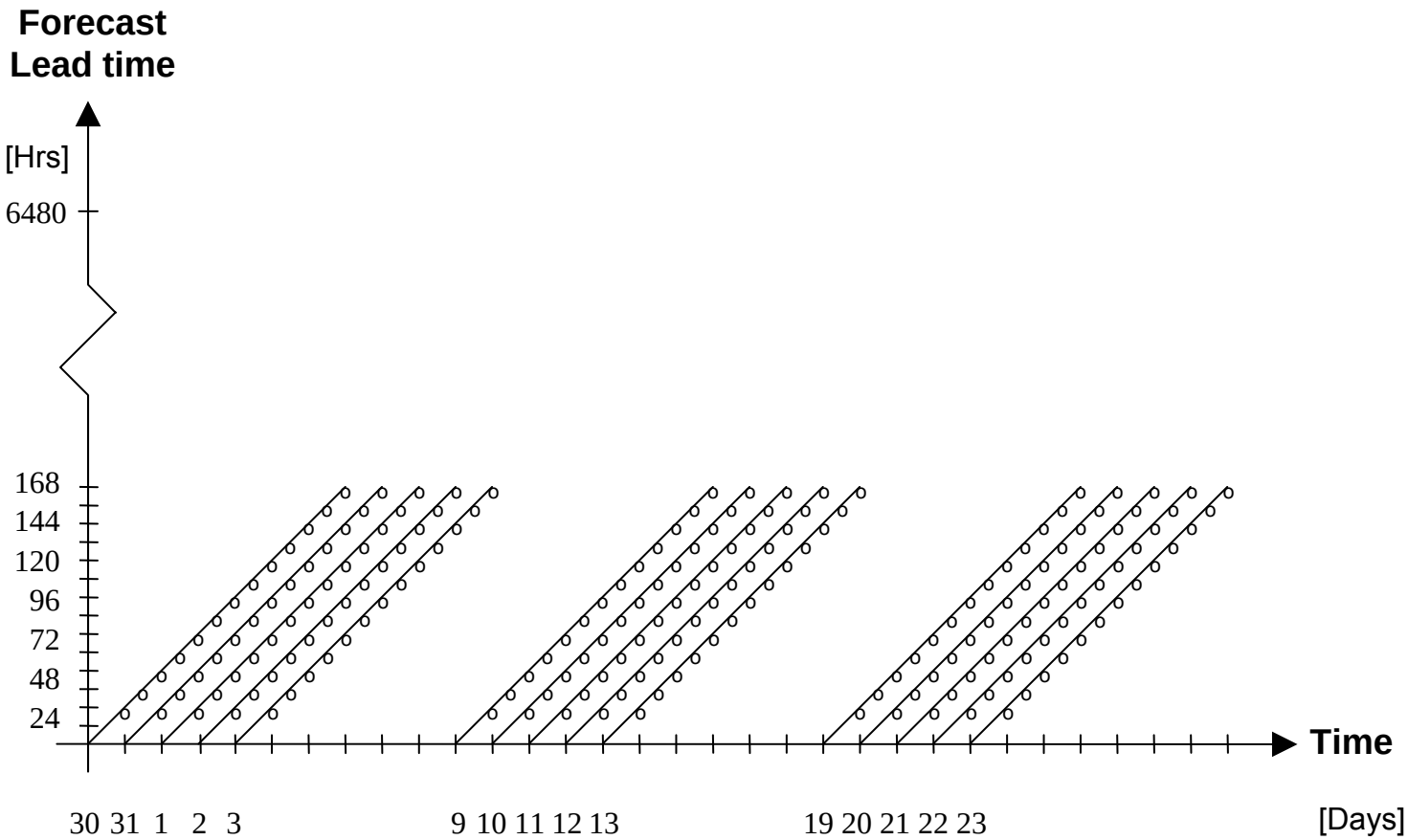
NCEP's NEW CFS Components for S/I Climate

- T62/64-layer version of the ~2003 NCEP atmospheric **GFS** (**G**lobal **F**orecast **S**ystem) model
  - Model top 0.2 mb
  - Simplified Arakawa-Schubert convection (Pan)
  - Non-local PBL (Pan & Hong)
  - SW radiation (Chou, modifications by Y. Hou)
  - Prognostic cloud water (Moorthi, Hou & Zhao)
  - LW radiation (GFDL, AER in operational wx model)
  - R2 Initial conditions for Atmosphere&Land
- GFDL Modular Ocean Model, version 3 (MOM-3)
  - 40 levels
  - 1 degree resolution, 1/3 degree on equator
- Global Ocean Data Assimilation (GODAS)
- **Fully coupled atmosphere-ocean (no flux correction)**

Day 5 AC-scores, using the harmonically smoothed model and observed climatologies

(which are more ‘competitive’ than the old-old-old monthly climo used on Pete Caplan’s page

Variables:
Extra-tropical Z500 (NH, SH)
PSI200 and CHI200


Forecast
Lead time
[Hrs]
6480
168
144
120
96
72
48
24
Time
30 31 1 2 3
9 10 11 12 13
19 20 21 22 23
[Days]

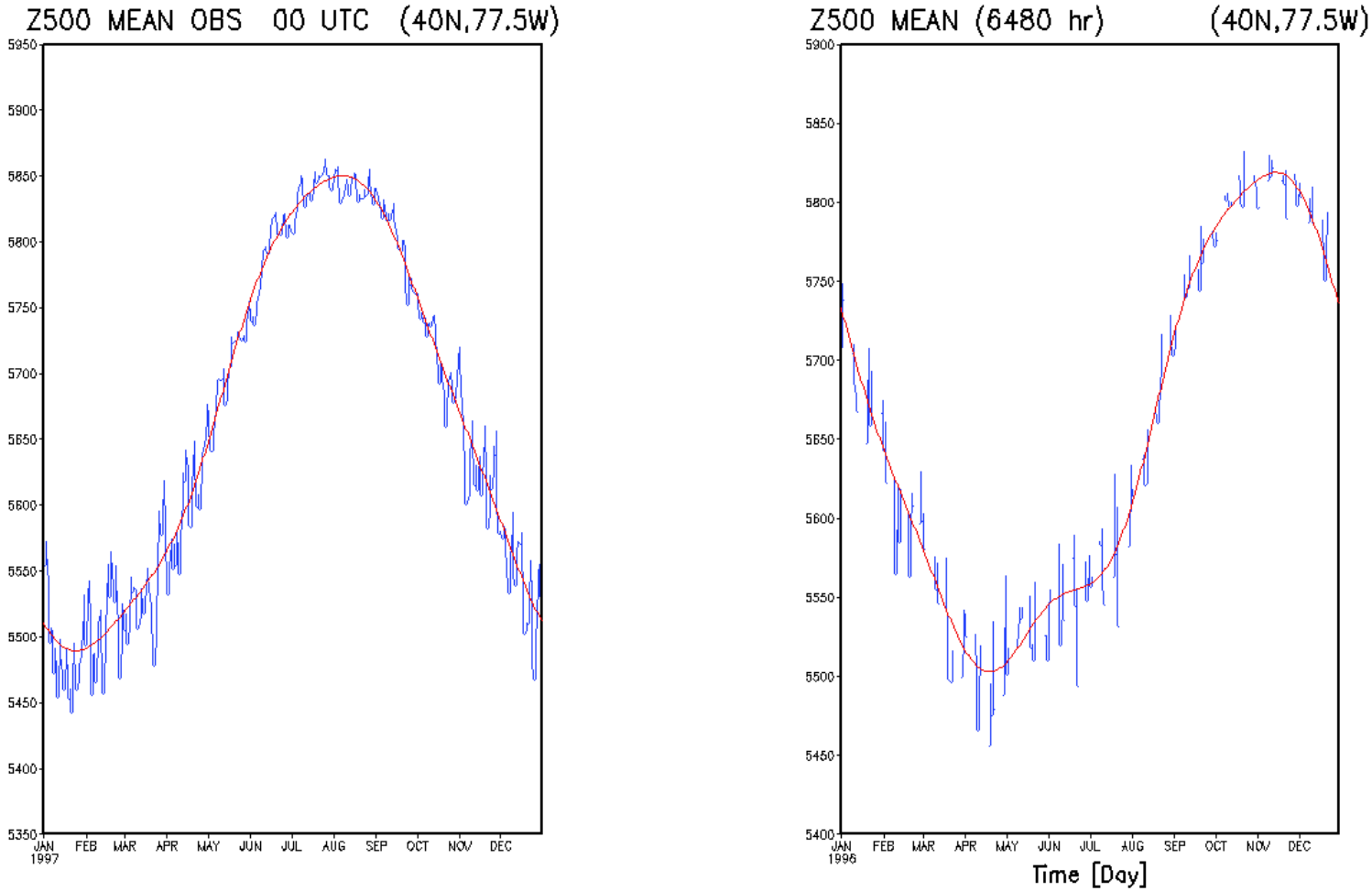
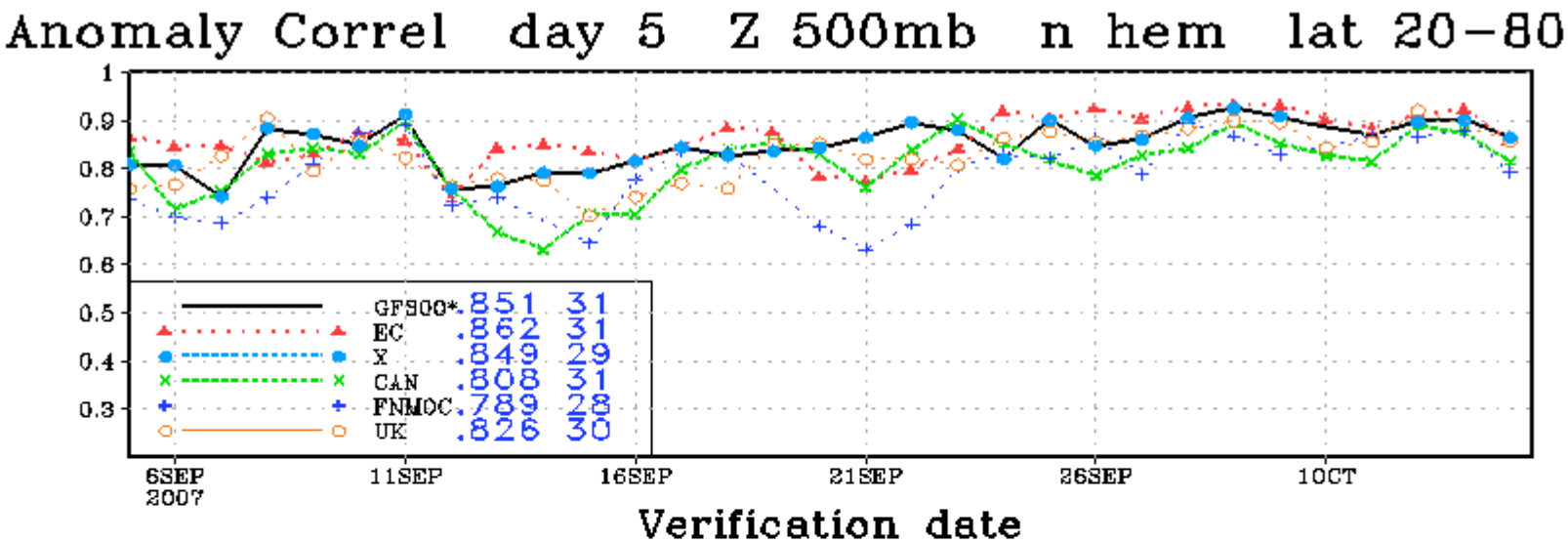


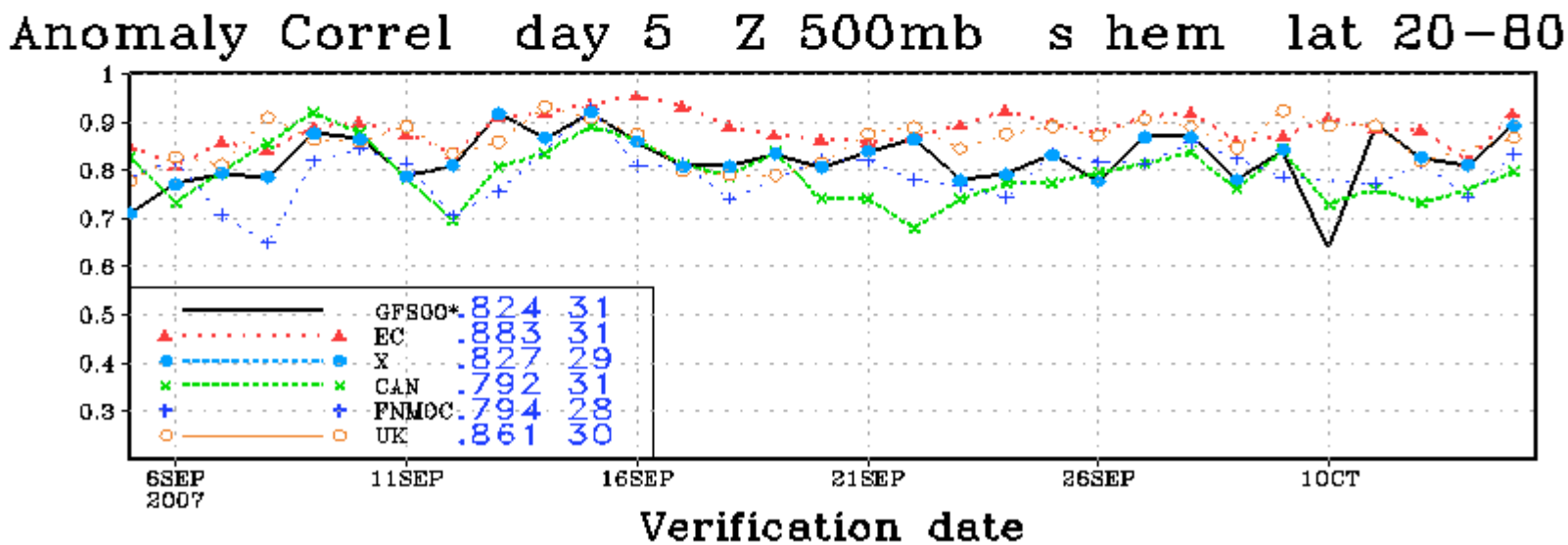
Fig. 2. The climatological annual cycle of the mean value of the geopotential height at a grid point close to Washington, DC, is shown as a smooth red curve. Shown are both the observed climatology (left panel) and the forecast climatology at a lead of 6480 hours (~ 9 months) (right panel). The 24-yr mean values as calculated directly from the data are shown by the blue curves. The reason for the discontinuities in the right panel is that these values are only available during roughly half of the days in a year. The time on the abscissa in the right panel is the initial time of the forecast, not the actual time when the forecast is valid. In this case this implies an offset by roughly nine months. Unit is m. ACKNOWLEDGEMENT: Ake Johansson

Day 5 AC-scores, using the harmonically smoothed model and observed climatologies

(which are more ‘competitive’ than the old-old-old monthly climo used on Pete Caplan’s page

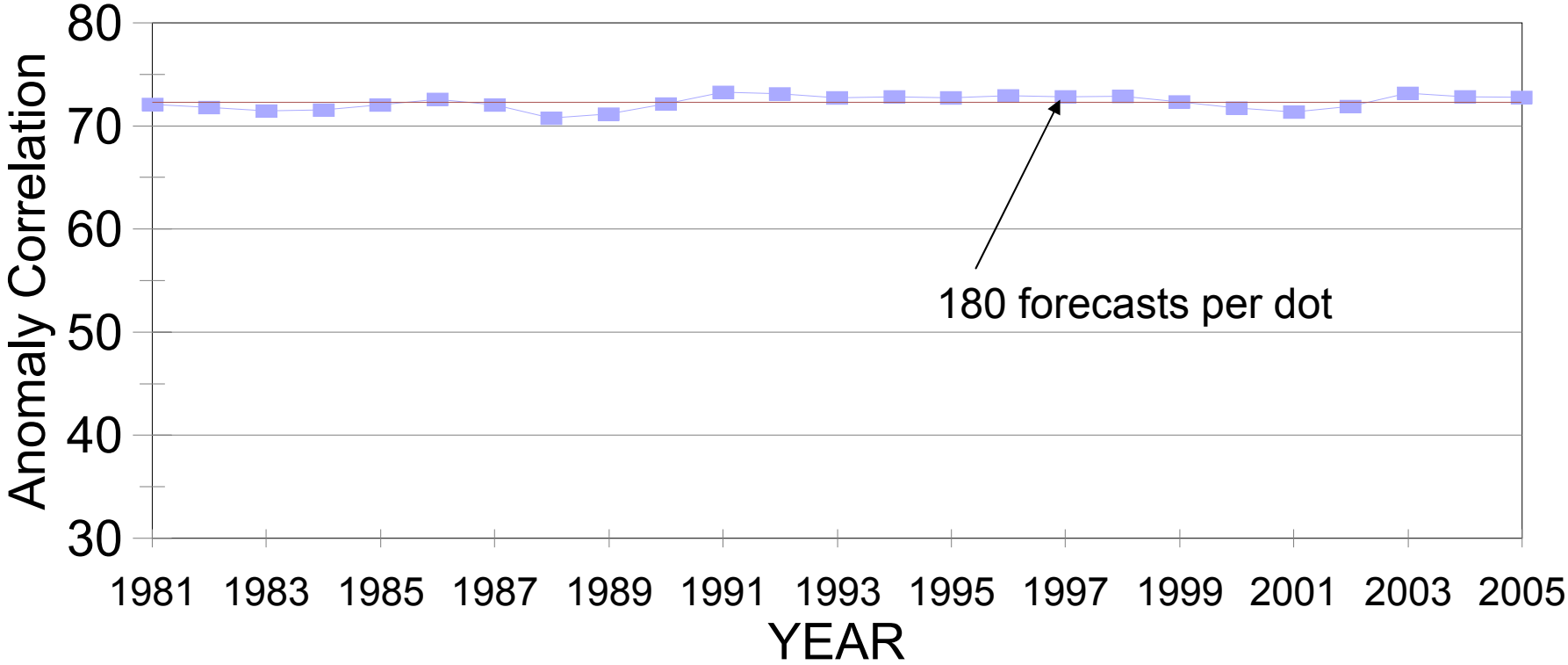
Variables:
Extra-tropical Z500 (NH, SH)
PSI200 and CHI200


Anomaly Correl day 5 Z 500mb n hem lat 20-80
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
GFS00* .851 31
EC .862 31
X .849 29
CAN .808 31
FNMOC .789 28
UK .826 30
6SEP 2007
11SEP
16SEP
21SEP
26SEP
10CT
Verification date


Anomaly Correl day 5 Z 500mb s hem lat 20-80
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
GFS00* .824 31
EC .883 31
X .827 29
CAN .792 31
FNMOC .794 28
UK .861 30
6SEP 2007
11SEP
16SEP
21SEP
26SEP
10CT
Verification date

Annual Mean CFS day 5 scores

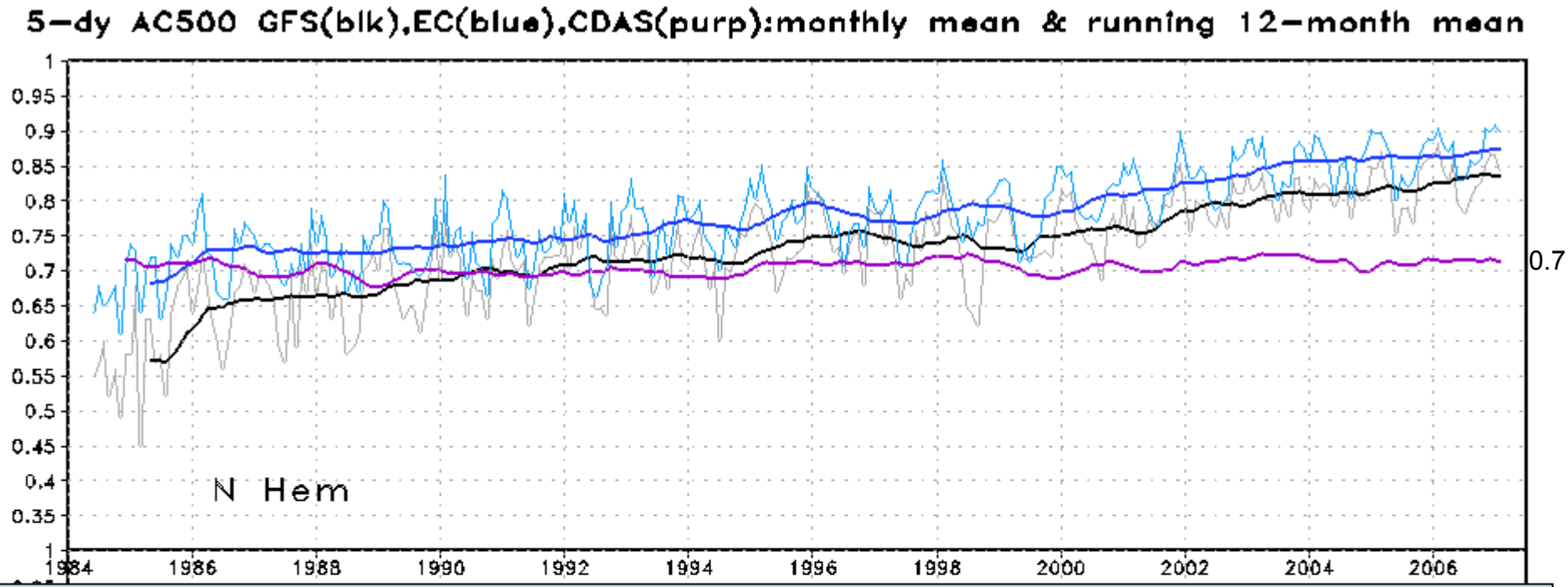
Z500 - NH 1981-2005



Z500: 72.3 (4500 cases, grand mean)

Congratulations with a constant system!

