The impact of model error on ENSO forecast performance

Tim Stockdale

With input from Magdalena Balmaseda, Paco Doblas-Reyes

Seasonal Forecast Group

European Centre for Medium-Range Weather Forecasts

T.Stockdale@ecmwf.int

Overview

- **Introduction to model error: a new operational seasonal forecast system**
  - How good are the forecasts, and is model error important?
- **Can we side-step model error?**
  - Bias correction terms and post-processing
  - Multi-model techniques
- **Some specific comments on model error**
  - Resolution / timestep
  - Coupling of surface currents

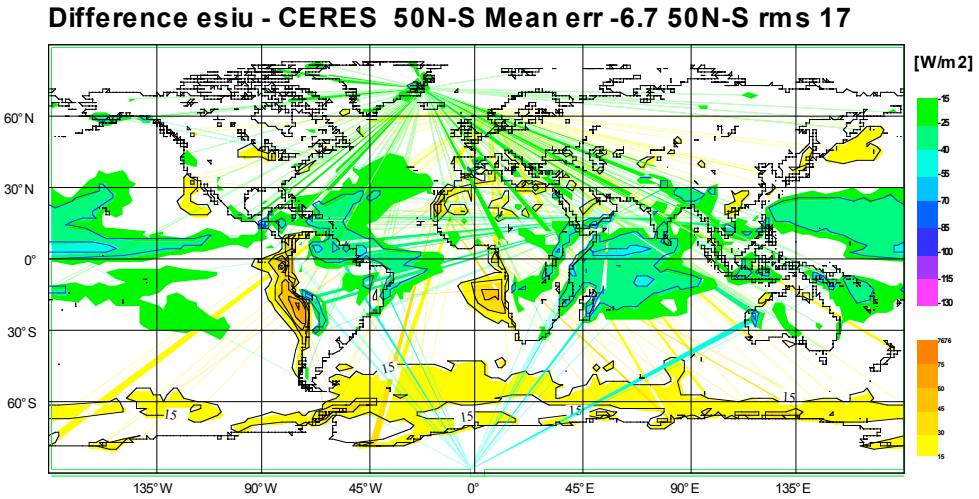
ECMWF seasonal forecast model (System 3)

- **IFS (atmosphere)**
  - T_L159L62 Cy31r1, 1.125 deg grid for physics (operational in Sep 2006)
  - Full set of singular vectors from EPS system to perturb atmosphere initial conditions (more sophisticated than needed …)
  - Ocean currents coupled to atmosphere boundary layer calculations
- **HOPE (ocean)**
  - Global ocean model, 1x1 mid-latitude resolution, 0.3 near equator
  - A lot of work in developing the OI ocean analyses, including analysis of salinity, multivariate bias corrections and use of altimetry.
- **Coupling**
  - Fully coupled, no flux adjustments, except no physical model of sea-ice

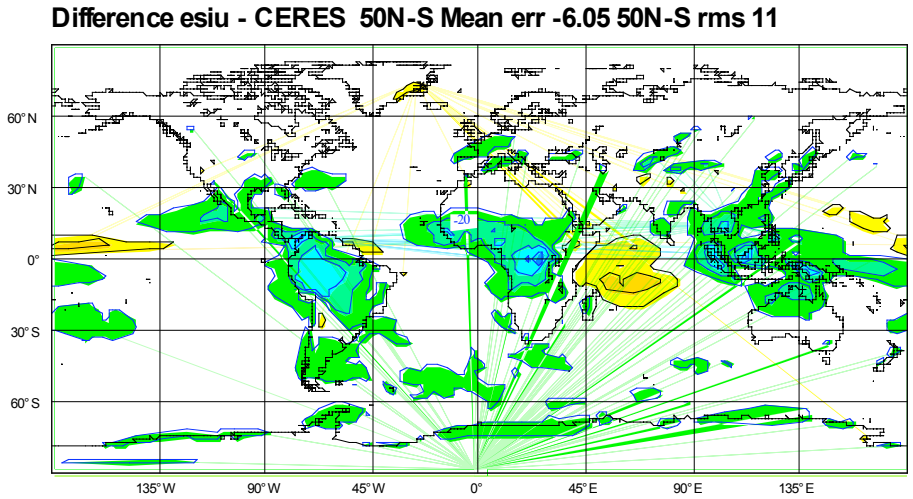
Cy31r1 uncoupled runs

Global mean net TOA flux balance = -4 Wm/2

CERES balance = +5 Wm/2 (for same 12 month period)