From Pete Caplan's EMC website:

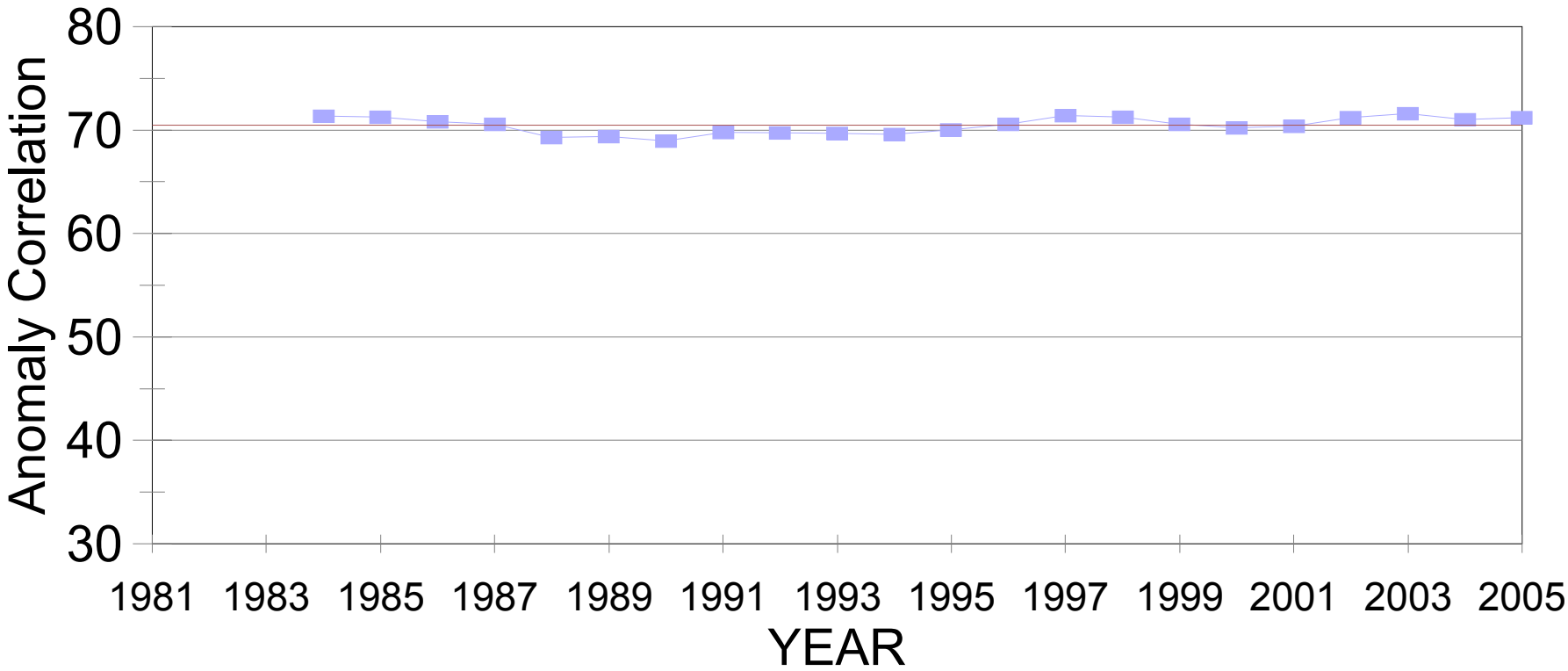


Grand Mean NH over 1984-2005: 70.5 (CDAS1)

((CDAS statistics has several problems))

Annual Mean [redacted] day 5 scores

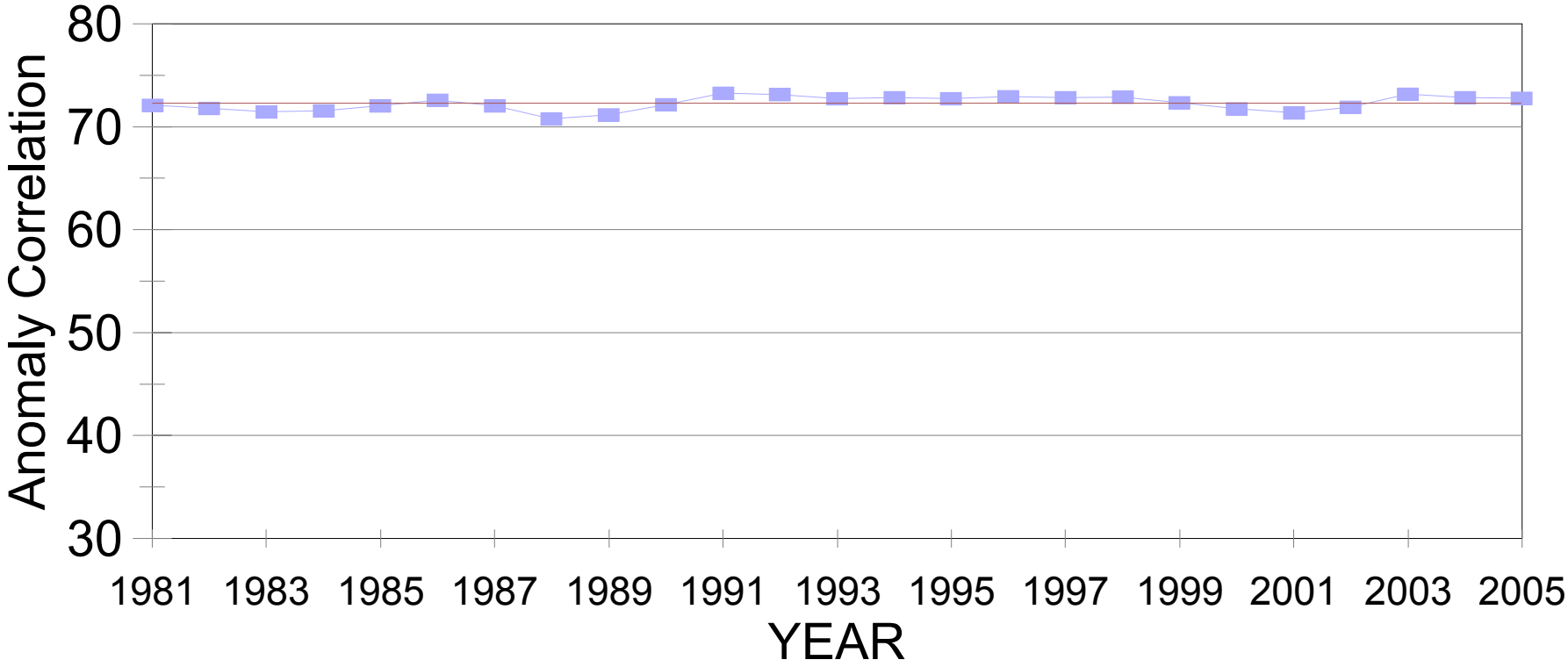
Z500 - NH 1981-2005 (CDAS)



CDAS

Annual Mean CFS day 5 scores

Z500 - NH 1981-2005

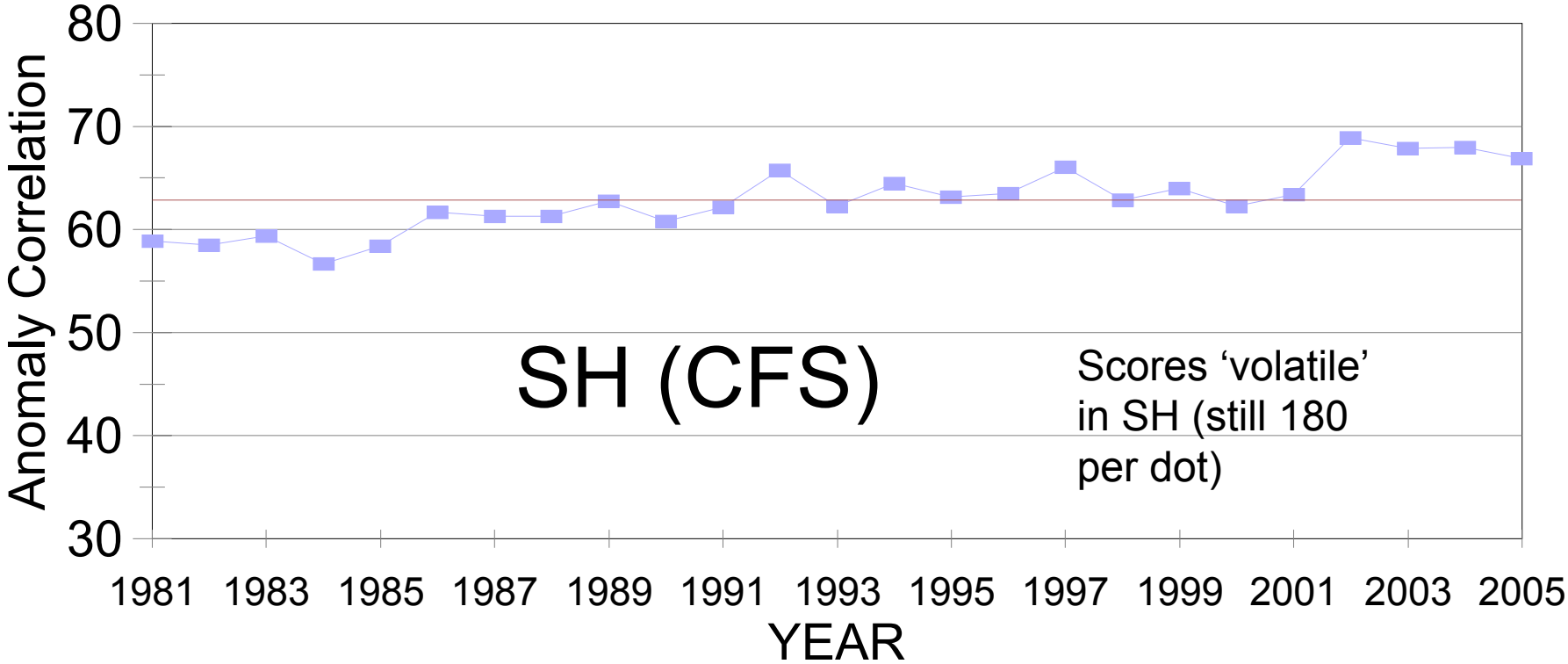


CFS

Next Two Slides: Excursion to SH Z500

Annual Mean CFS day 5 scores

Z500 - SH 1981-2005

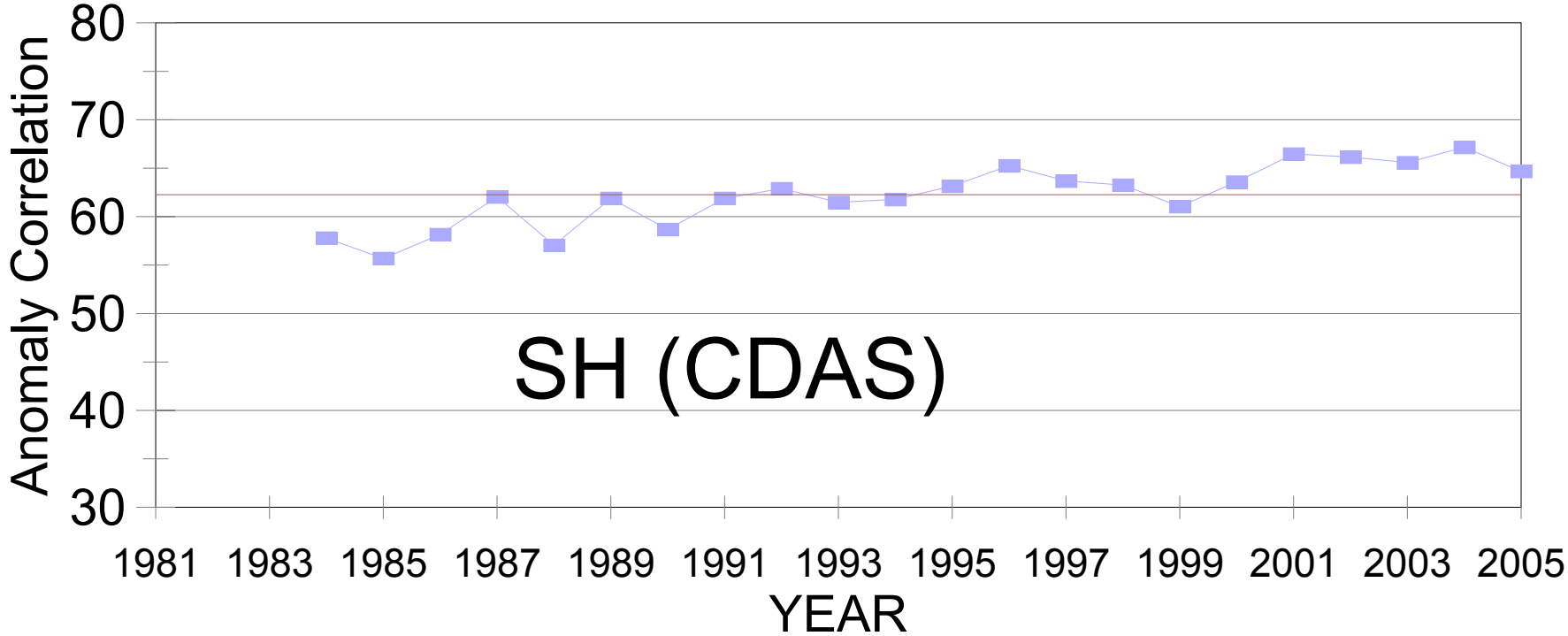


Z500: 62.9 (4500 cases, grand mean)

Scores go UP and UP! "Congratulations" to whom ???

Annual Mean [redacted] day 5 scores

Z500 - SH 1981-2005 (CDAS)



Doing the best we can: Comparing CFS to CDAS day 5 Z500 scores)

	NH	SH
CFS 1981-2005 T62L64, coupled, R2, harmonic daily climo	0.723	0.629
CDAS 1984-2005 T62L28, R1, 'old' monthly climo	0.705	0.623

Flawed

Warning: Number of forecasts per month differ. Climos differ!!! and change in '96 for CDAS

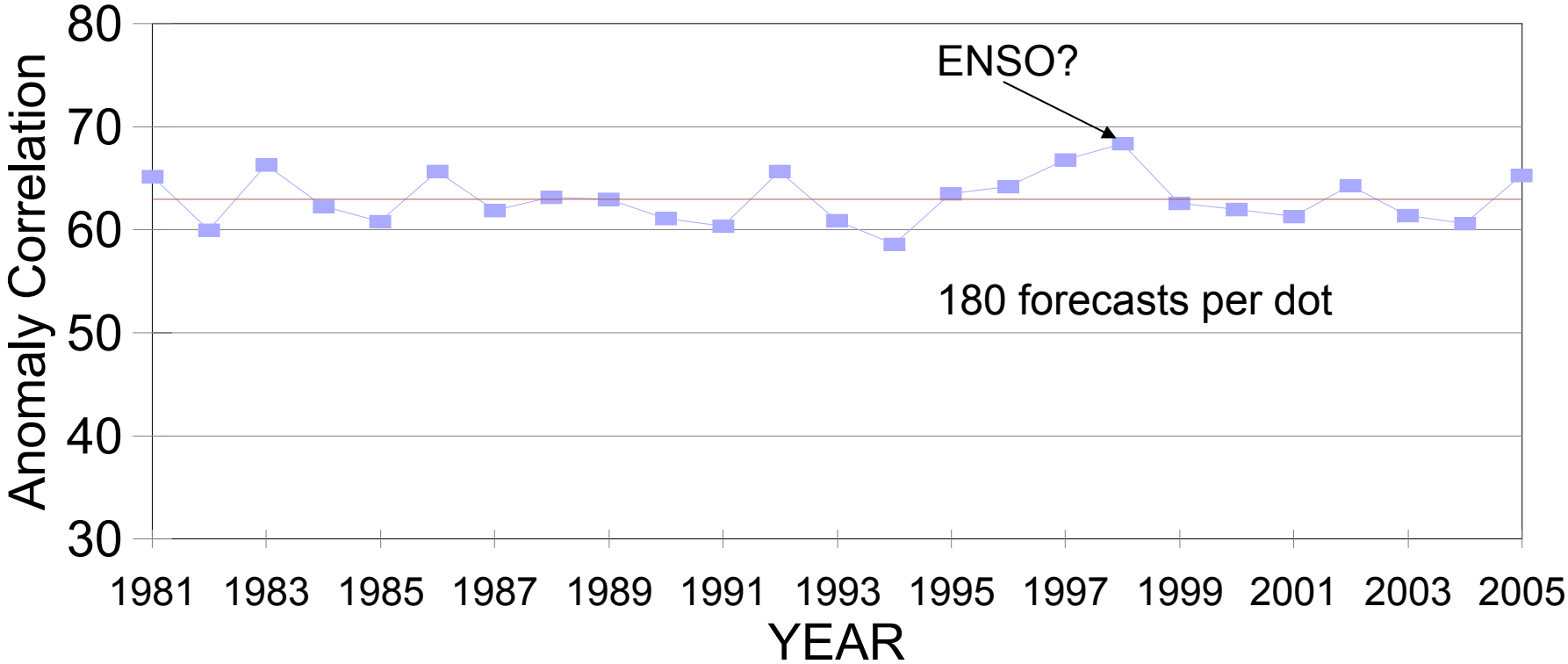
Prelim Conclusion

- R2/CDAS2/CFS 5-day forecasts are (slightly) better than R1/CDAS.
- NH much better than SH (typical for pre-2000 technology) – this will change largely in next CFSRR (Jan2010??)
- Trend-issues and non-constancy of system (will get worse)

Next Two Slides: Excursion to TROPICS

Annual Mean CFS day 5 scores

PSI200 - TROPICS 1981-2005

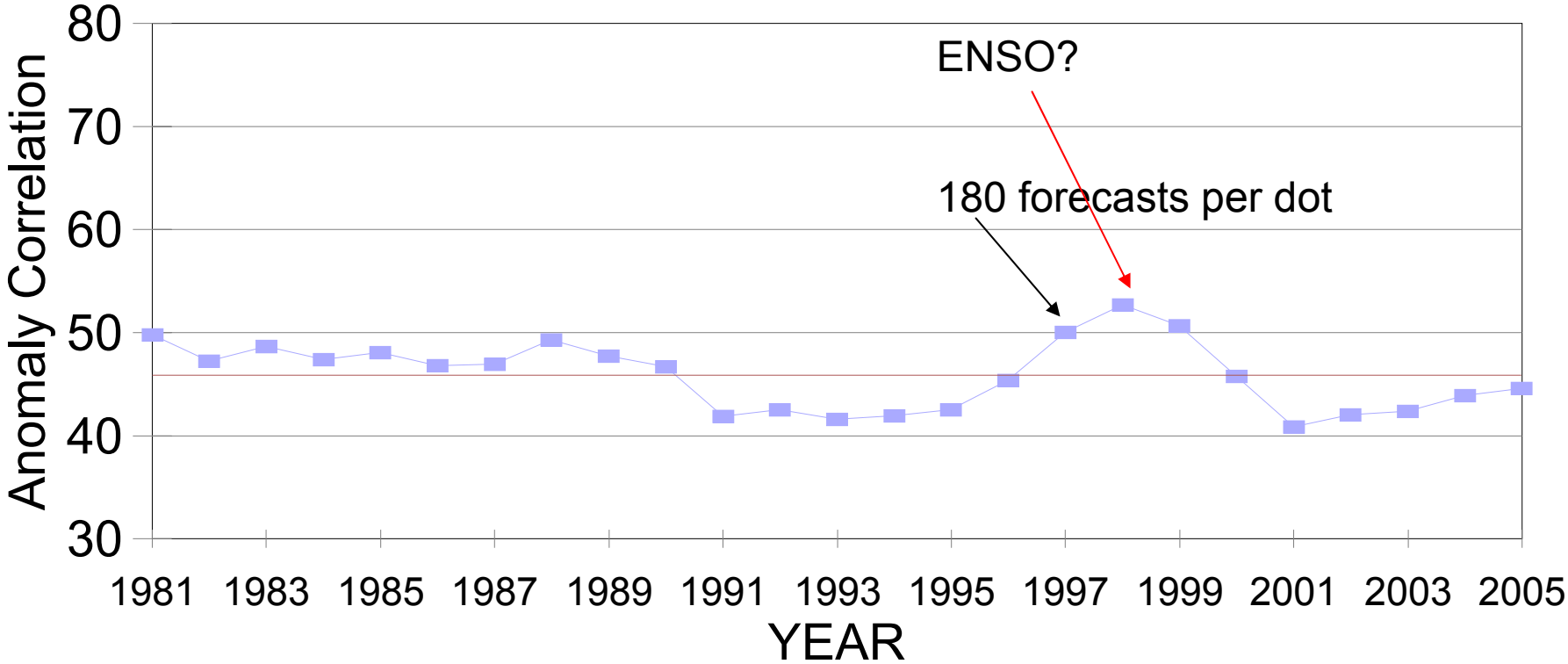


PSI200: 63.0 (4500 cases grand mean)

No increase???

Annual Mean CFS day 5 scores

CHI200 - TROPICS 1981-2005



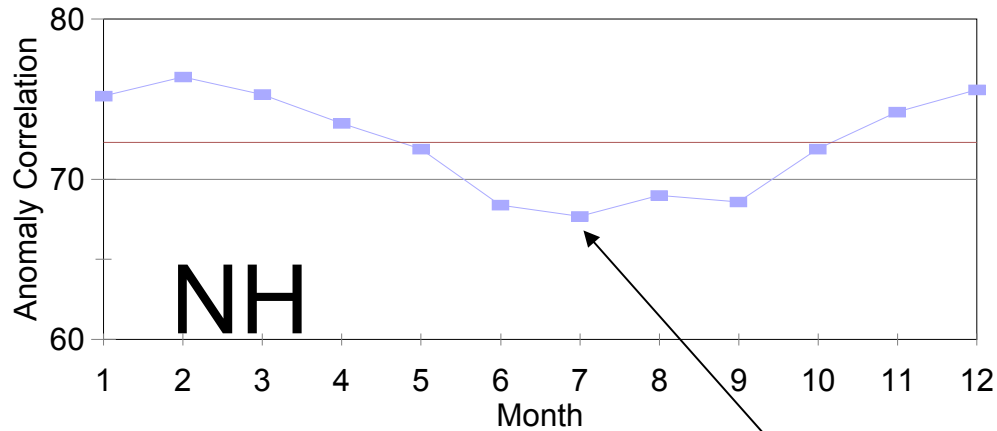
CHI200: 45.9 (4500 cases grand mean)
This may be lowish for MJO type prediction.

Definitely no increase over time. Perhaps a decrease!!!

Climatological Annual Cycle of day-5 scores

Climo Day 5 Scores by Month

Z500 - NH 1981-2005

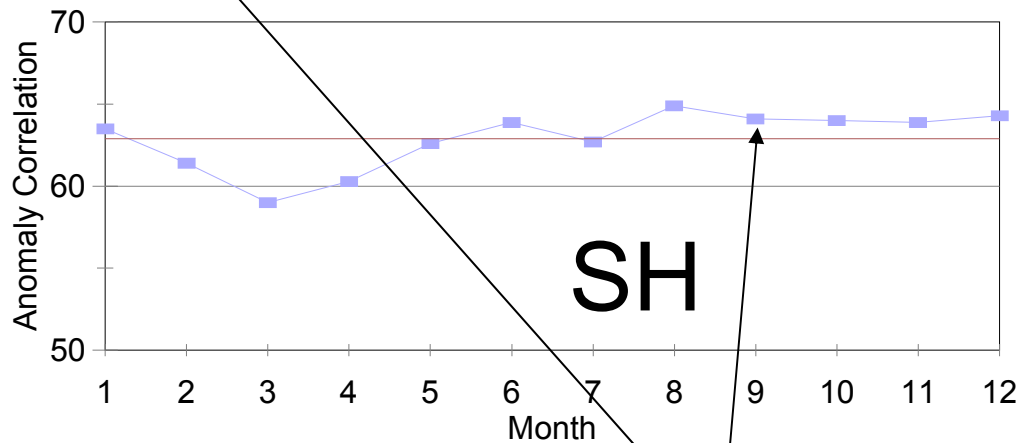


NH: Best month is Feb (76.4), Worst month is July (67.7). Near sinusoidal variation.
Range=8.7pnts

Should we be surprised that the winter-summer difference in NH is NOT larger!??

Climo Day 5 Scores by Month

Z500 - SH 1981-2005



SH: Best month is Aug (64.9), Worst month is March (59.0). Typically 62-64, except Feb-Apr.
Range=5.9pnts

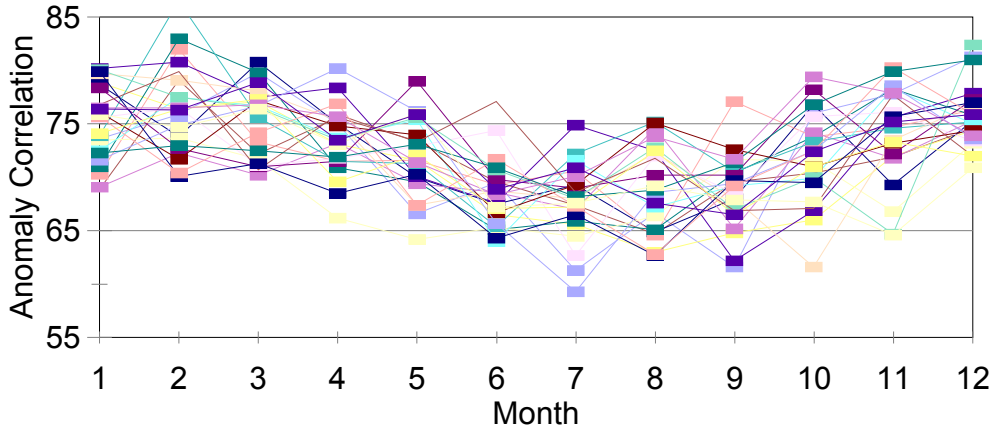
Forecasts verifying in month shown

375 forecasts per month

Curious: March is best/worst in NH/SH. Aug is best/worst in SH/NH.

Climo Day 5 Scores by Month

Z500 - NH 1981-2005

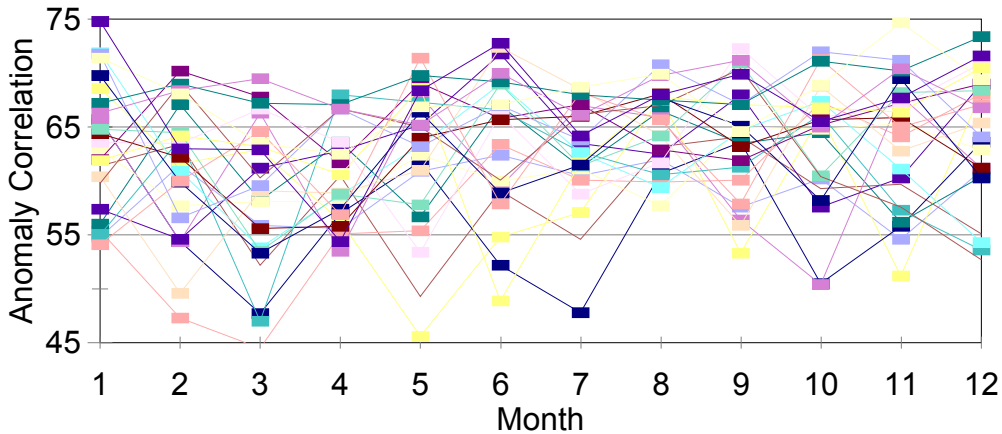


The NH has an annual cycle in skill, every year.

(Each color curve is a different year.)

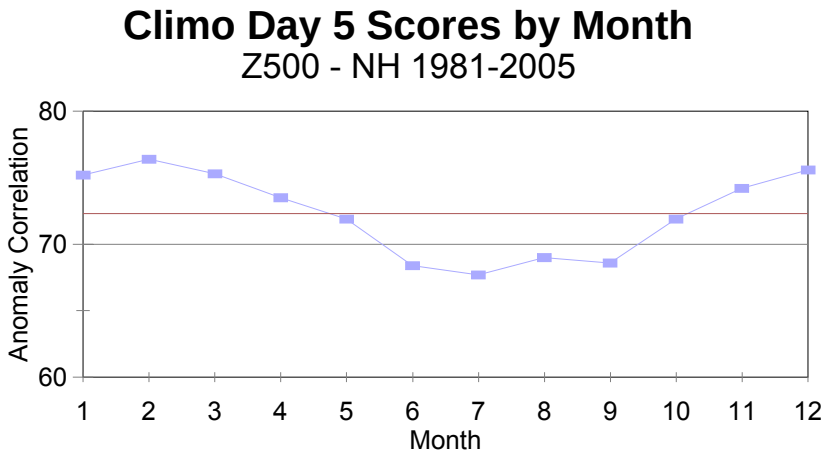
Climo Day 5 Scores by Month

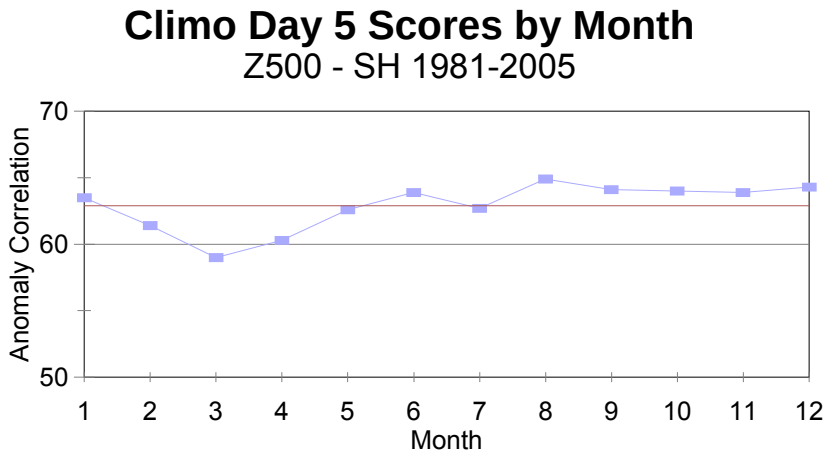
Z500 - SH 1981-2005

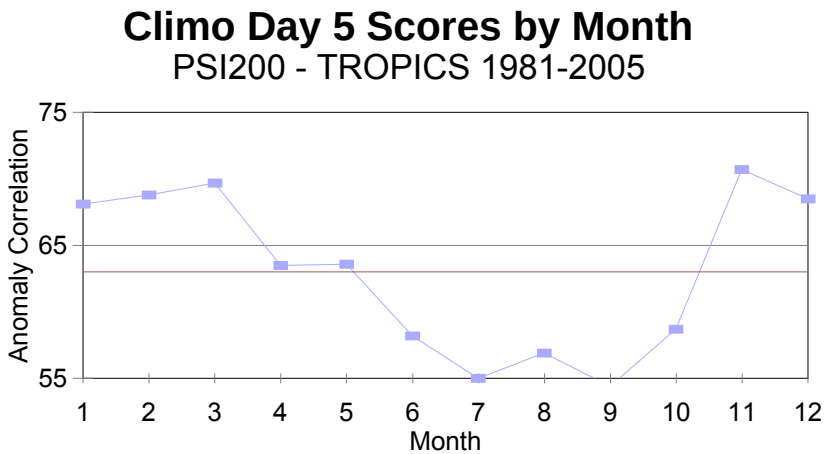


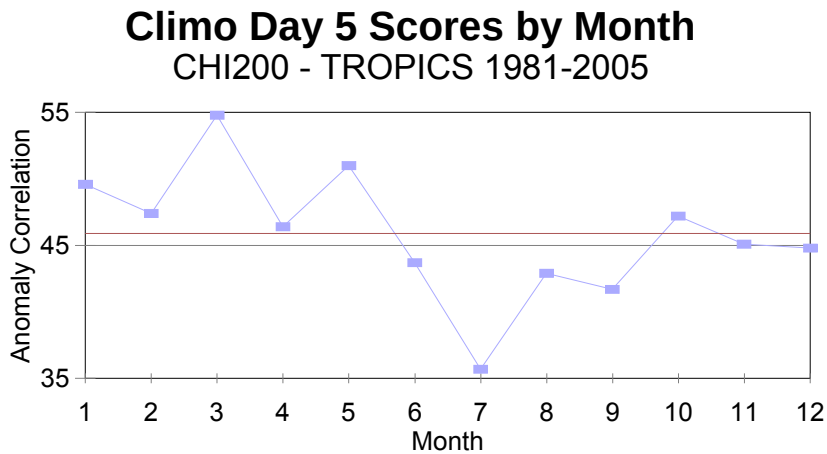
The SH does NOT have a clear annual variation in skill each year.

15 forecasts per month









forecasts verifying in month shown

375 forecast per dot(month)

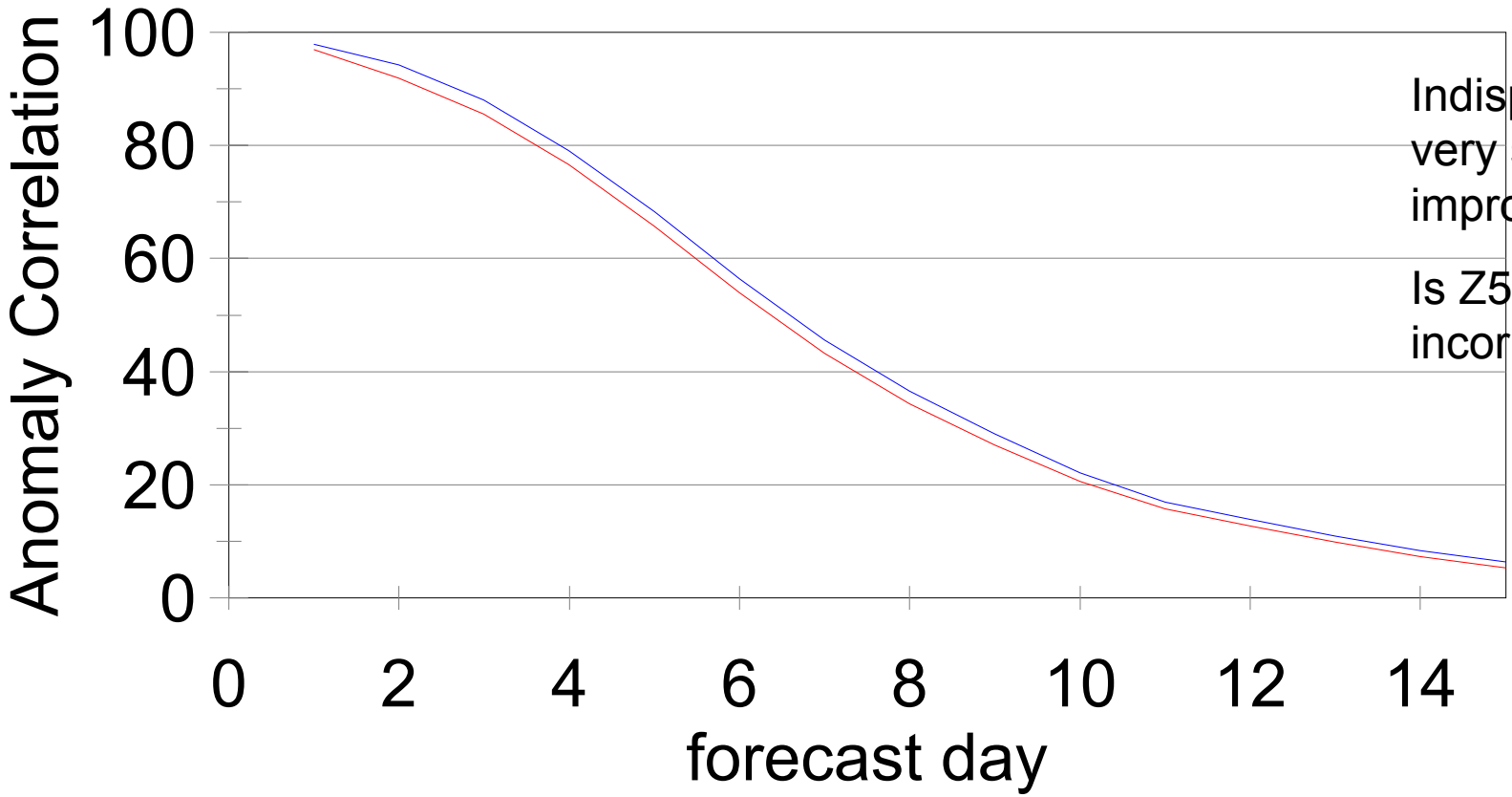
Large annual variation in the tropics! But volatile.

How about Bias Correction??? One of the claimed usages of hindcasts

Daily CFS Scores

Z500 1981-2005 NH July

Based on 375 forecasts.



Indisputable but very small improvement.

Is Z500 incorrigeable?

The looks of a die-off curve

raw

bias corrected

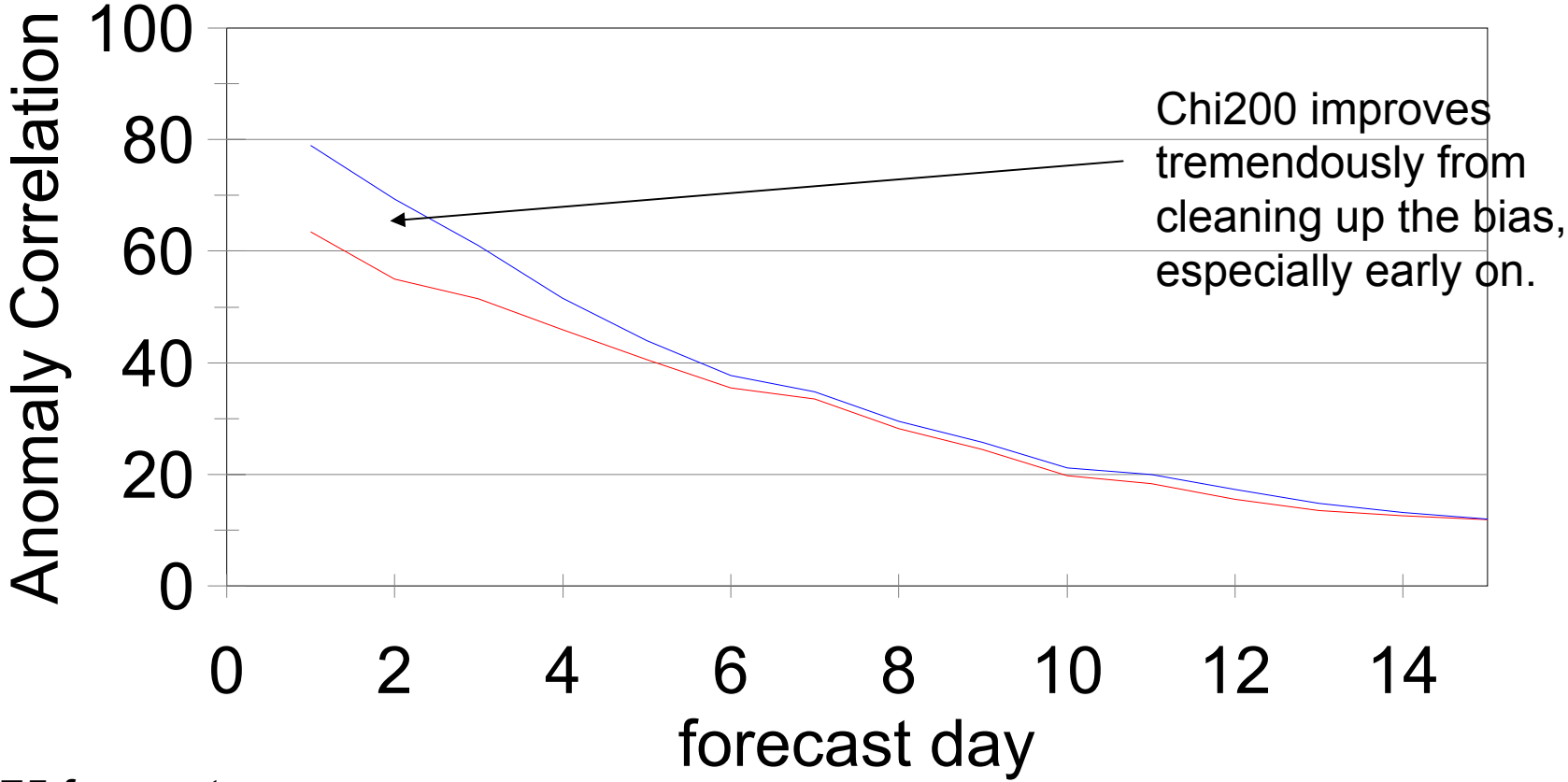
The gain due to bias correction in a few selected months. Day 5 scores Z500 1981-2005

	Raw	Bias Corrected
NH DEC	73.0	74.5 (+1.5)
NH JUL	65.8	68.3 (+2.5)
SH FEB	61.8	63.5 (+1.7)
SH AUG	62.2	63.7 (+1.5)

Is Z500 incorrigeable? Largely Yes, because the systematic error is small.
Wait till you see CHI200

Daily CFS Scores

CHI200 1981-2005 TR December



375 forecasts

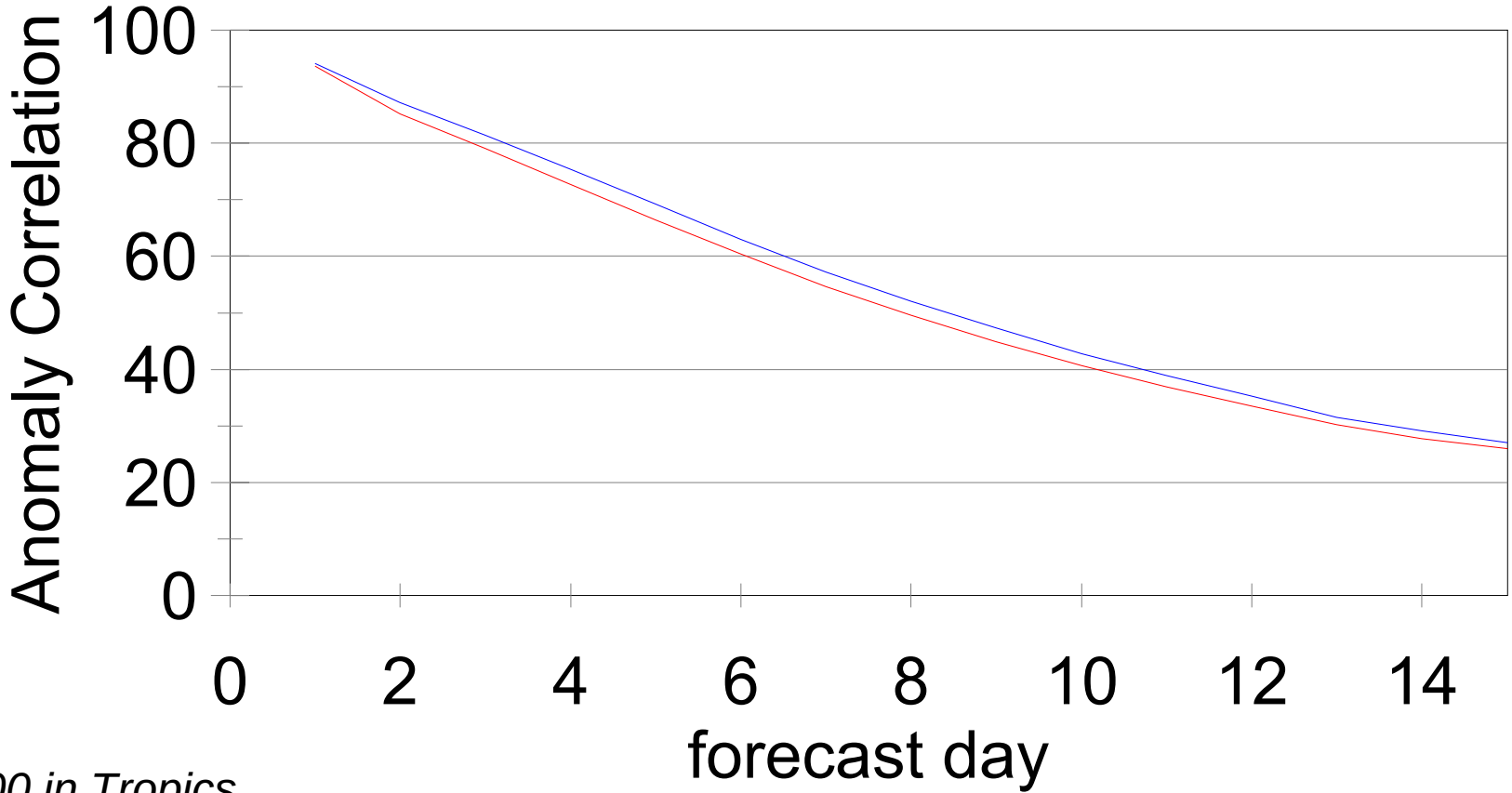
Can we do MJO forecasts ??

raw

bias corrected

Daily CFS Scores

PSI200 1981-2005 TR February



PSI200 in Tropics does not improve very much from bias removal. !!!!