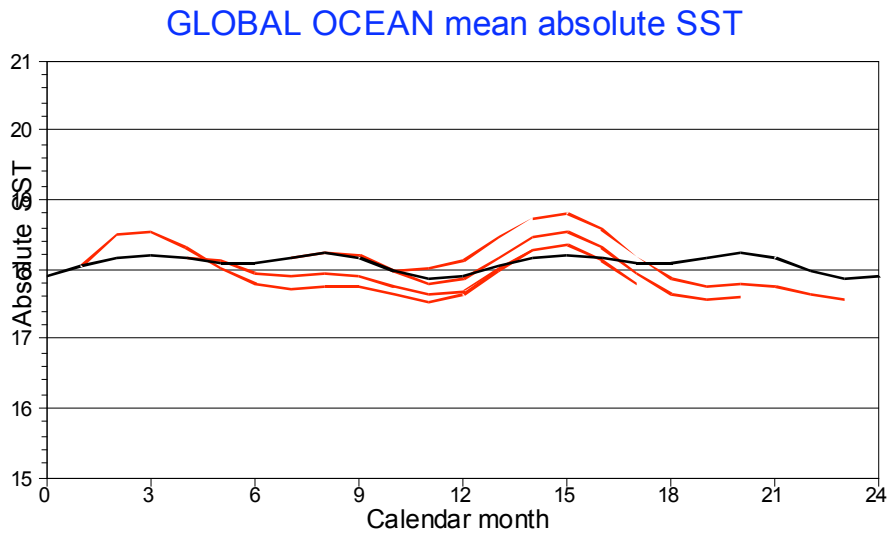
TOA SW error

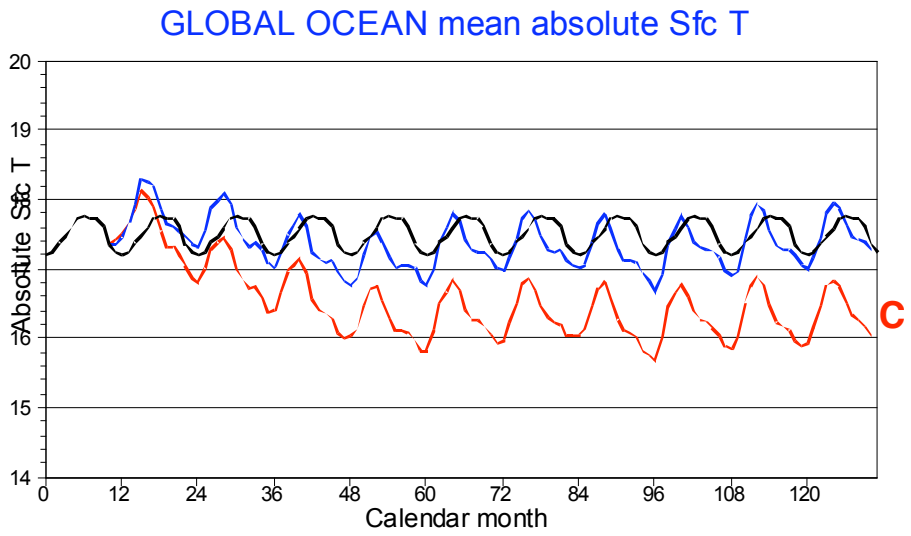


TOA LW error

ECMWF

Cy31r1 coupled runs





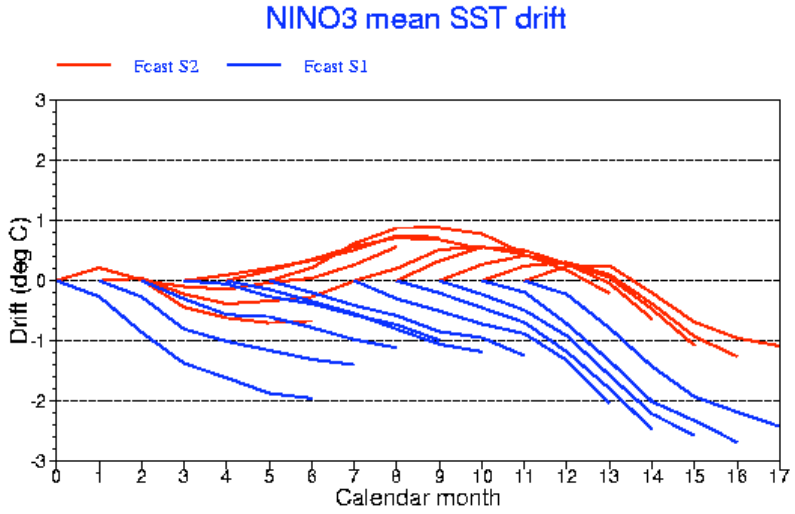
13 month integrations

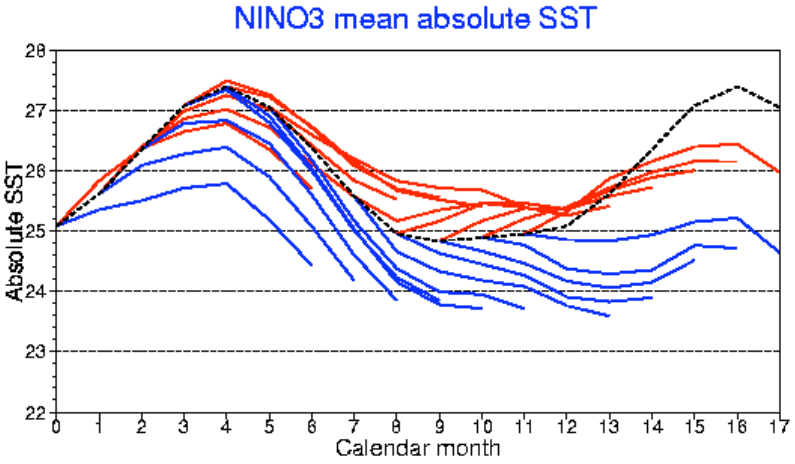
10 year integrations

ECMWF

System 3 configuration

- **Real time forecasts:**
  - **41 member ensemble forecast to 7 months**
  - SST and atmos. perturbations added to each member