The looks of a die-off curve

raw

bias corrected

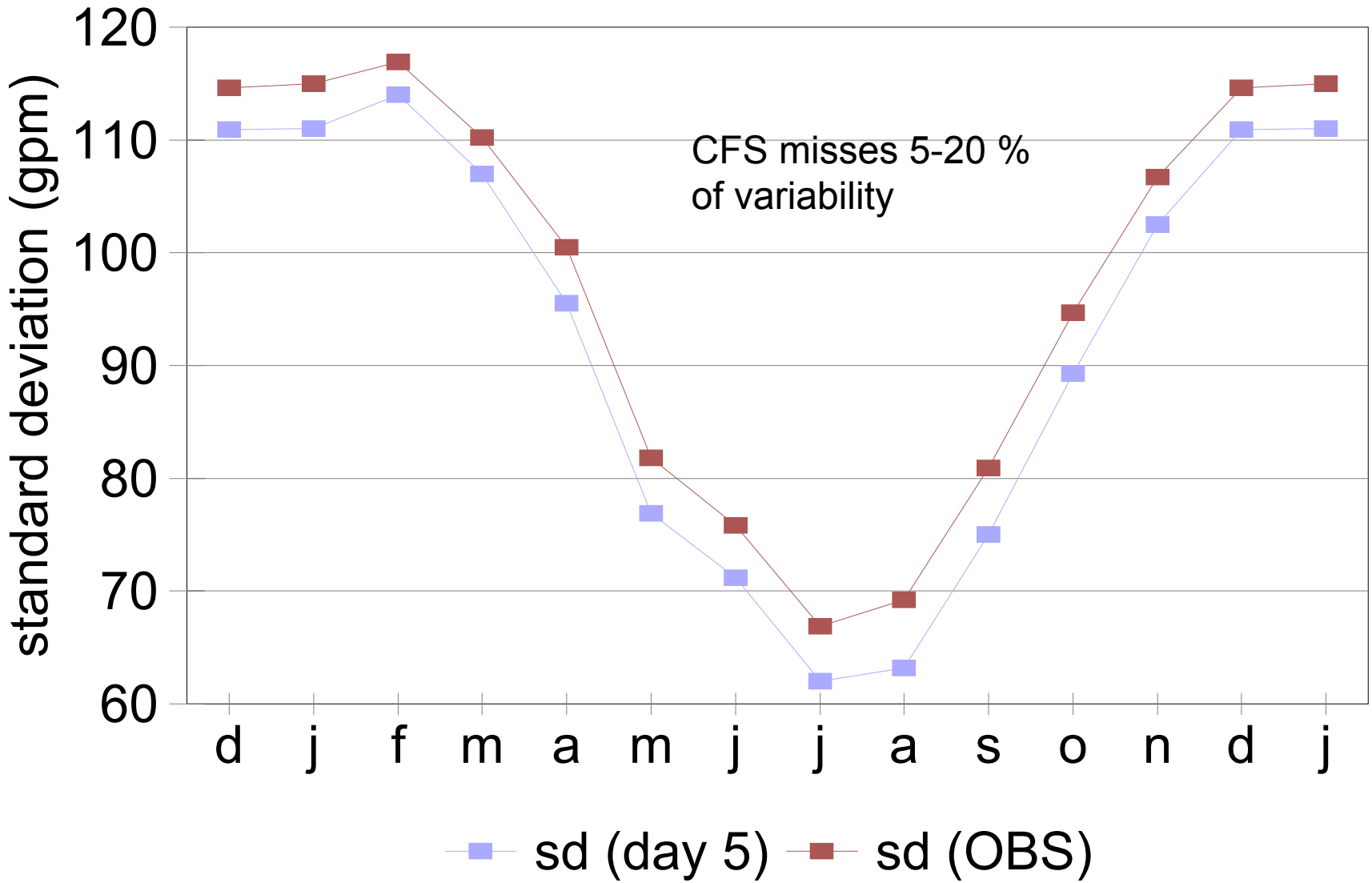
Day 5 scores

- AC (obviously; already shown),

But also:

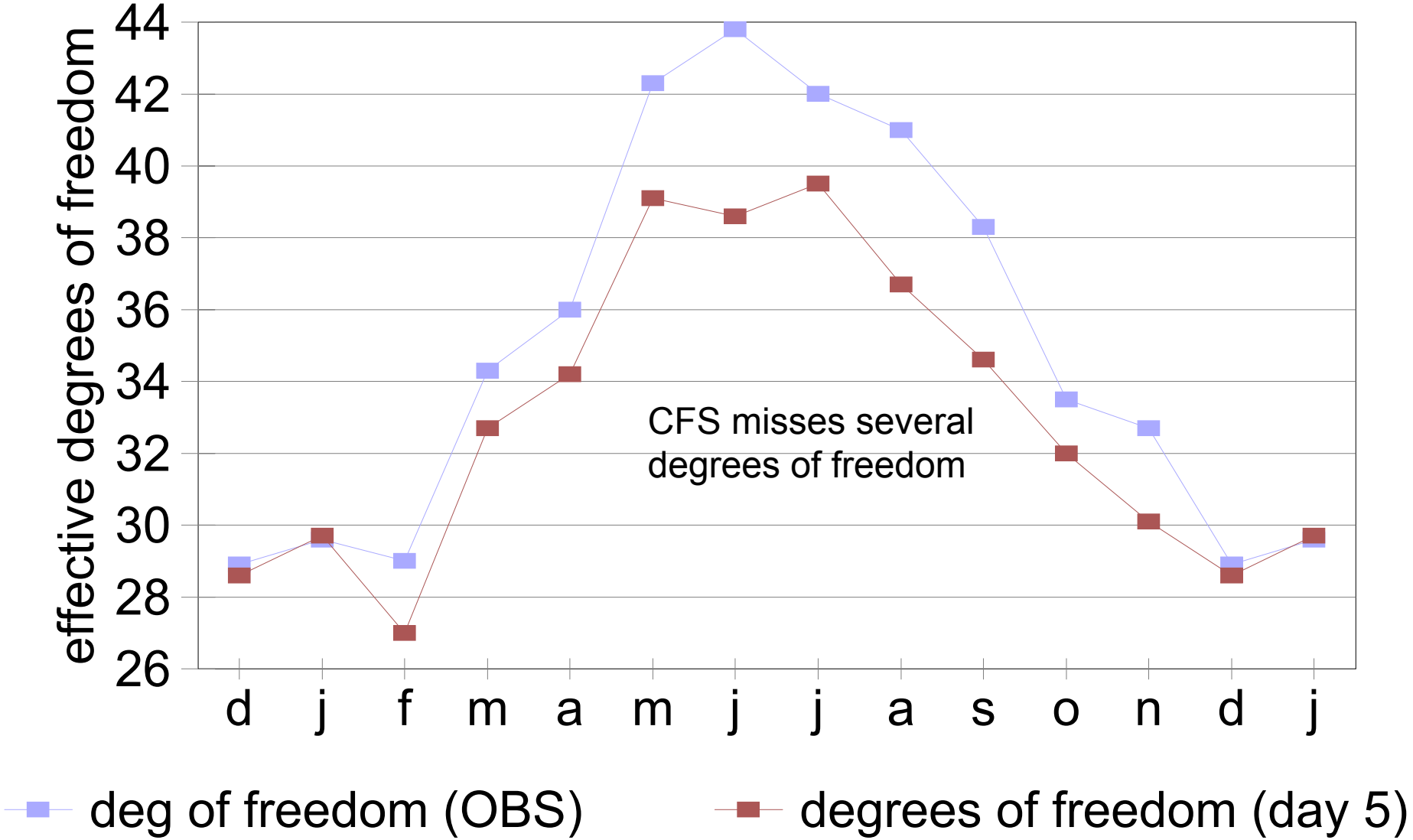
- SD
- eDOF

CFS daily scores 1981-2005 Z500 NH



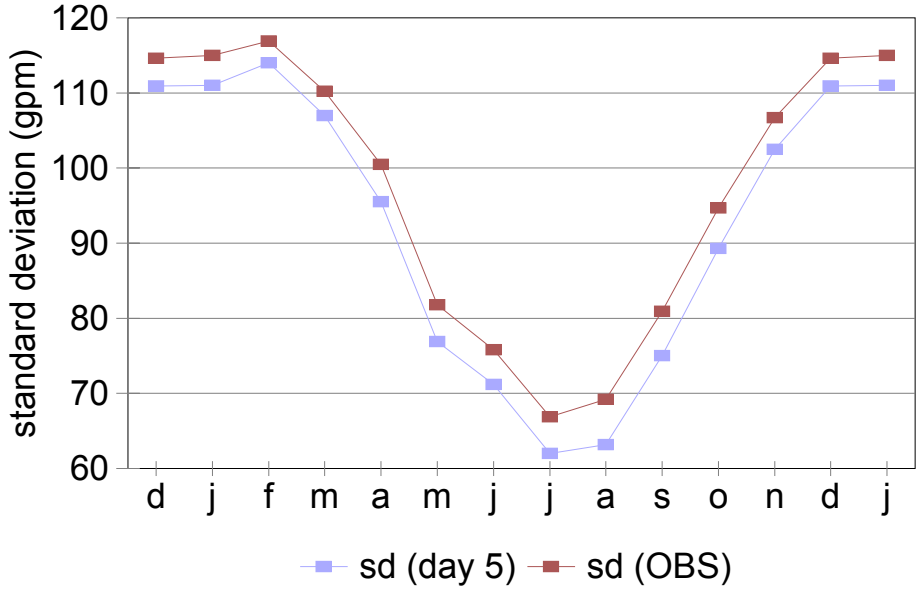
Loss of variance does not increase beyond 10-15 days and is never more than 20%

CFS daily scores 1981-2005 Z500 NH

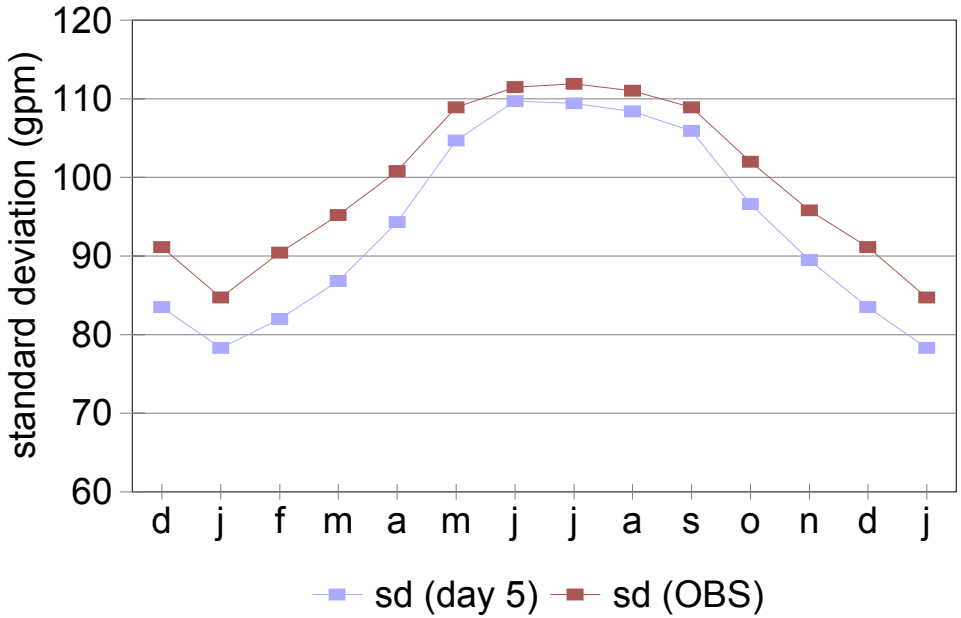


Loss of dof does not increase beyond 10-15 days and is never more than shown above

CFS daily scores 1981-2005 Z500 NH



CFS daily scores 1981-2005 Z500 SH



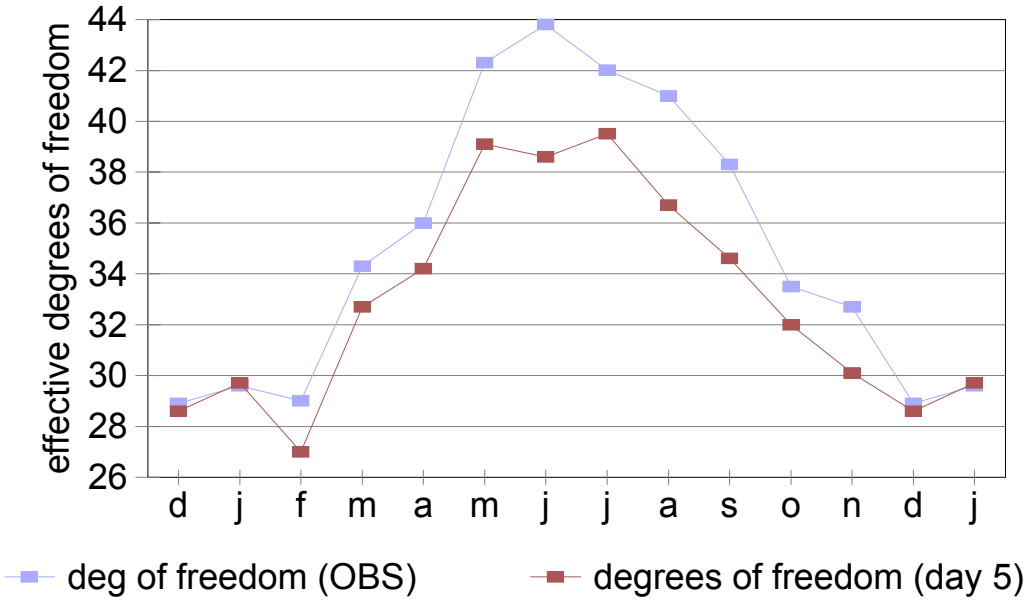
Compare the SD reduction in the

NH

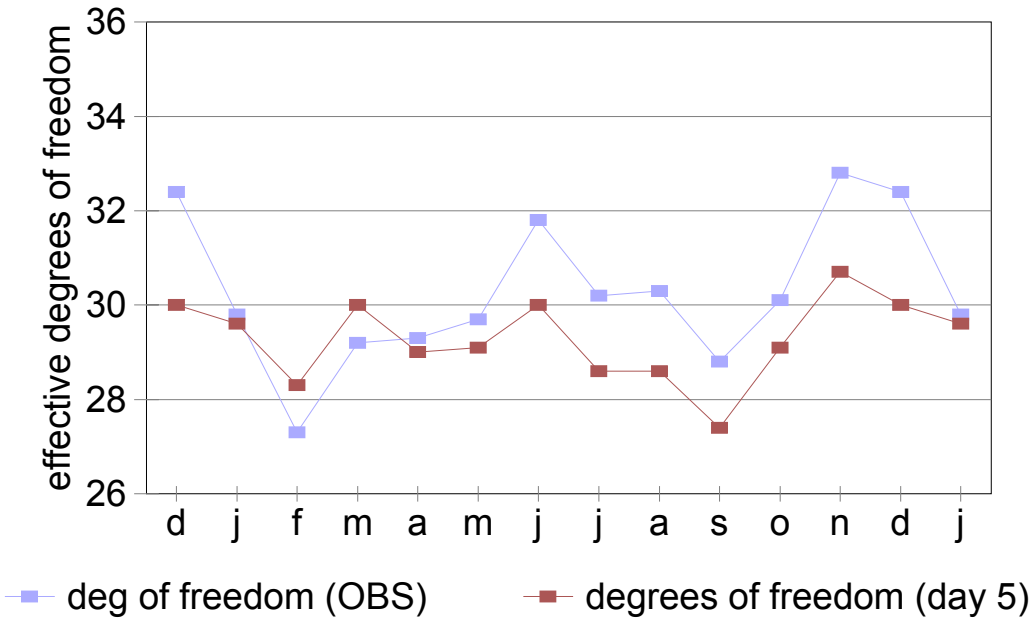
To that in the

SH

CFS daily scores 1981-2005 Z500 NH



CFS daily scores 1981-2005 Z500 SH



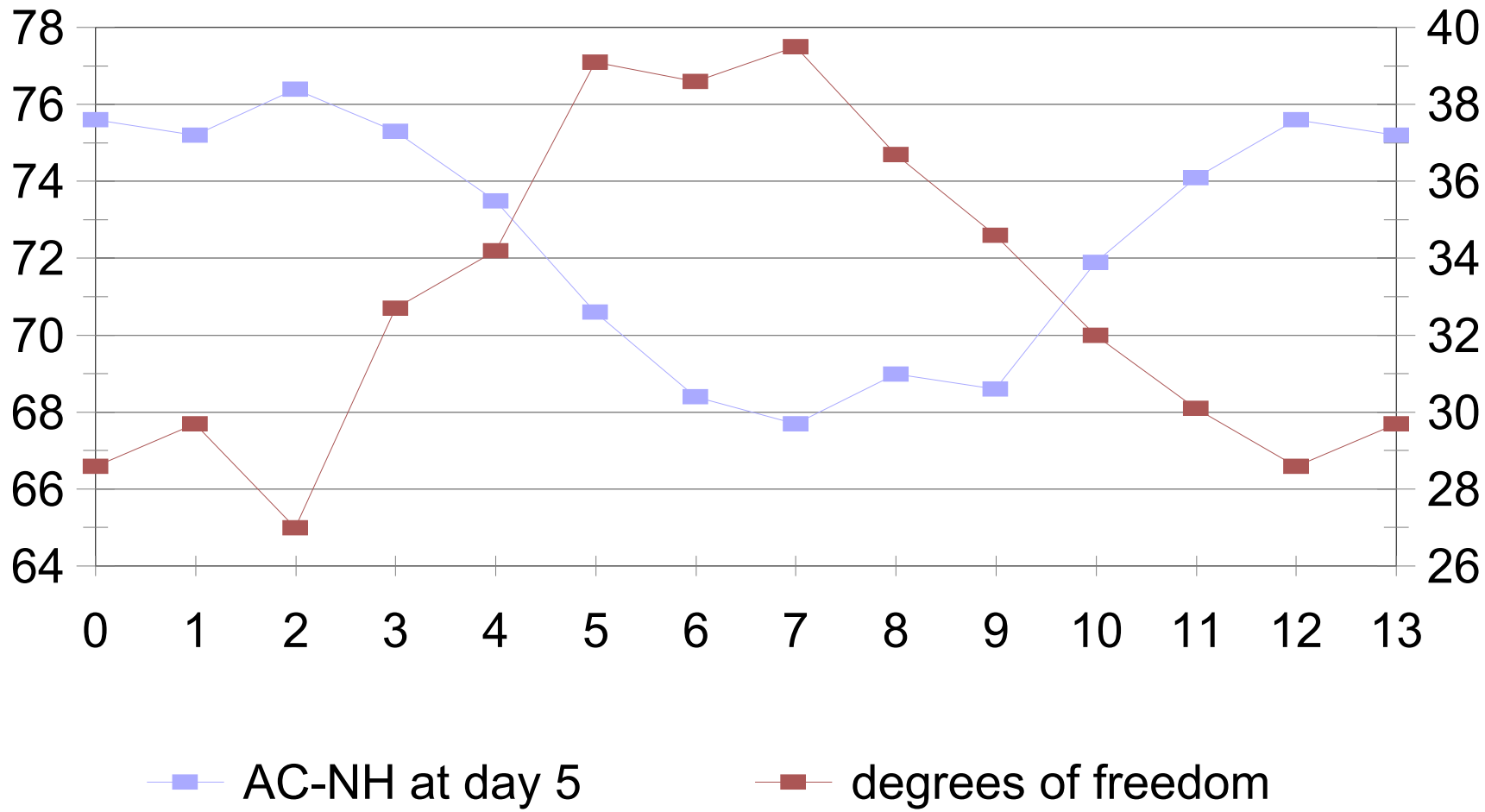
Compare the DOF reduction in the

NH

To that in the

SH

CFS daily scores 1981-2005 Z500 NH



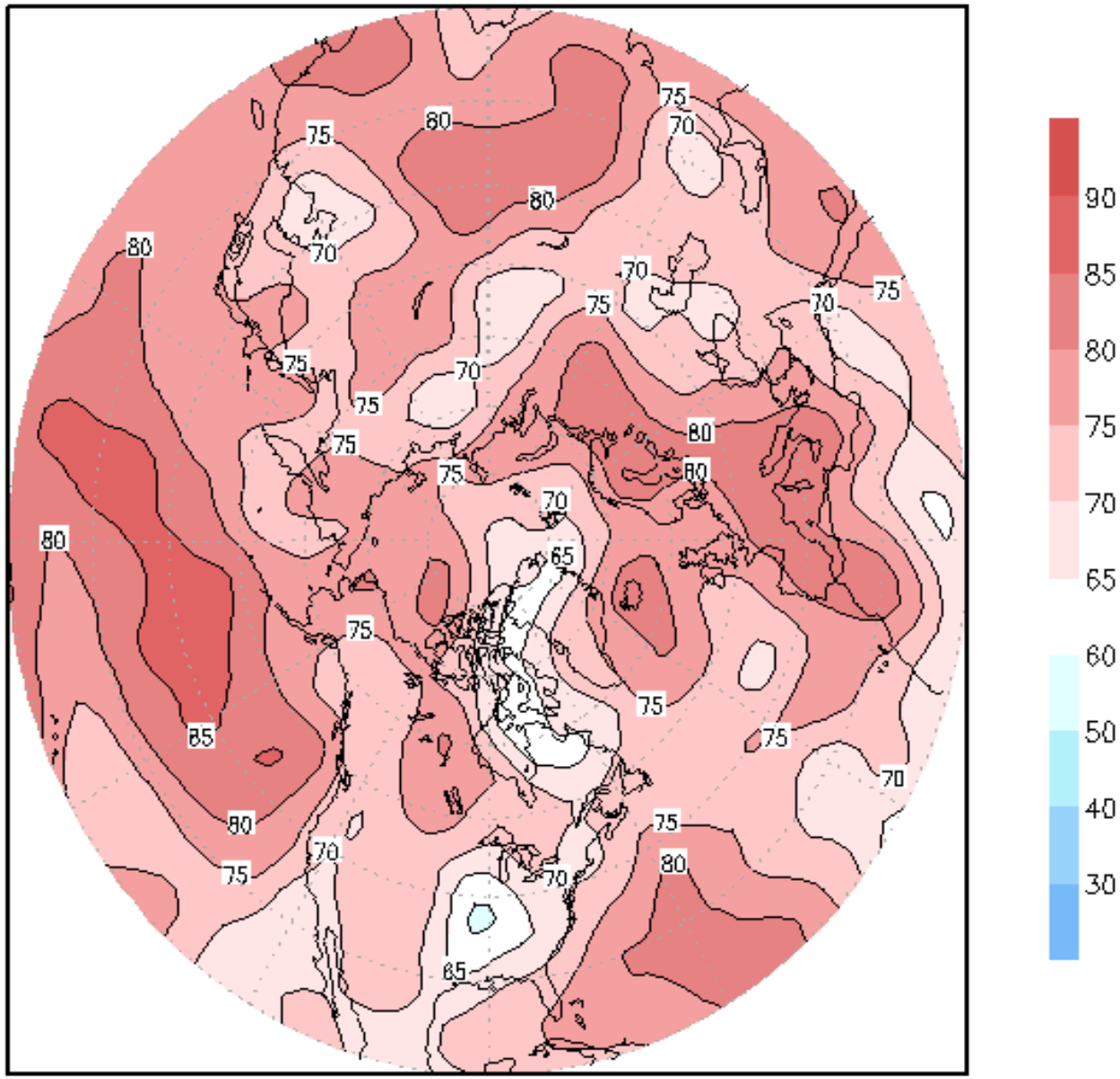
The climatological AC variation can be explained by eDOF and sd: low dof ←> high score, high sd ←> high score.

Distribution of Skill in space is important

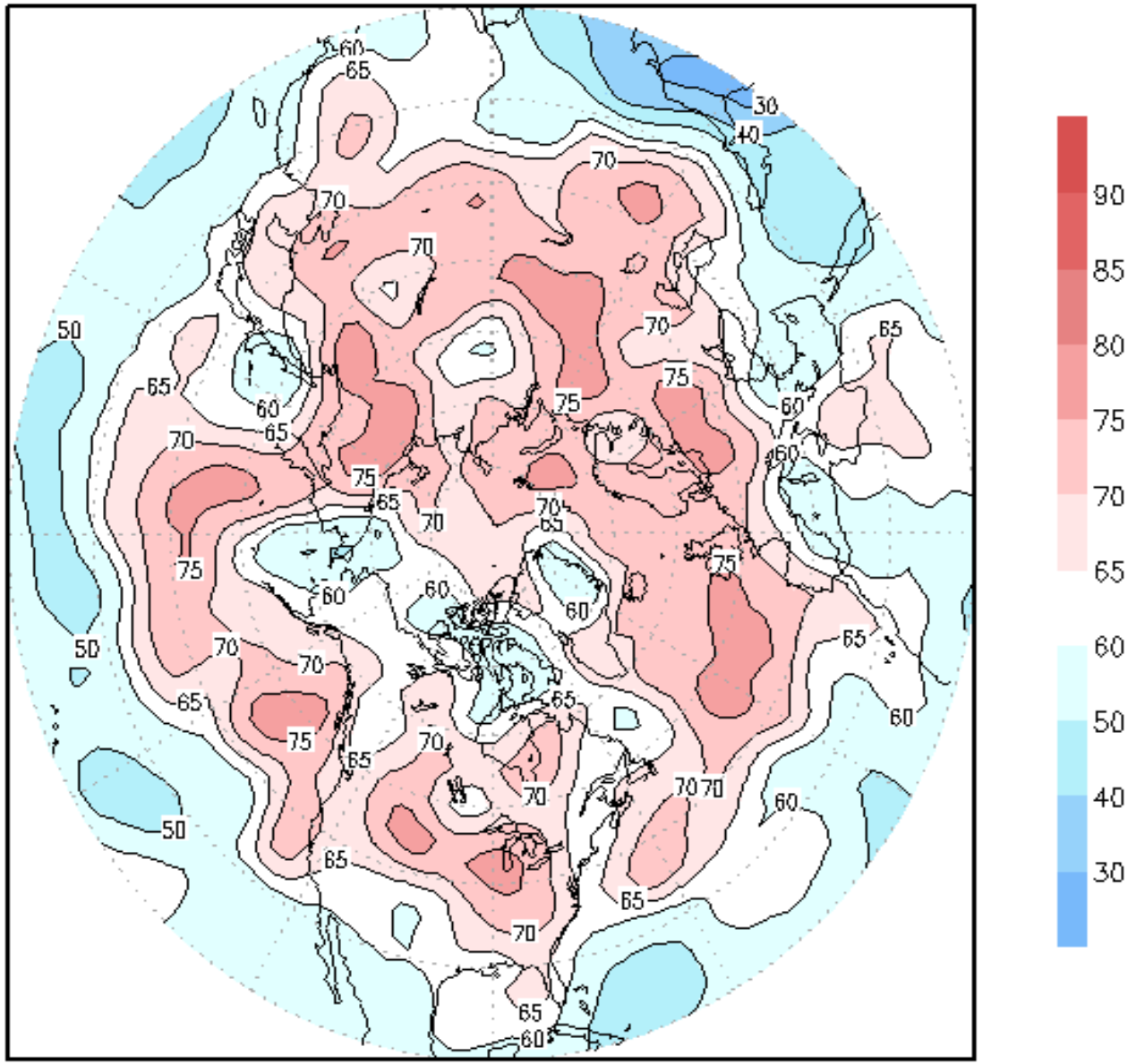
and largely unexplored and unexplained

(because we never had enough data)

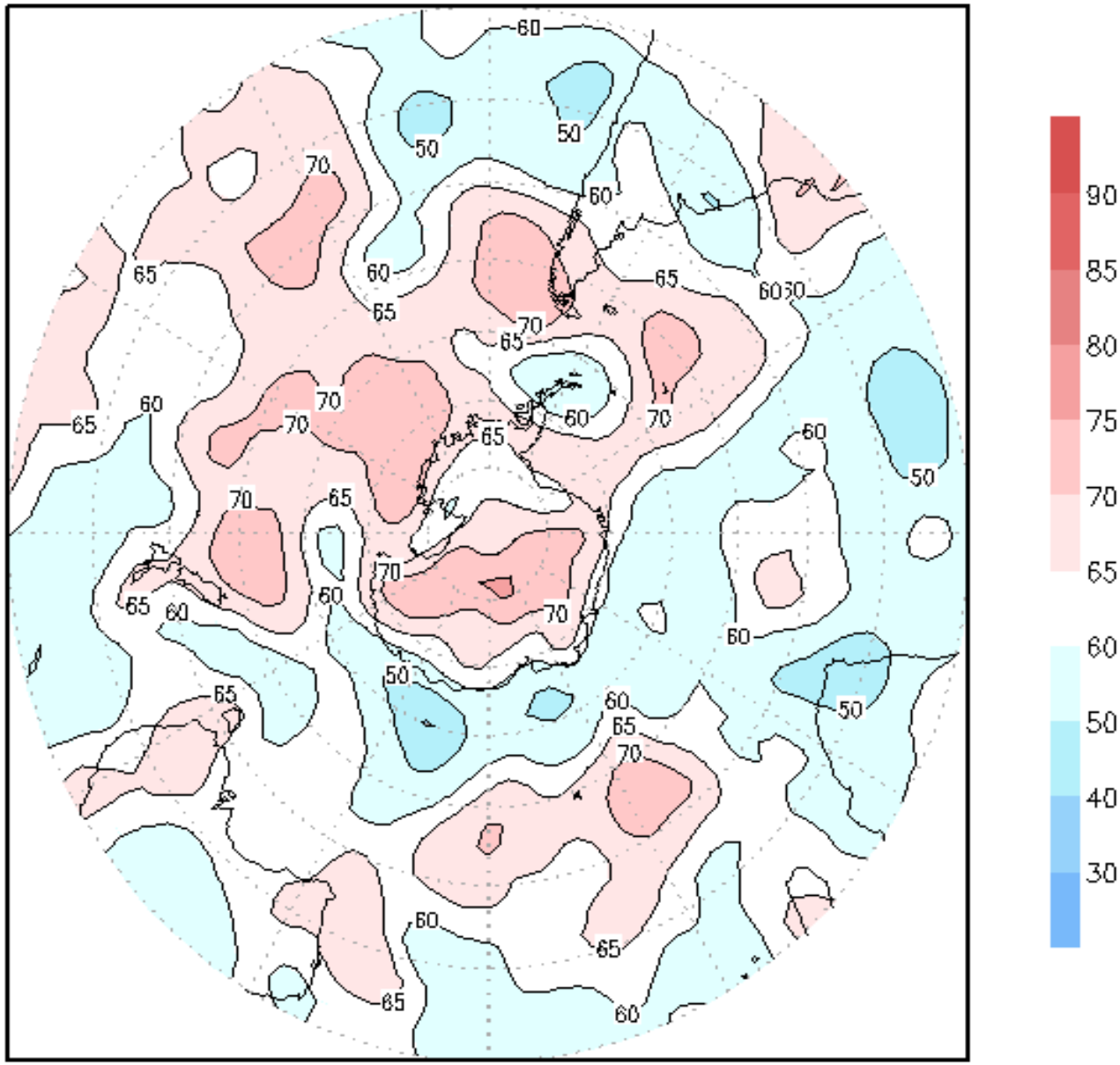
Skill (AC) NH 5-day Z500 CFS forecast for Feb IC



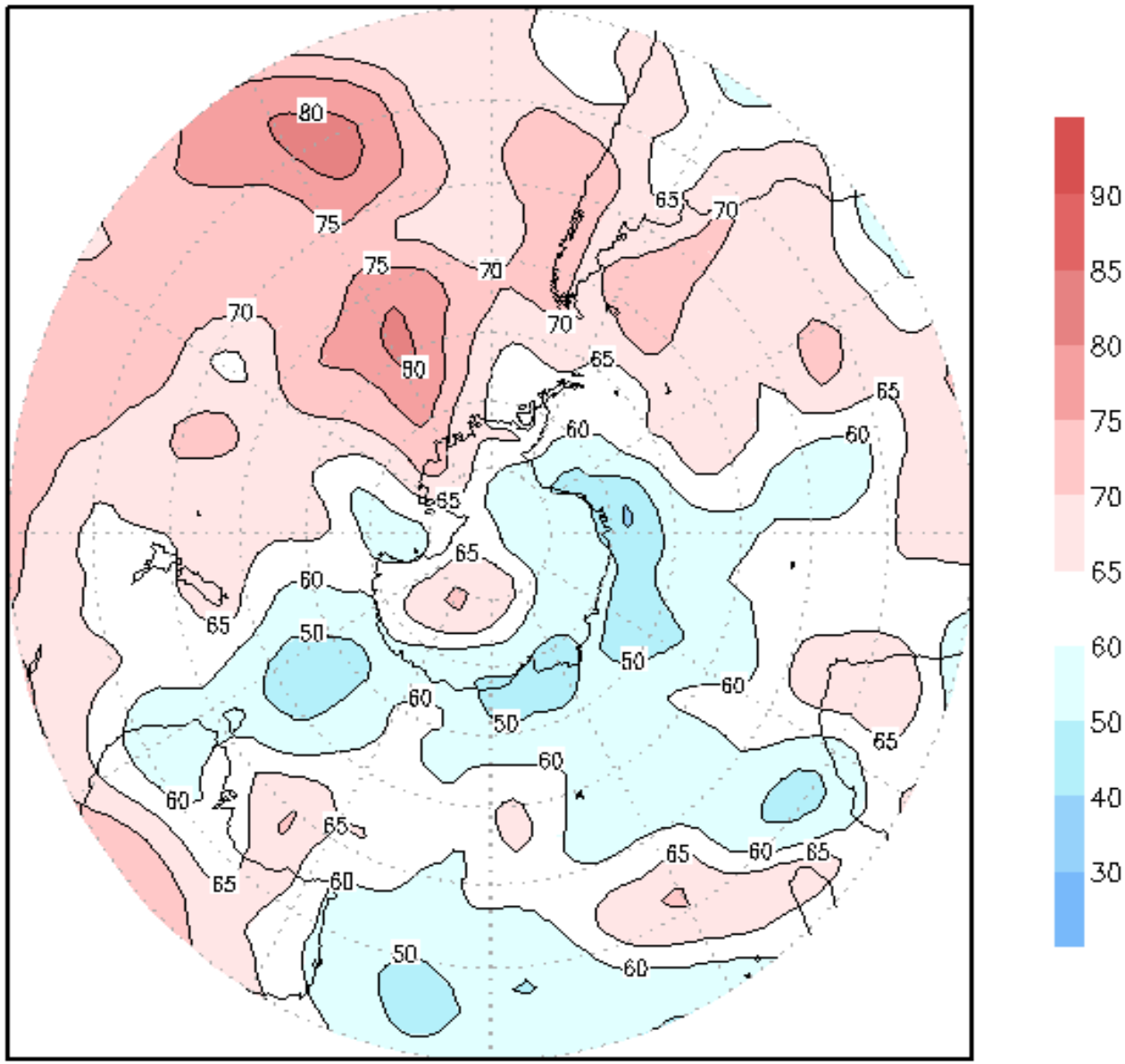
Skill (AC) NH 5-day Z500 CFS forecast for Aug IC



Skill (AC) SH 5-day Z500 CFS forecast for Feb IC



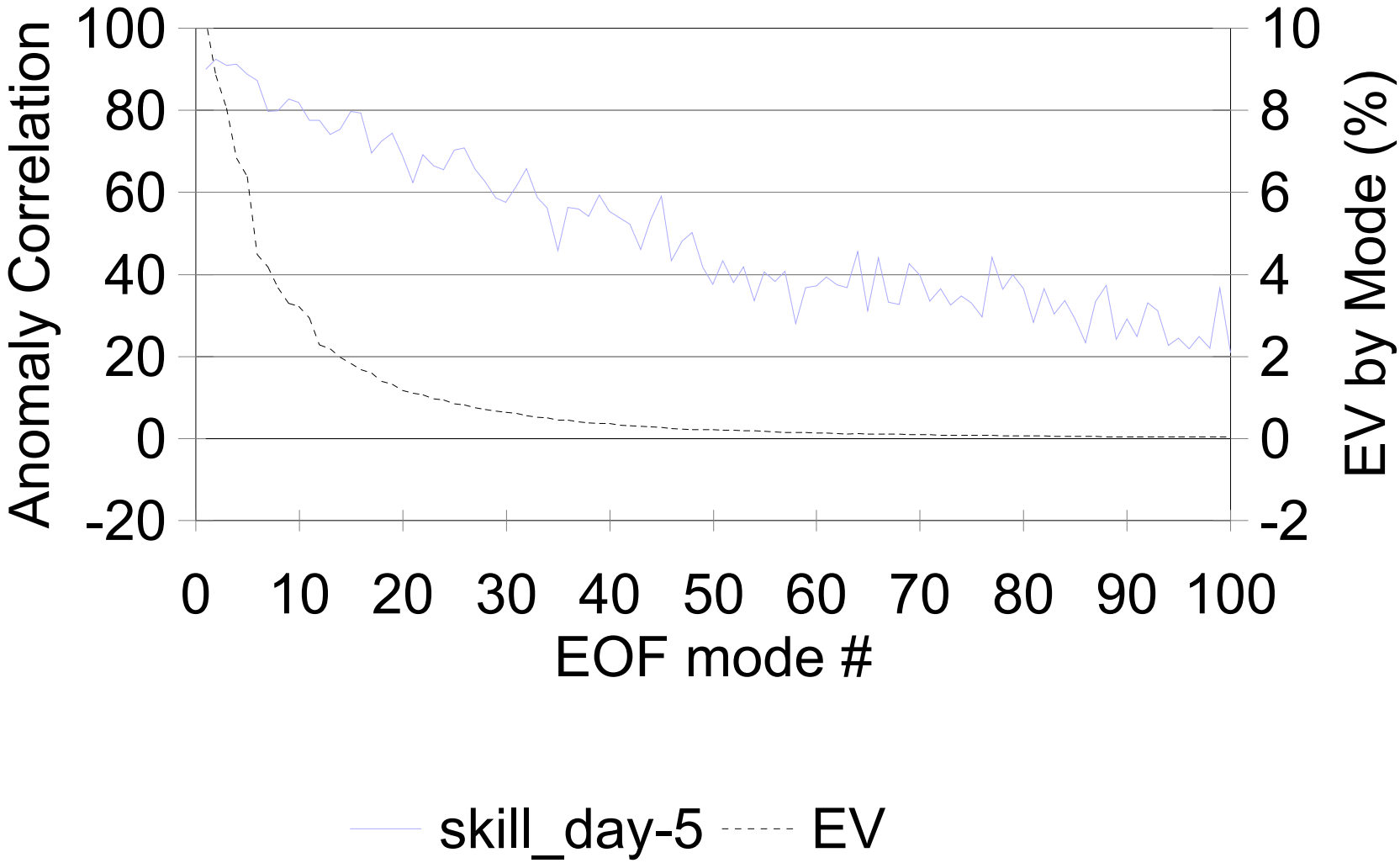
Skill (AC) SH 5-day Z500 CFS forecast for Aug IC



Skill as a function of (EOF) mode is interesting

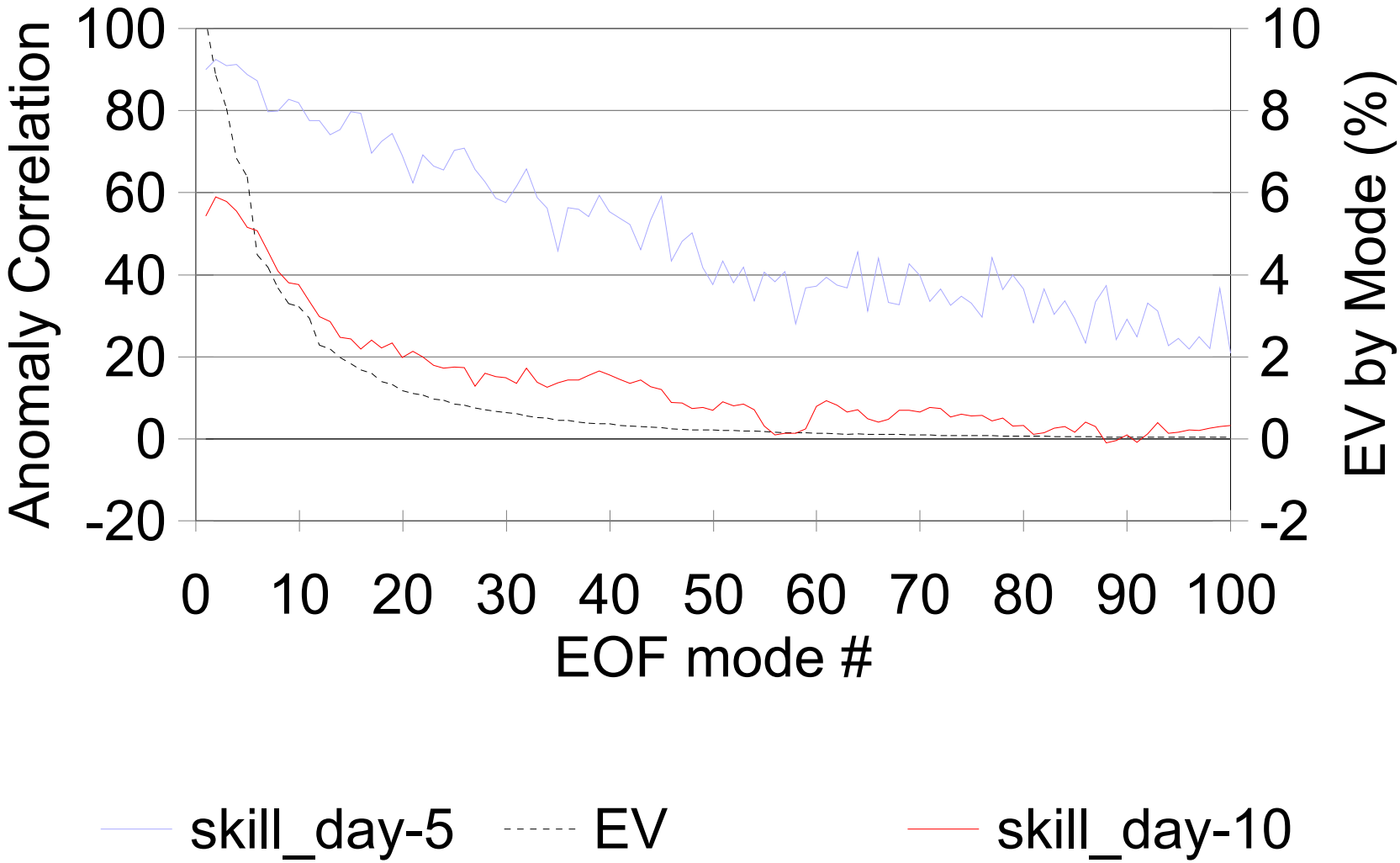
CFS Skill in Daily Z500 prediction

February 1981-2005 NH



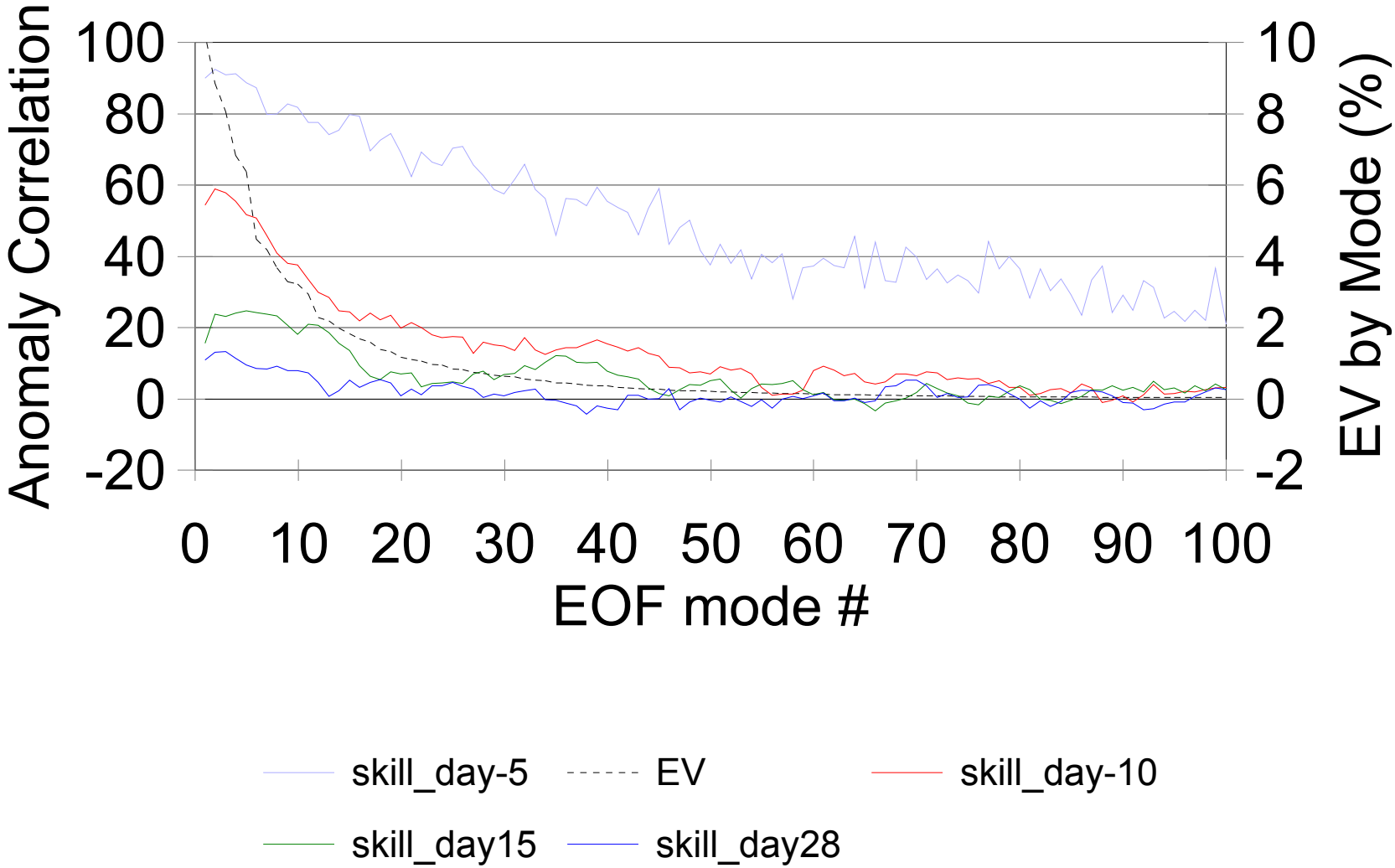
CFS Skill in Daily Z500 prediction

February 1981-2005 NH



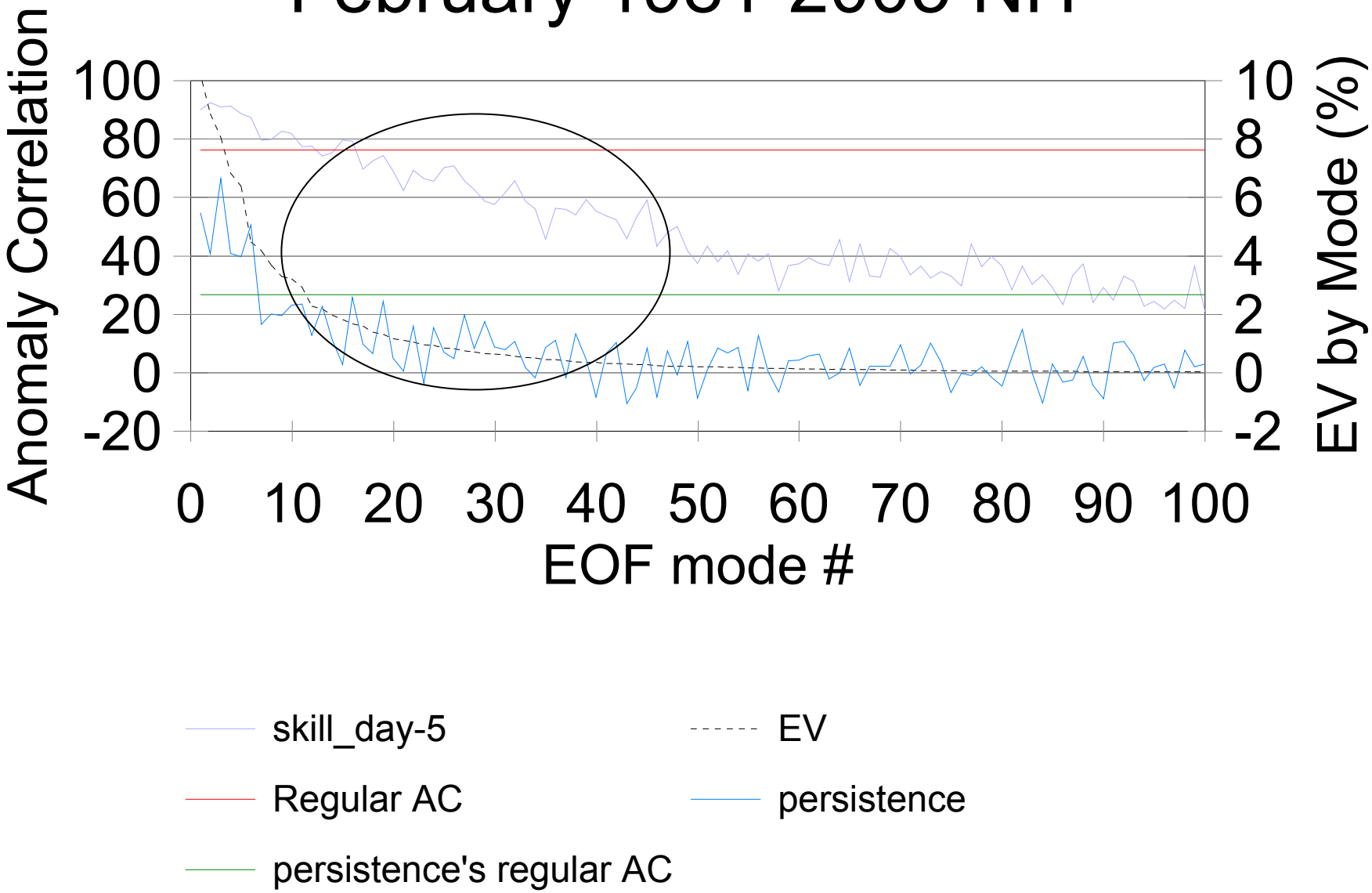
CFS Skill in Daily Z500 prediction

February 1981-2005 NH

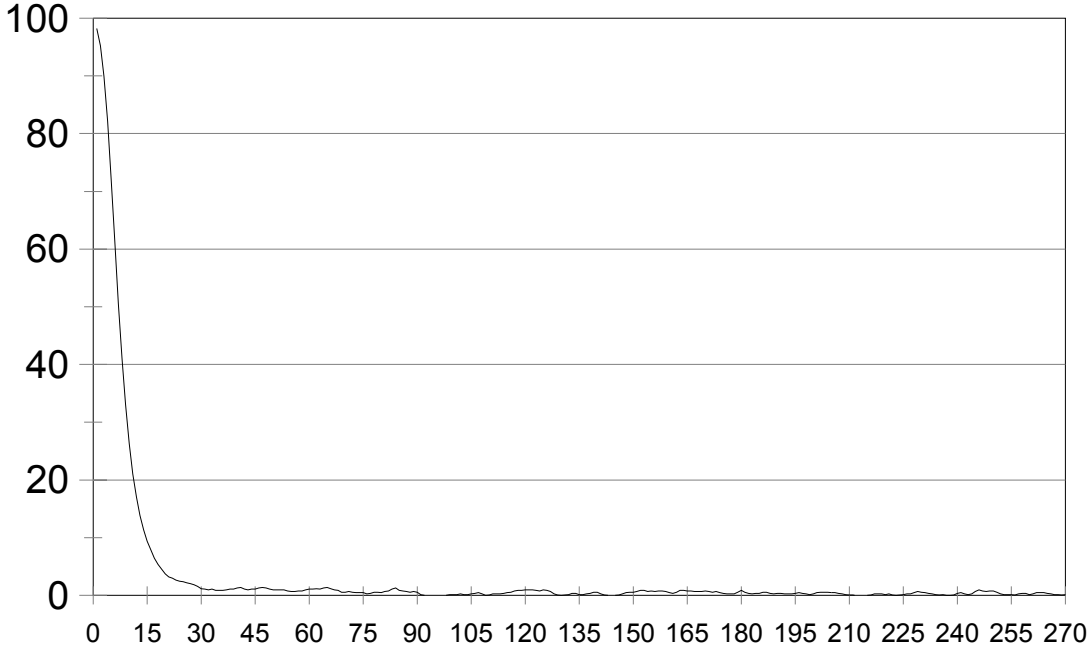


CFS Skill in Daily Z500 prediction

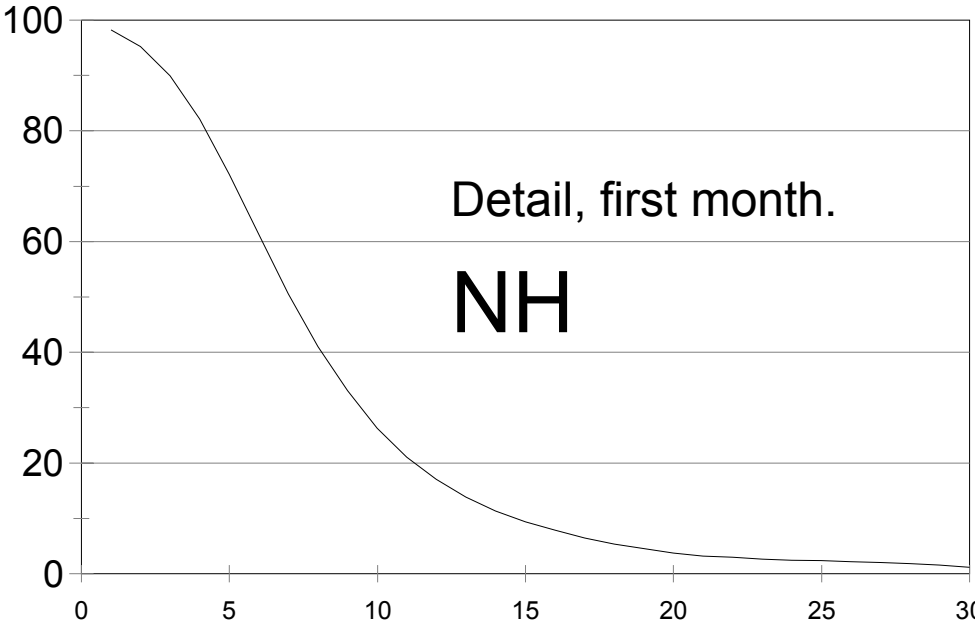
February 1981-2005 NH

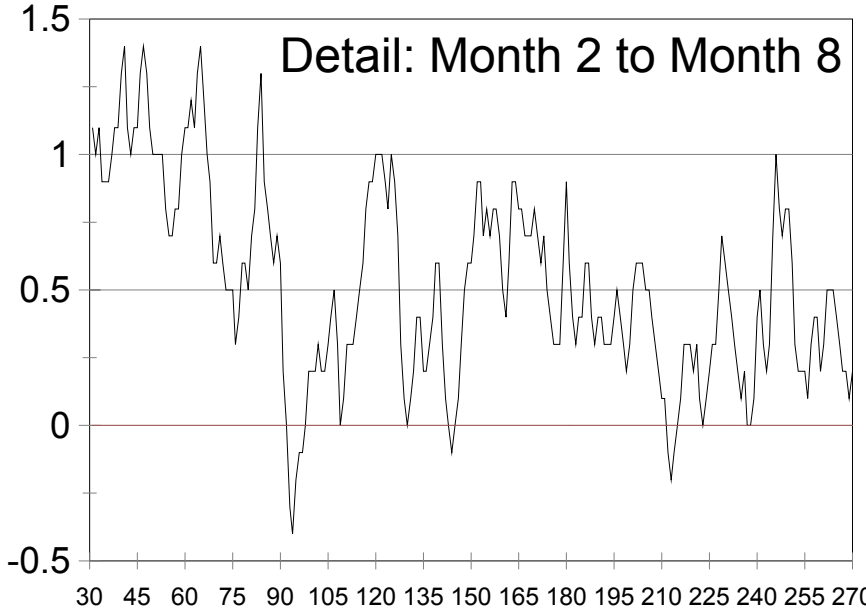


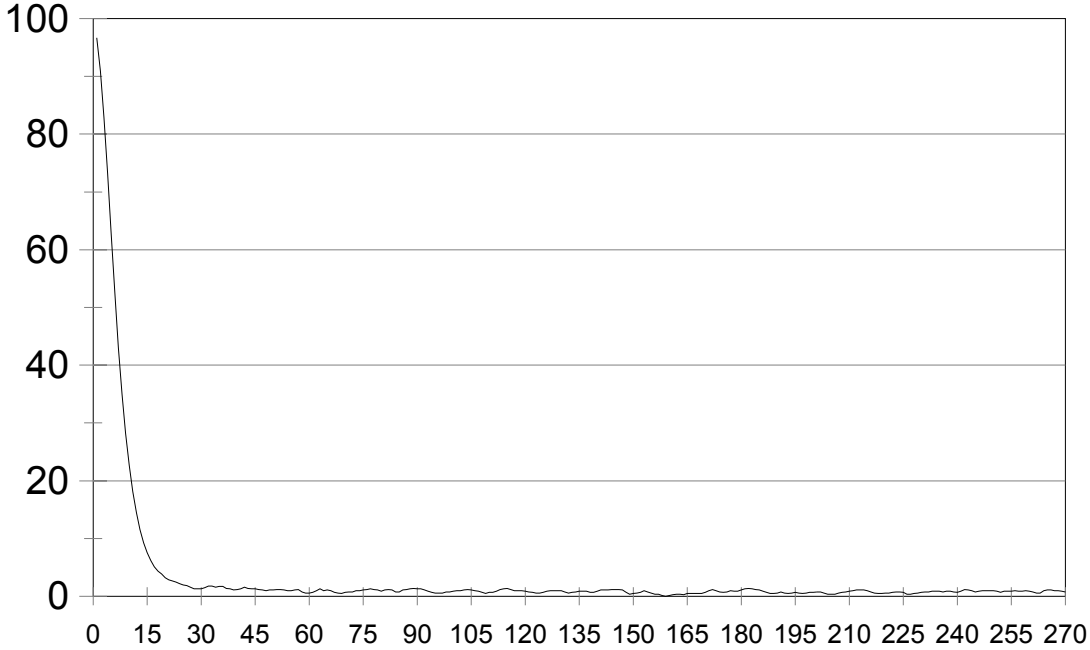
Now: OUT TO 270 DAYS !!



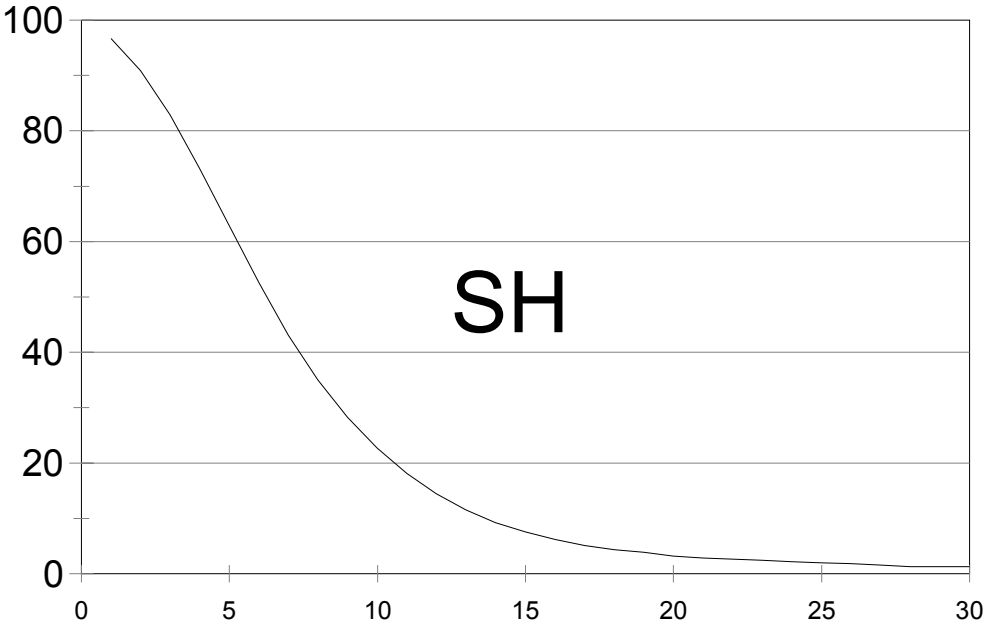
Die-off curve, Z500, NH, 1981-2005, 4500 forecasts, all seasons aggregated

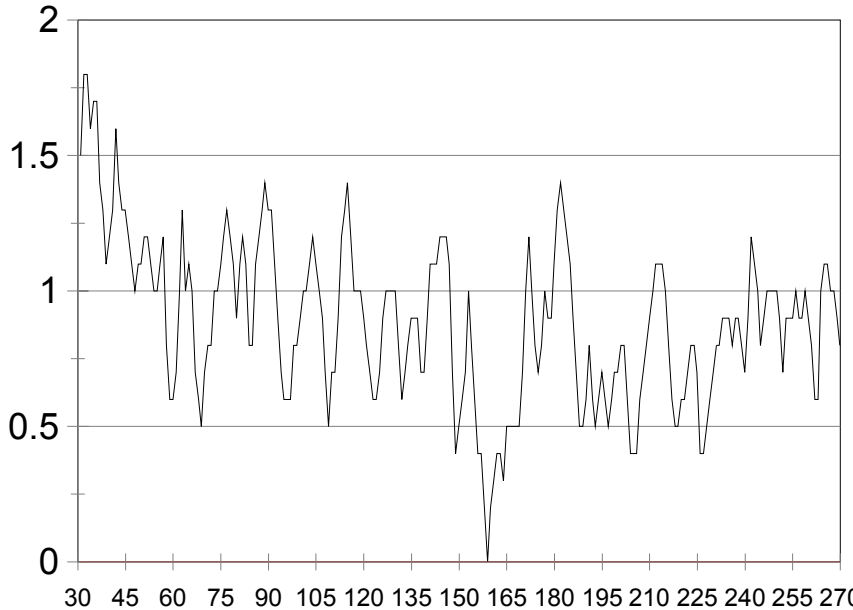






Die-off curve, Z500, SH, 1981-2005, 4500 forecasts, all seasons aggregated





FOCUS

- Week 3
- Week 4
- = days 15-21 and 22-28
- Physical basis of wk3/wk4
- Ocean interaction is as much a liability as a promise at this point in time.

	DAY 15			DAY 28	
	CFS	Pers		CFS	Pers
NH	15.2	4.2		6.1	3.9
TR	36.2	29.3		29.6	27.5
SH	9.6	6.2		2.0	3.4

Anomaly correlation of CFS and ‘Persistence’ Z500 prediction in Feb 1981-2005

$$\frac{\text{signal}}{\text{NOISE}} \quad \text{ratio}$$

Improving signal to noise ratio: (often cosmetic) by reducing noise by applying some operator and, hopefully, not hurting the signal by this operator

1) Take a time mean (7 days)

2) Ensemble mean (not done here)

3) EOF filters (or better (maximally predictable modes))

Effect of a 7 day mean……

	Day 15	Wk3	Wk4	Day 28	
NH	15.2	15.7	9.2	6.1	
TR	36.2	46.7	42.9	29.6	
SH	9.6	10.6	5.2	2.0	

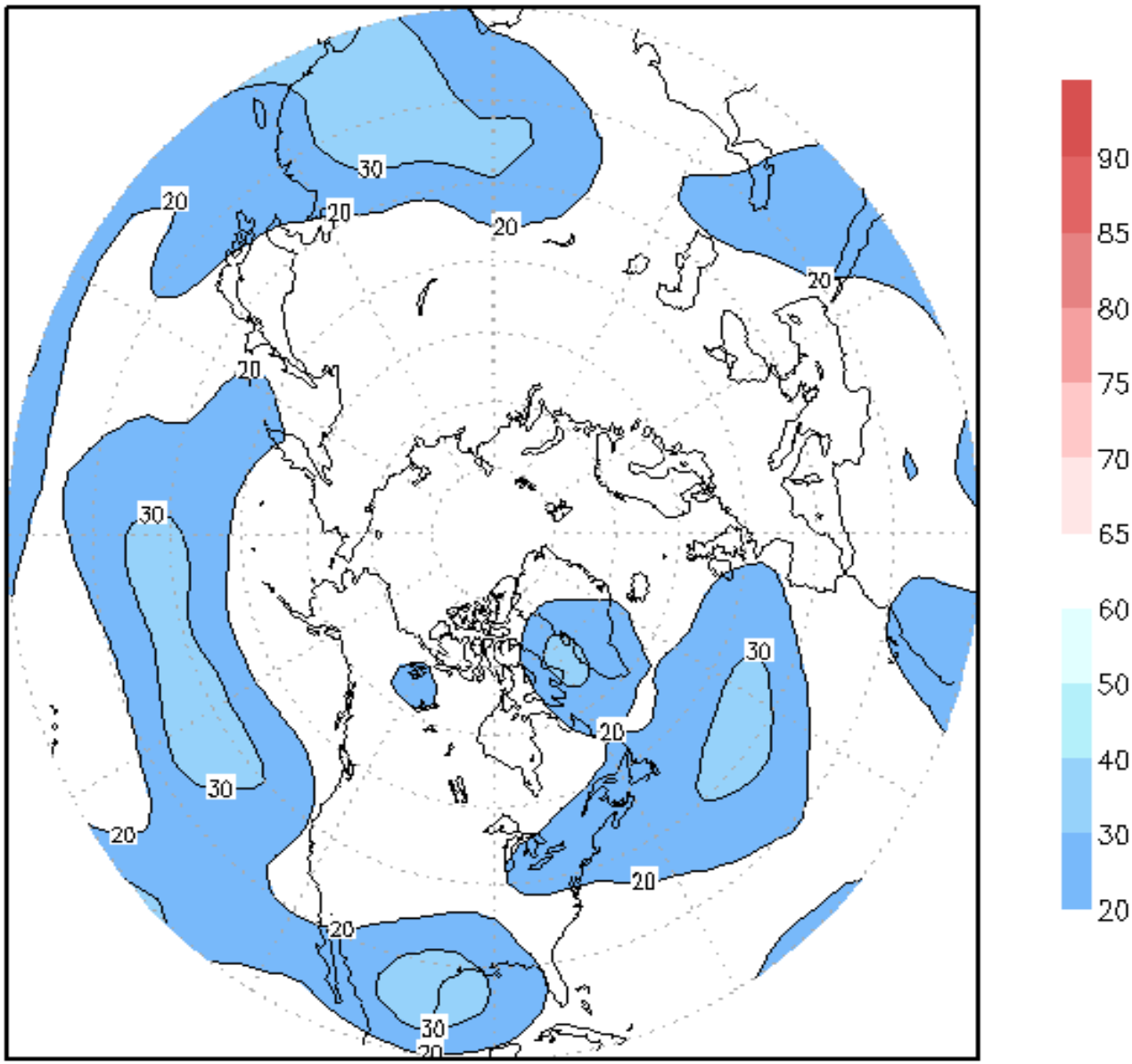
Anomaly correlation of CFS Z500 prediction, daily as well as weekly, in Feb 1981-2005

Effect of EOF truncation……

Wk3	Wk3 EOF1	WK3 EOF1-10	WK4	WK4 EOF1	WK4 EOF1-10	
15.7	22.0	18.7	9.2	23.9	12.0	

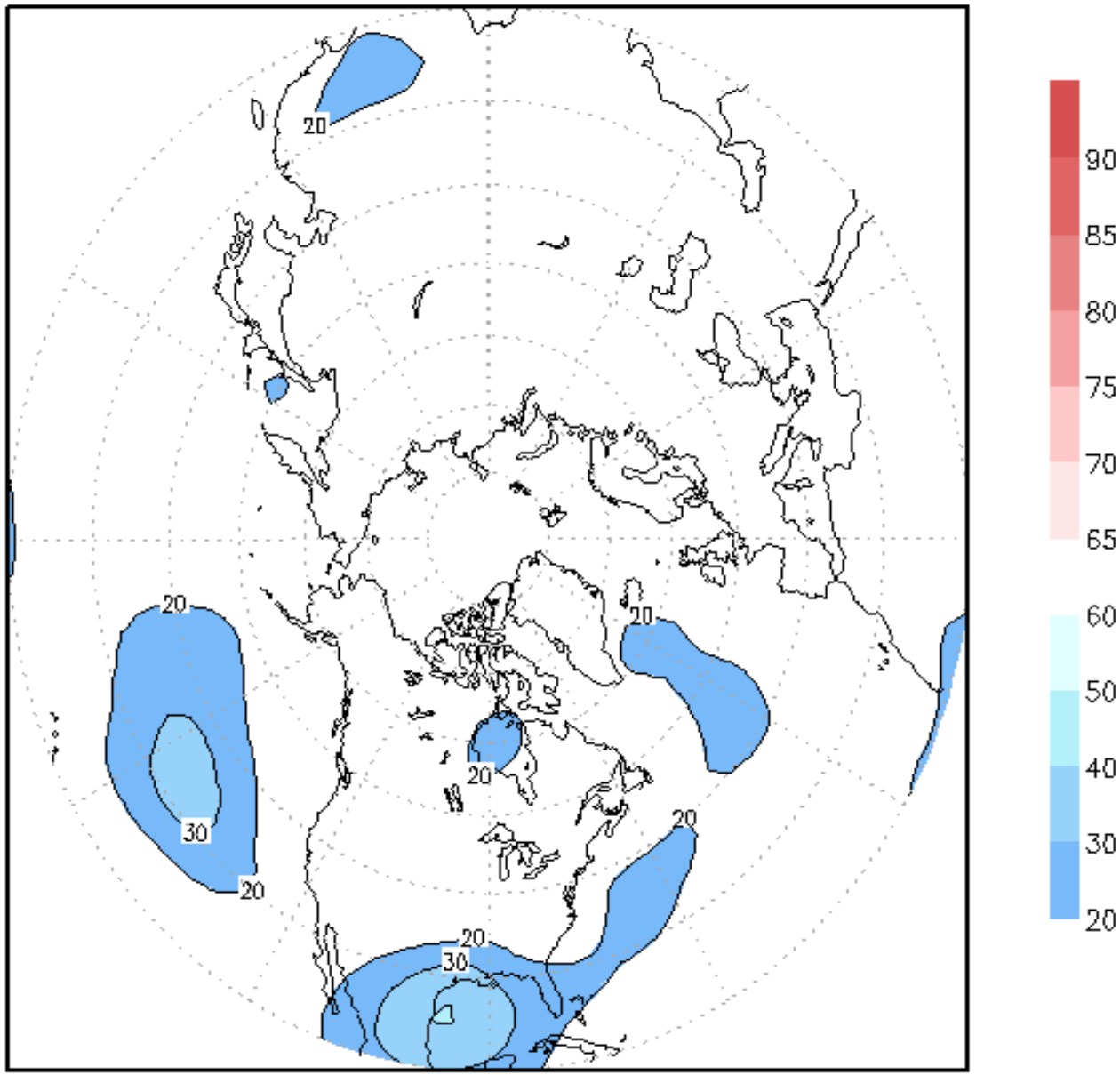
So, overall, we went from daily scores (15.2….6.1) to weekly scores (15.7 and 9.2) to filtered weekly scores (22.0 and 23.9 at best)

Skill (AC) NH wk3 Z500 CFS forecast for Feb IC



15.7

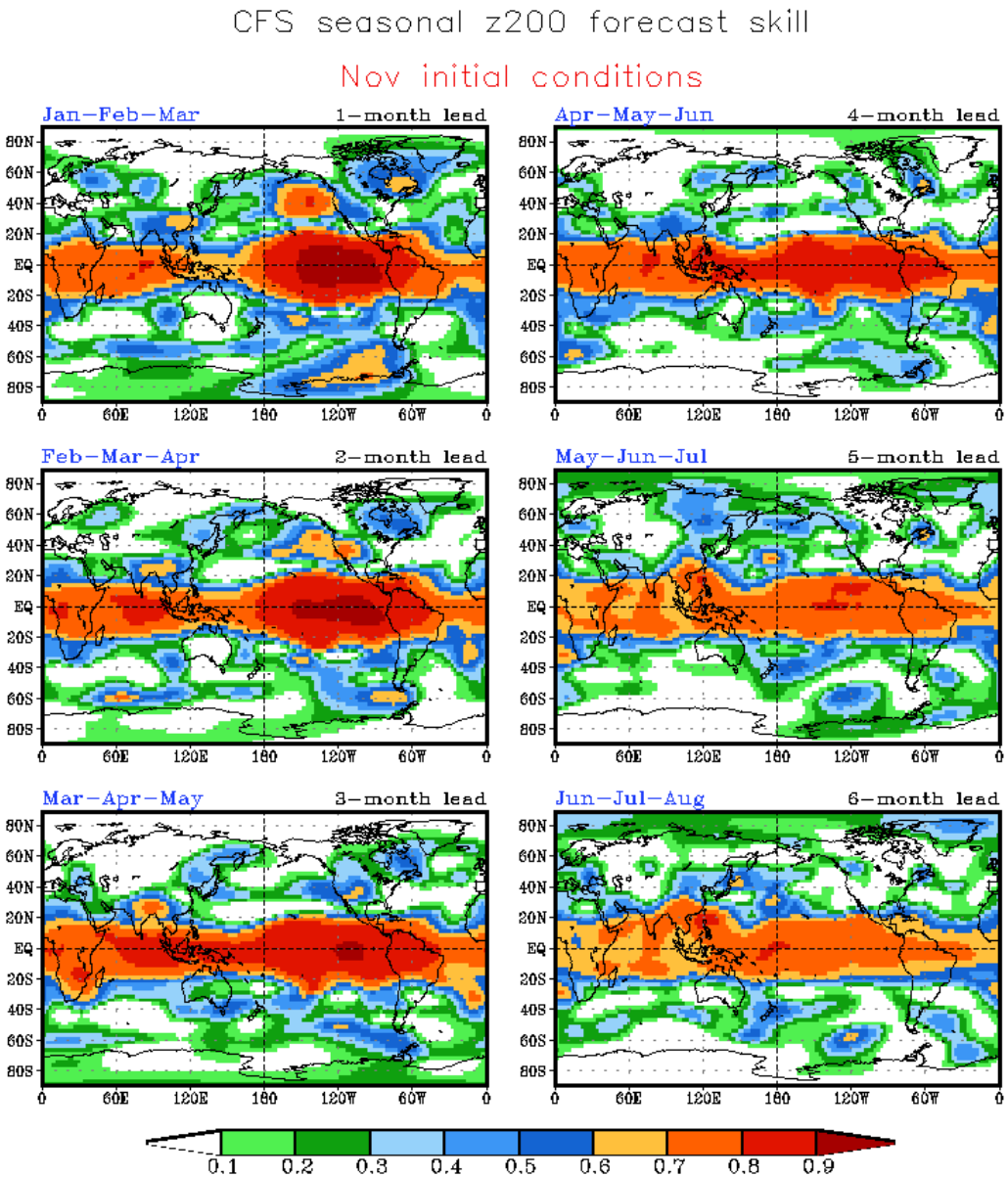
Skill (AC) NH wk4 Z500 CFS forecast for Feb IC



9.2

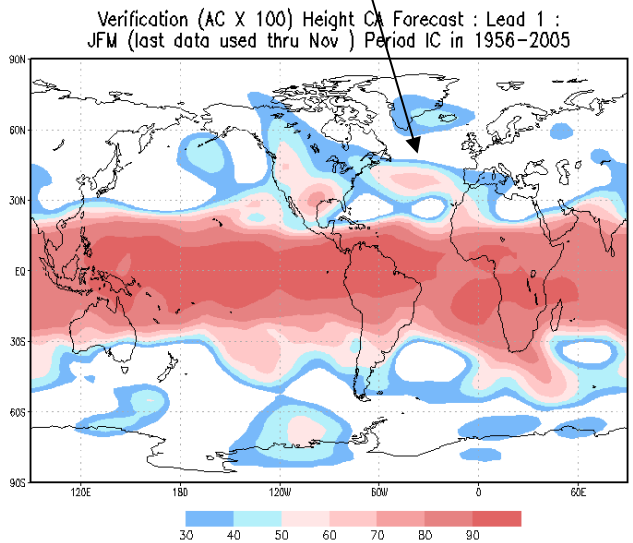
NWS/NCEP

Longer time average, and N member ensemble average helps.



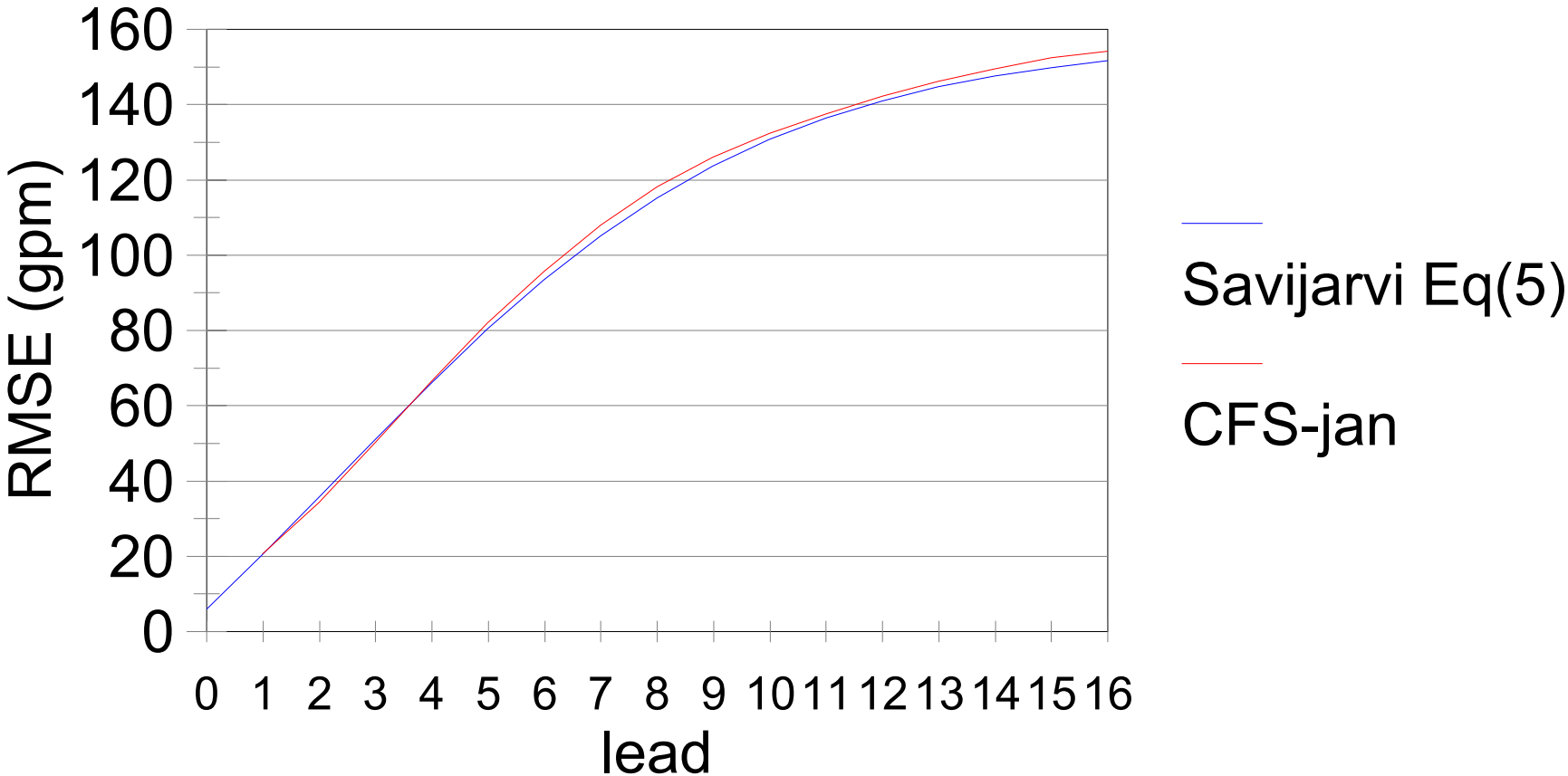
From Wanqiu Wang's page.

Constructed Analogue



RMSE growth CFS NH

NOSEC



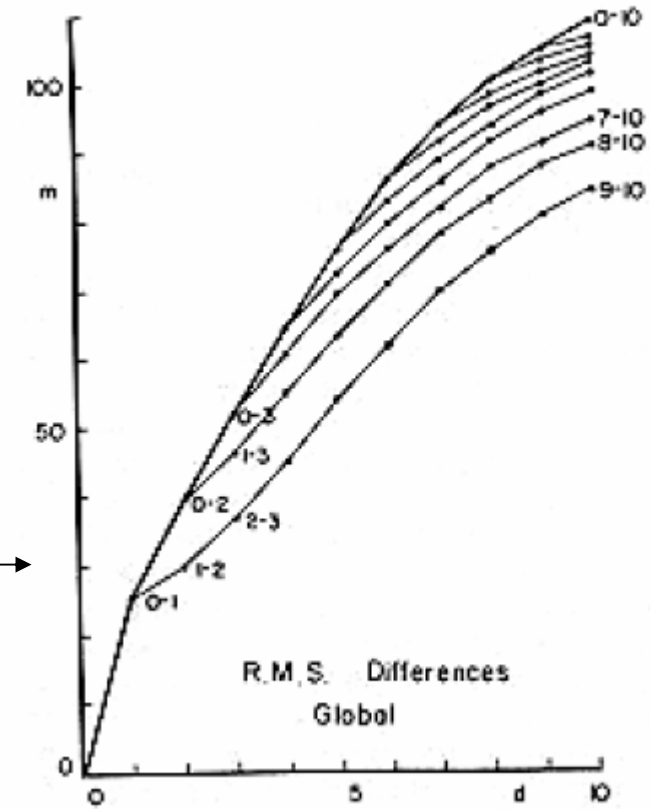
Initial error=6.1 gpm ; Systematic error growth s=13.5 gpm/day ; Small error amplification a=0.181/day ; e-infinity=157.7gpm ; NOSEC

$$\frac{de}{dt} = (ae + s)\left(1 - \frac{e}{e_i}\right), \qquad (5)$$

Savijarvi's equation 5:

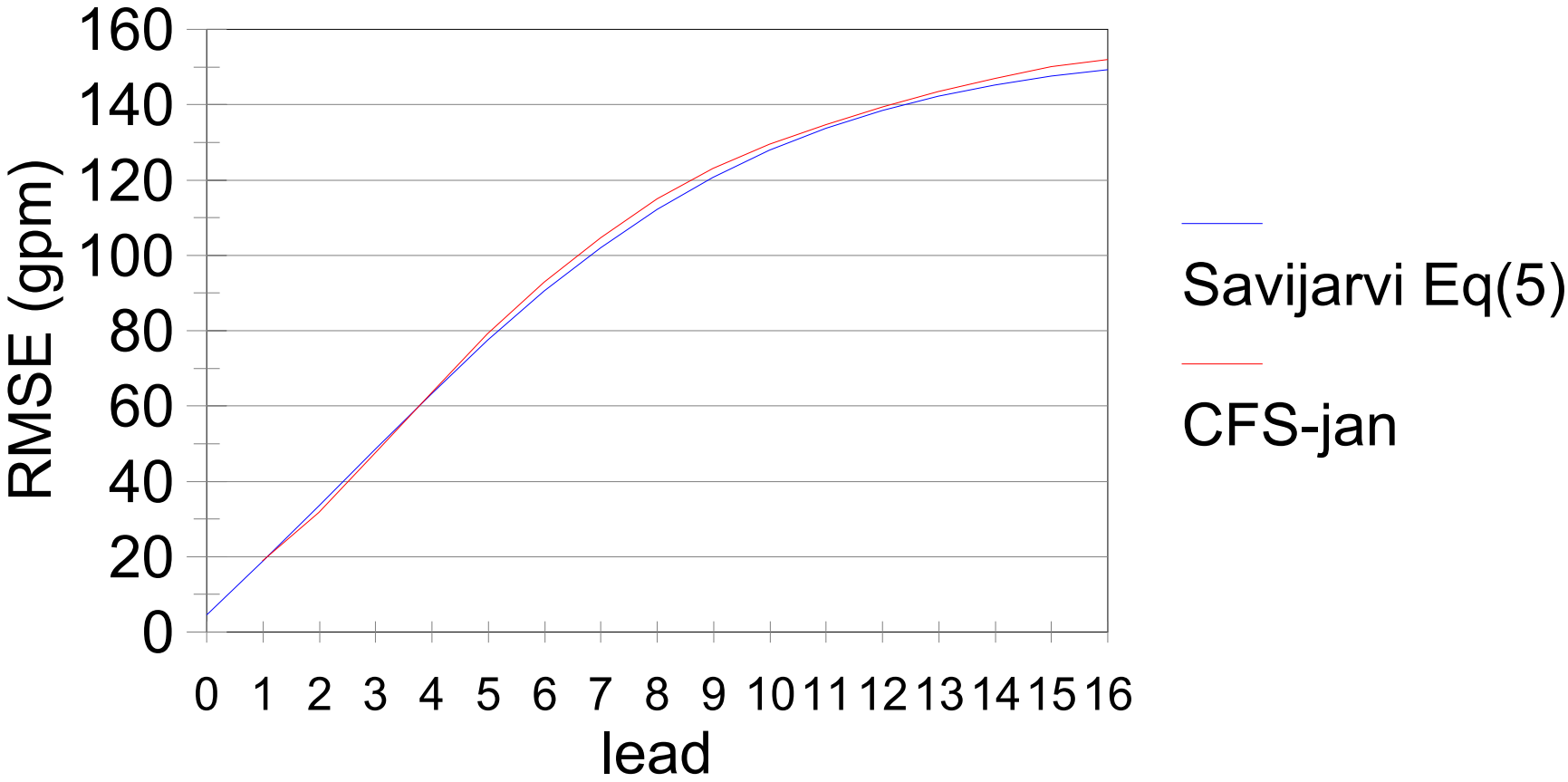
$$\frac{de}{dt} = (ae + s)\left(1 - \frac{e}{e_i}\right), \qquad (5)$$

Lorenz Tellus 1982



RMSE growth CFS NH

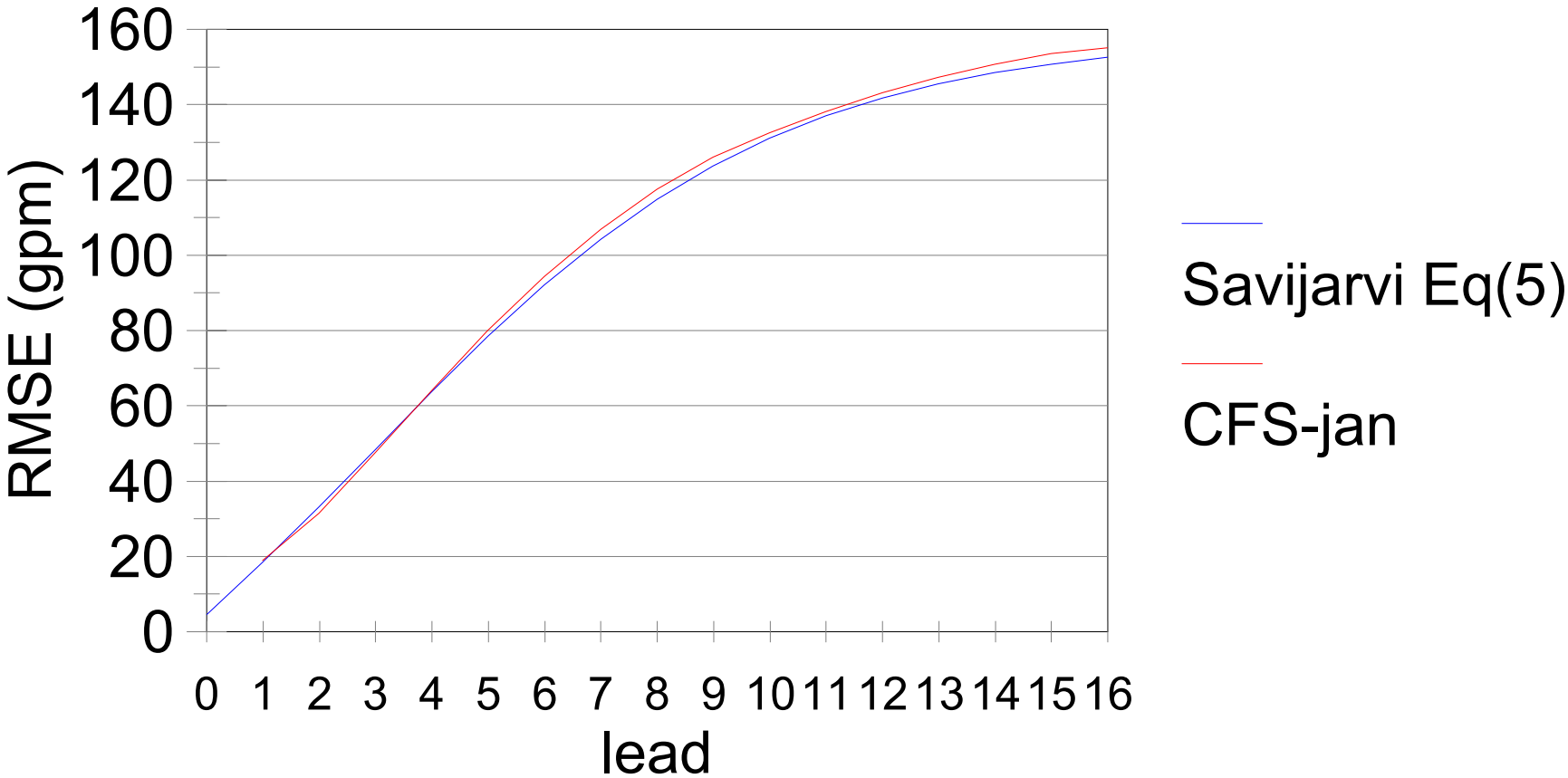
SEC



Initial error=4.7 gpm ; Systematic error growth s=13.1 gpm/day ; Small error amplification a=0.181/day ; e-infinity=155.7gpm ; SEC

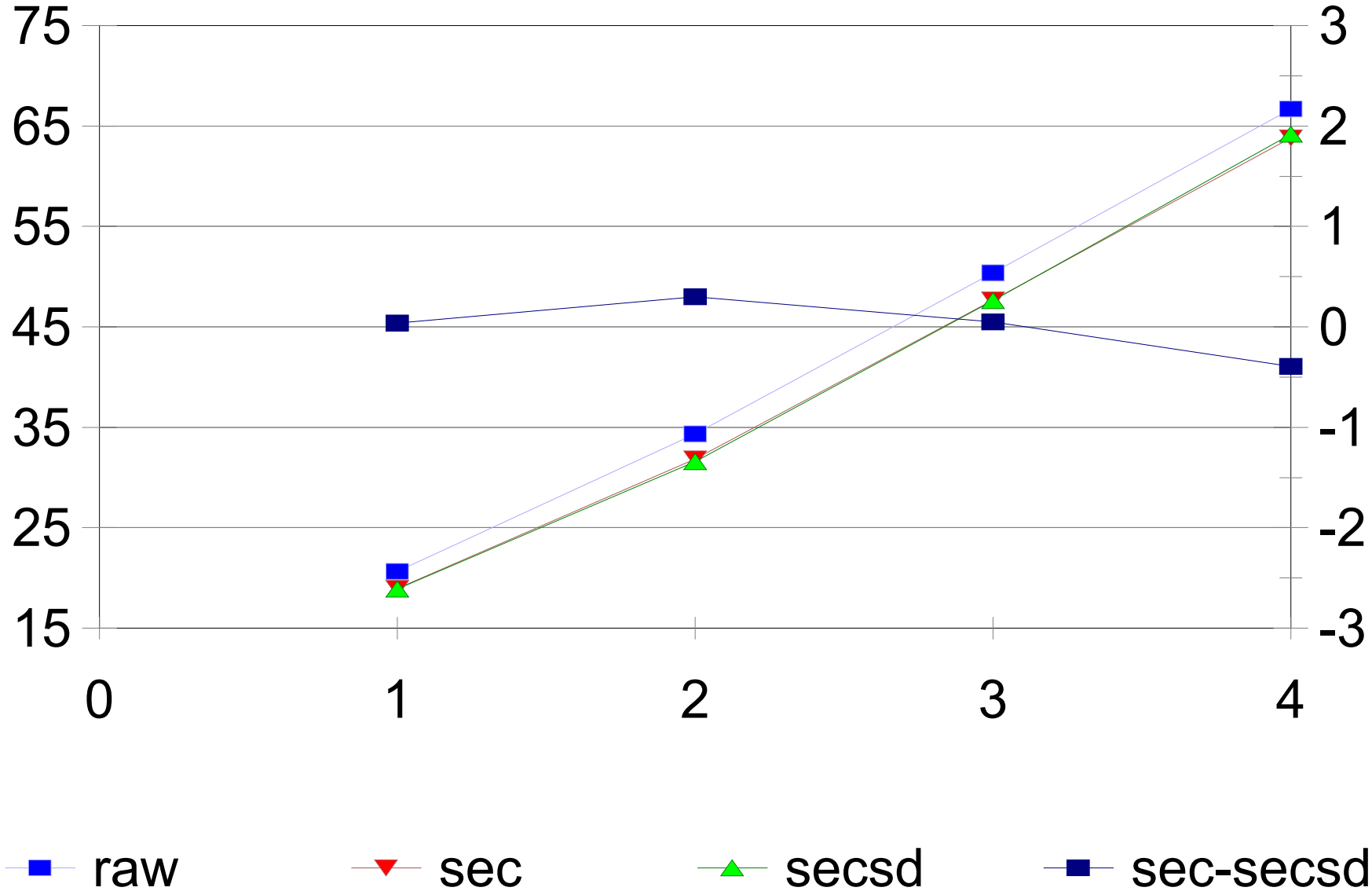
RMSE growth CFS NH

SECSD

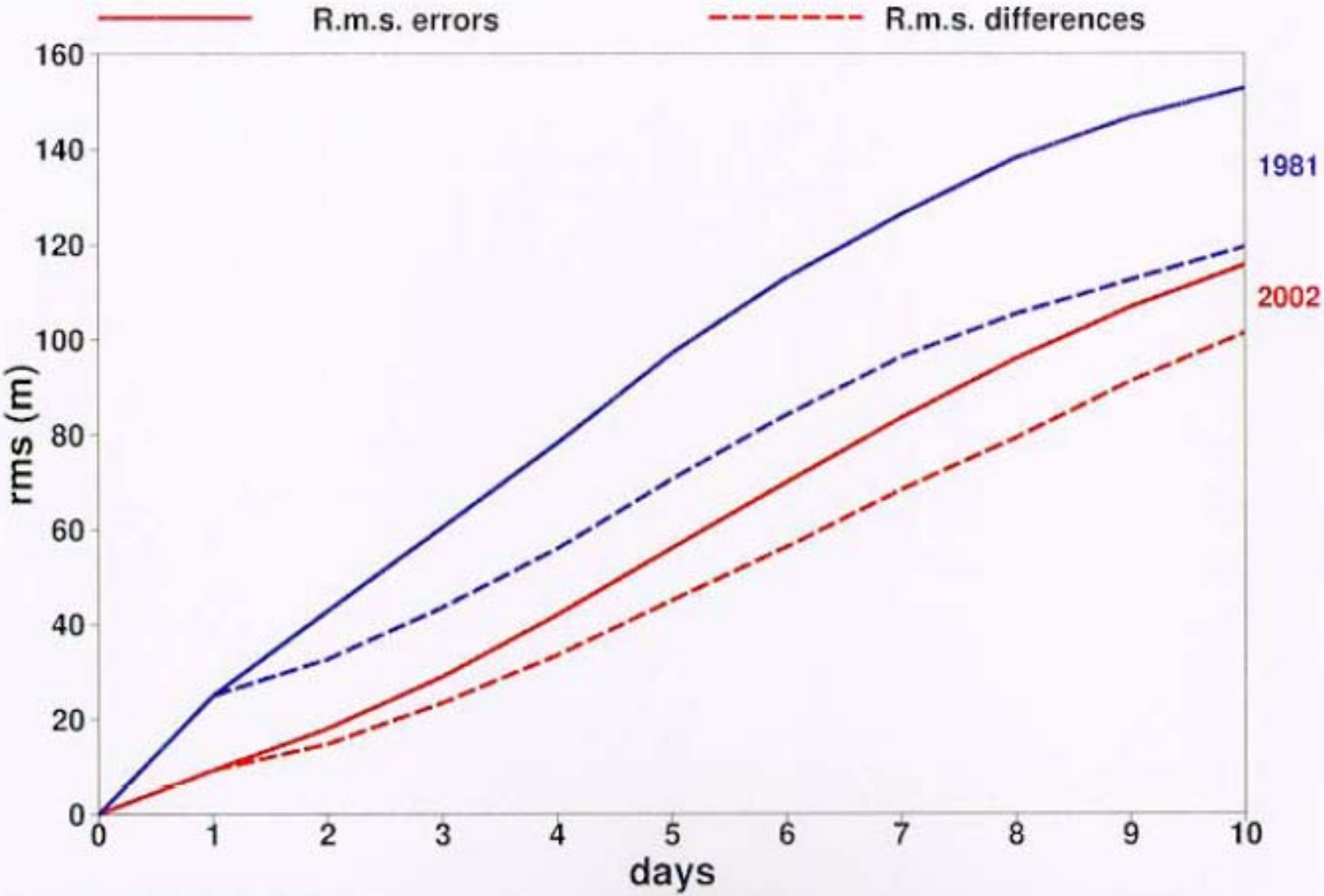


Initial error=4.7 gpm ; Systematic error growth s=12.6 gpm/day ; Small error amplification a=0.200/day ; e-infinity=158.3gpm ; SECSD

RMSE growth CFS NH 'jan' (Z500)

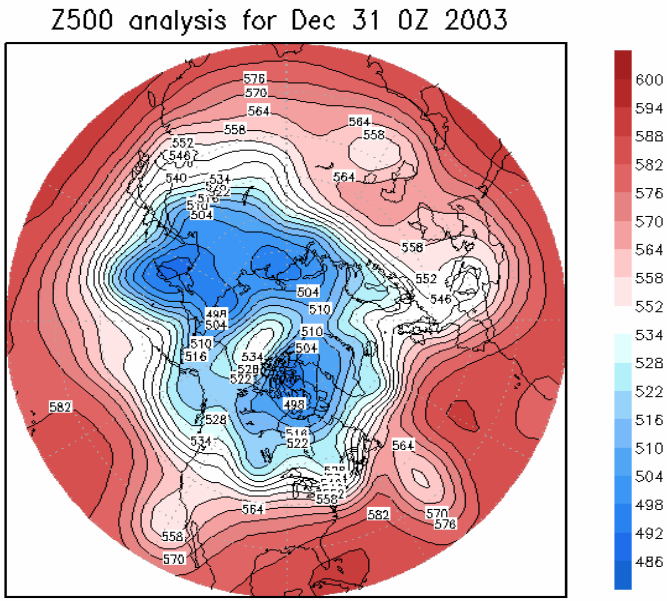


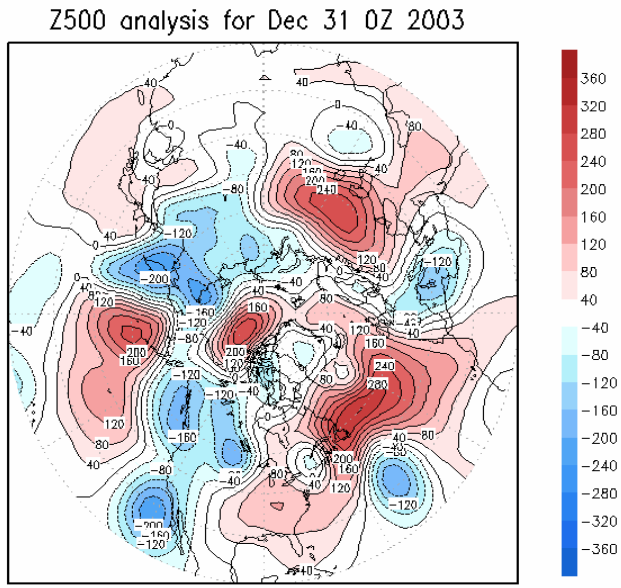
Northern hemisphere 500hPa height Winter



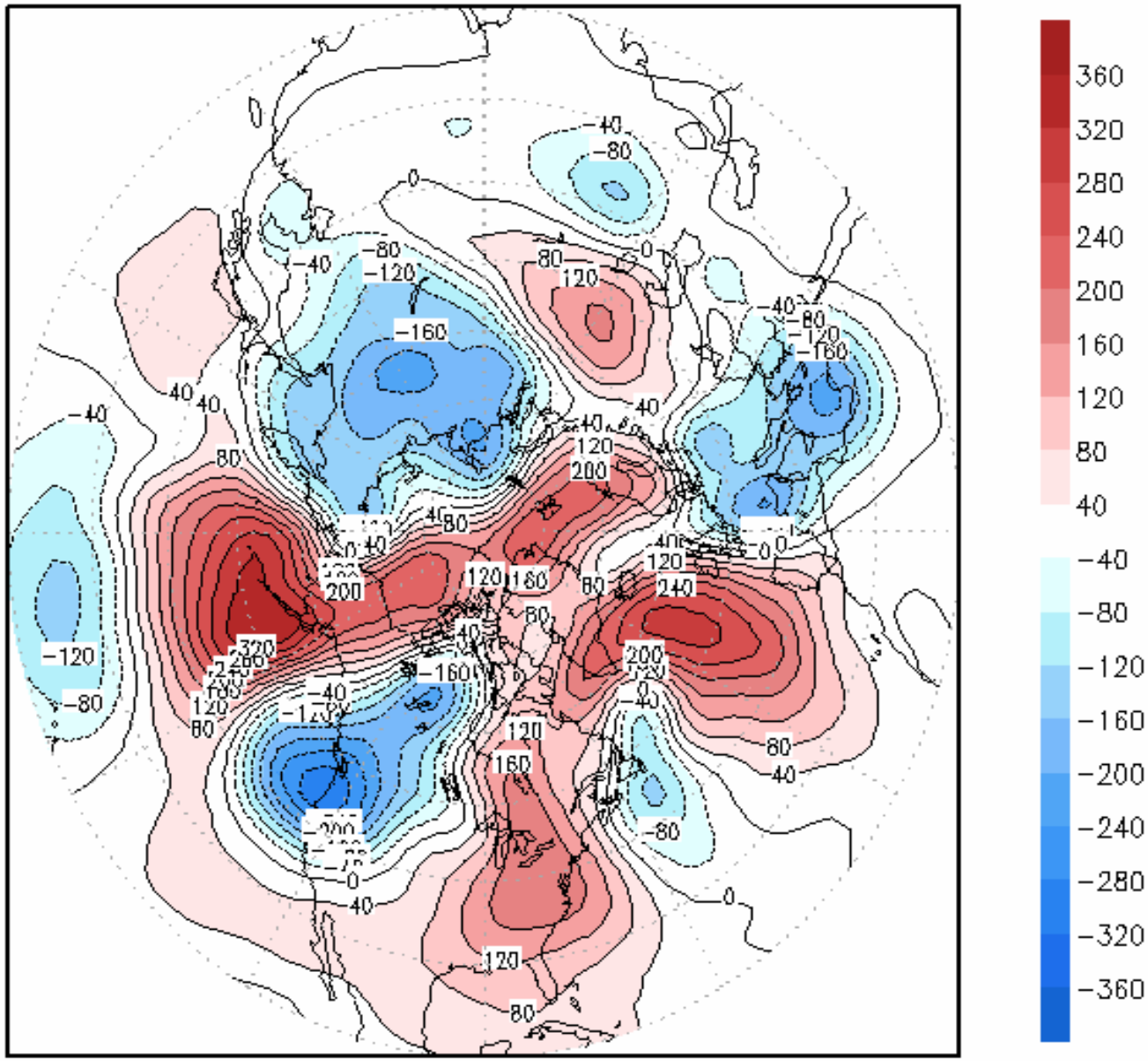
Among 4500 CFS daily forecasts:

- Was the RMSE at day 4 ever smaller than at day 3????? (Model vs Analysis)
- Yes, in one single case: the forecast from Dec, 31, 2003. Raise the champagne???

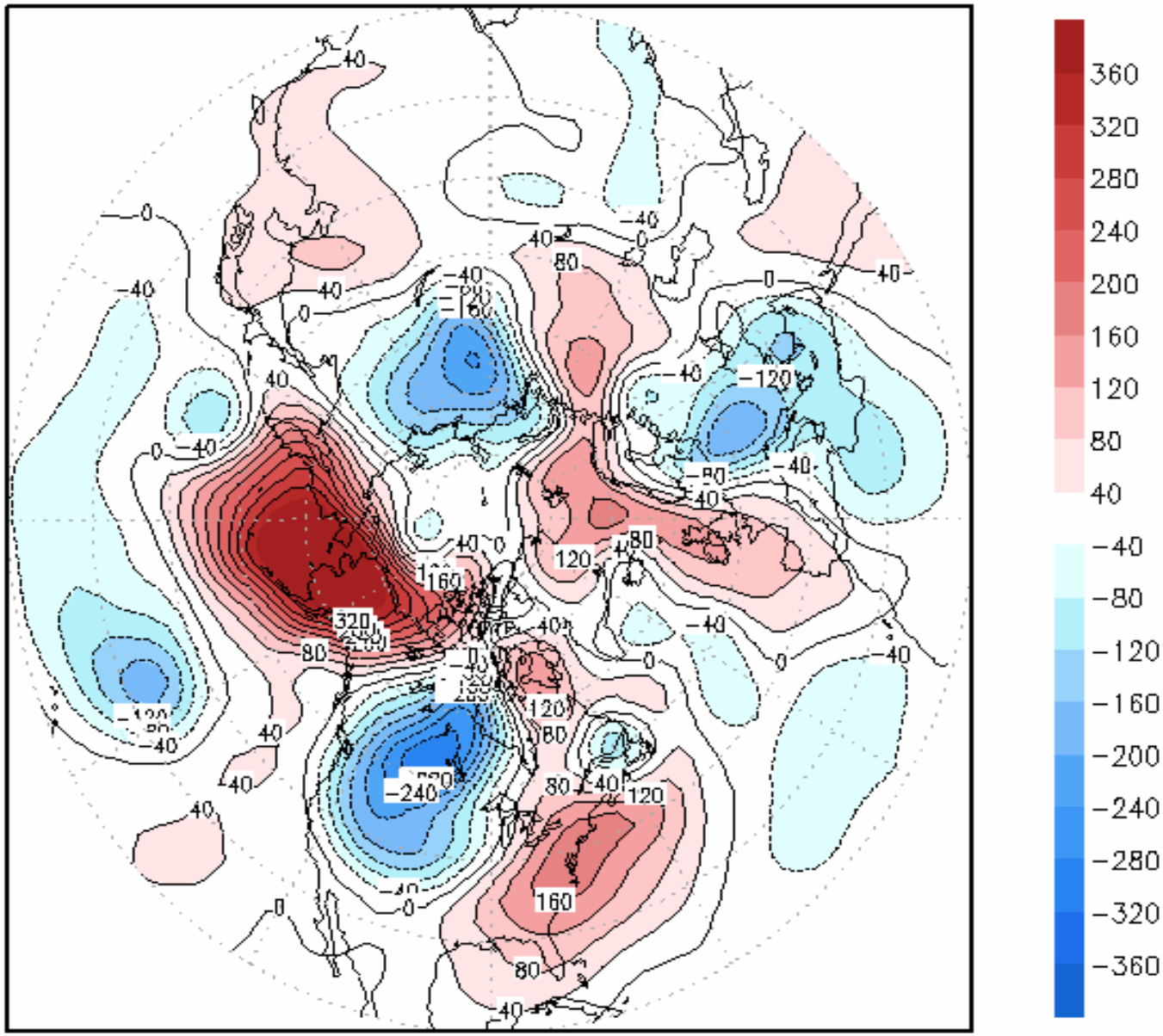




Z500 analysis for Jan 2 0Z 2004



Z500 analysis for Jan 5 0Z 2004



	1	2	3	4	5	6		9
AC	98.8	98.6	97.9	98.5	97.0	91.6		
RMSE	17.1	20.2	24.6	20.5	27.4	45.2		87.9
	Jan 1				Jan 5			

An exceptionally “good” forecast, especially for T62 (CDAS2).

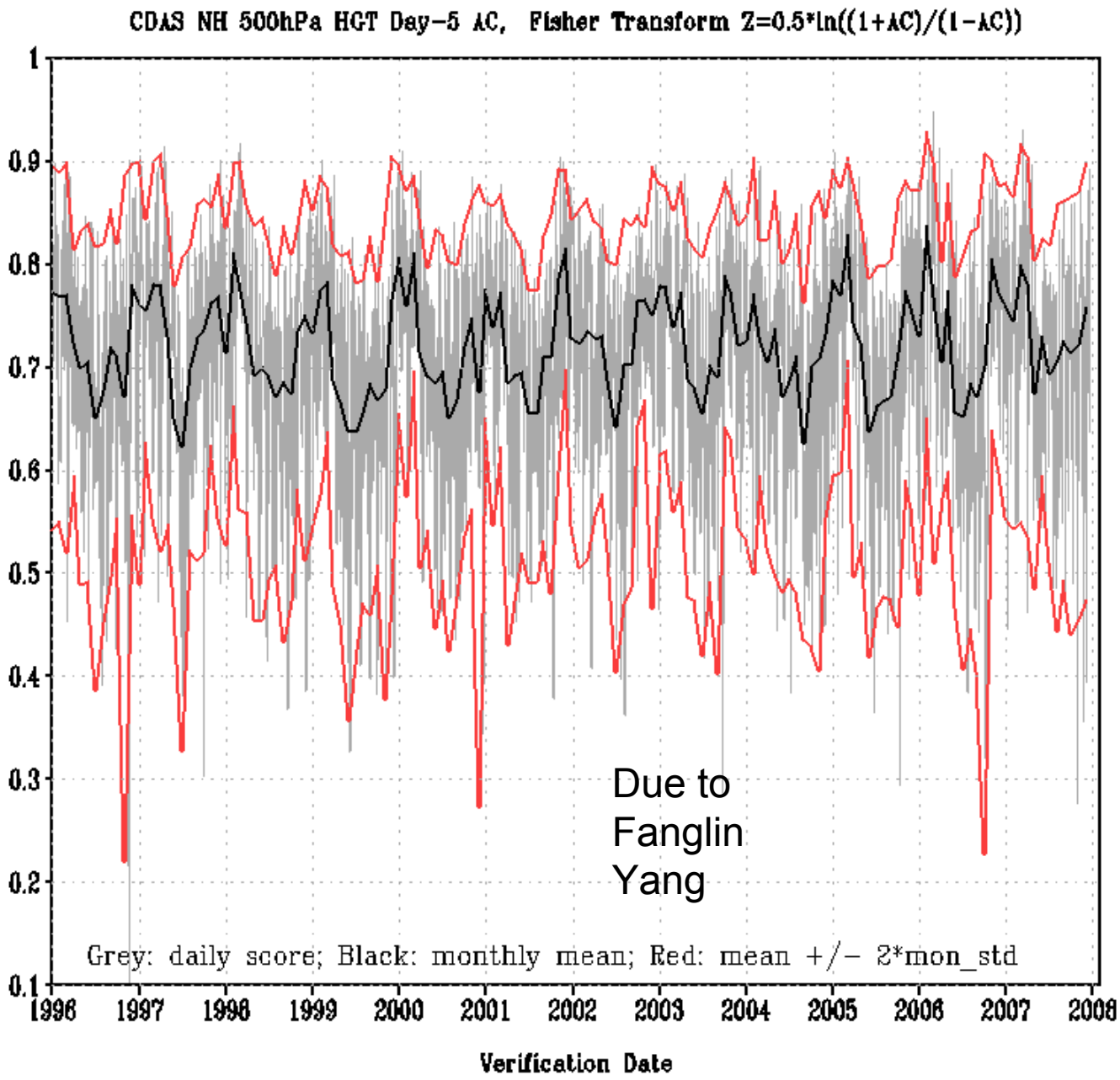
The best forecast (for day 5) ever seen in this building, by any model

Flow Situation is NOT particularly persistent

Forecasts from Dec 30 or Jan 1, 2 are not that good. Only Dec 31, 0Z.

Other models not that good for this case

Too good to be true? The date raises some suspicion


CDAS NH 500hPa HGT Day-5 AC, Fisher Transform Z=0.5*ln((1+AC)/(1-AC))
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
Verification Date
Grey: daily score; Black: monthly mean; Red: mean +/- 2*mon_std
Due to
Fanglin
Yang

Another piece of evidence:

- Scores in the SH for 2003123100 were also exceptionally good (0.92, 2nd highest is 0.77).
- It is unlikely that both hemispheres produce ‘hit-the-jackpot’ type forecasts.
- Could something be the matter with the R2 analyses.

What could be wrong in the (R2 or any) analyses verifying the CFS forecast from 31 Dec, 0Z??

- Drunken observers around the planet in a wave-like pattern (New Years Eve distracts!
- Satellites flying in and out of the year (in local time), i.e. time problems. Indeed much satellite data is rejected (for days in early January), sometimes ‘wholesale’.
- But can this derail the analysis enough to explain a ‘very good’ forecast?

ABOUT THE FUTURE

5 day forecast anomaly correlation

Annual Mean

Reanalysis	NH	SH
OPR	85	80
CFSRR-lite	74	72
CFS (R2)	72	63
R1	71	62

Conclusions

- CFS is (slightly) better than CDAS (Z500).
- Over time (1981-2005), the CFS “system” appears quite constant with various qualifiers
- If we need an as-constant-as possible in-house system, look no further than CFS (better than CDAS)
- Bias correction: small +ve impact on Z500 and Ψ200 (1-2pnts)
- Bias correction: huge +ve impact on CHI-200 (15 pnts), but this is worrisome
- CFS loses 5-20% in terms of SD and eDOF in the first few days, then, admirably, stays nearly constant out to 270 days

More conclusions

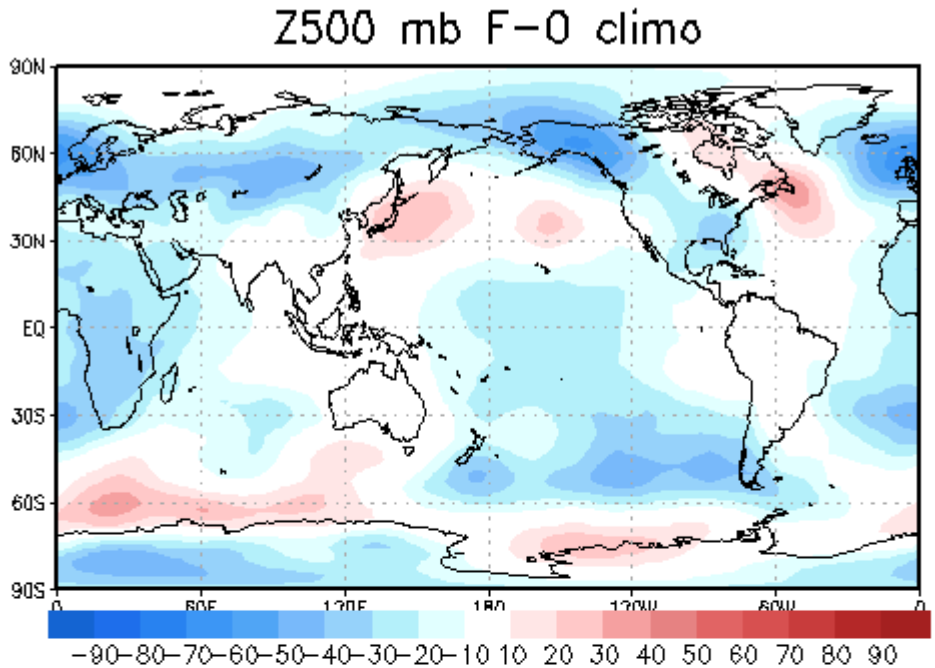
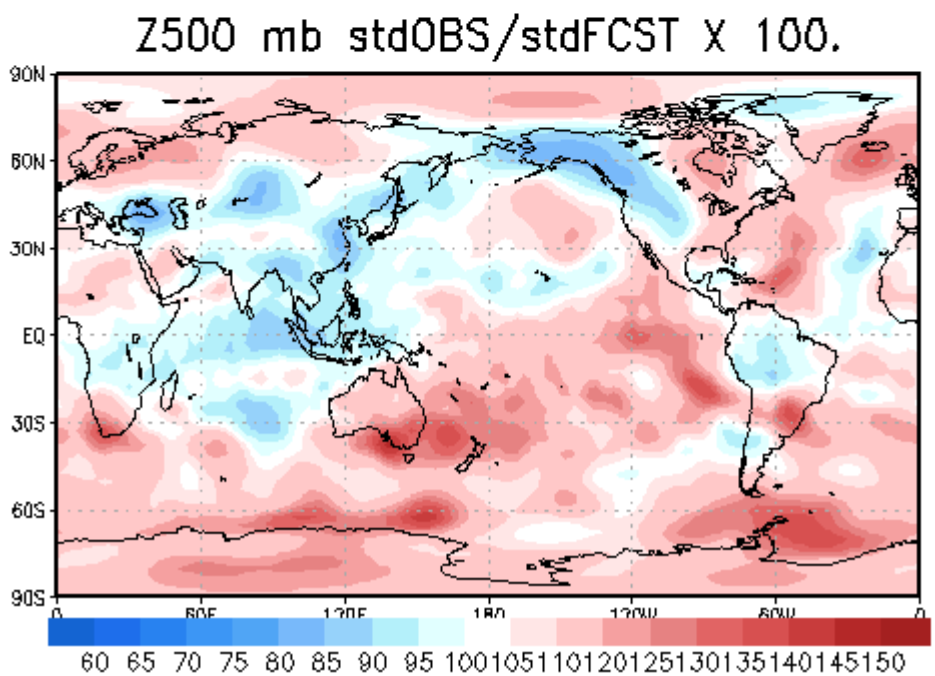
- Scores in SH/TR more volatile than in NH.
- SH lacks a proven annual cycle in scores
- Spatial distribution of scores not easy to understand
- Chi200 scores at day 1 point to serious problems in R2 and ‘consistency’
- Scores as a function of EOF non-surprising
- Error growth equation fits (Lorenz/Savijarvi) indicates large (mainly random) error growth due to systematic model error. Surprising. Needs more study. The good news about it is …..
- We have one perfect forecast.

More Conclusions

- (Very) modest skill in wk3 and wk4
- Even with 2 of 3 steps for signal to noise improvements in place, the AC is only 0.20-0.25. (No ensemble average here)
- Waiting for the next CFS and CFSRR (higher Res, consistent IC)

Conclusions:

- The day 1 – 3 forecasts appear to be too damped, and damp faster than a regression would. Increasing anomaly amplitude as a postprocessor (undoing the sd decrease) actually improves the rms error early on. A curiosity?
- Probably: initial conditions are damped as well.

Example:

Systematic errors for mid-January at day 5.