  - **11 member ensemble forecast to 13 months**
  - Designed to give an 'outlook' for ENSO
  - Only once per quarter (Feb, May, Aug and Nov starts)
- **Back integrations from 1981-2005 (25 years)**
  - 11 member ensemble every month
  - 5 members to 13 months once per quarter
- **5 member ensemble ocean analysis/reanalysis**
  - Includes representation of ocean initial condition uncertainty

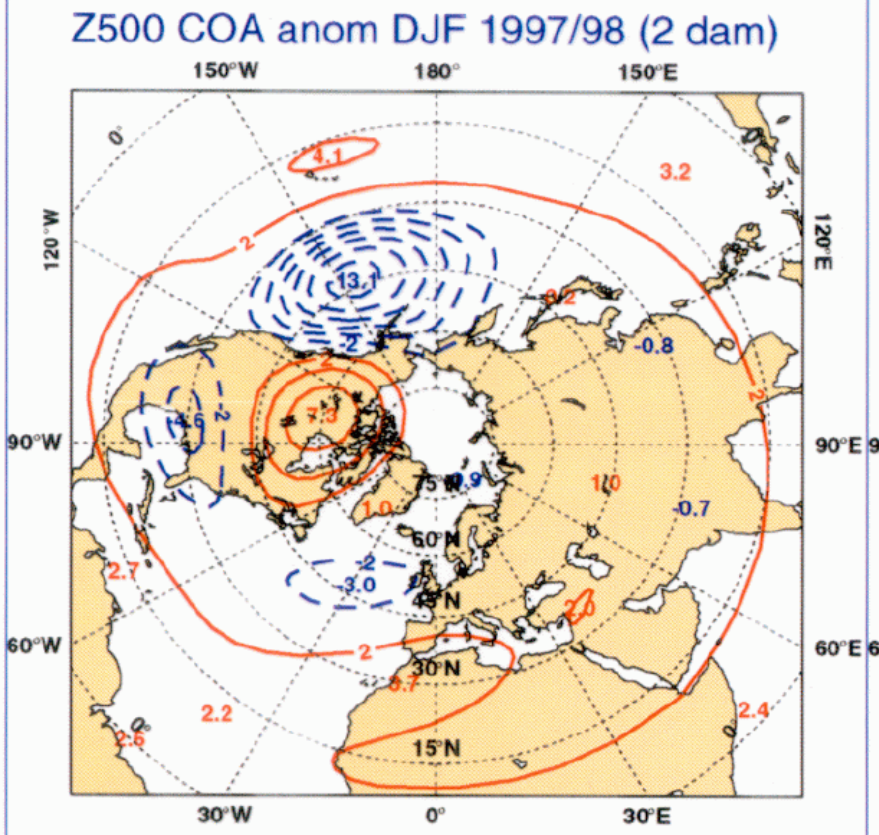


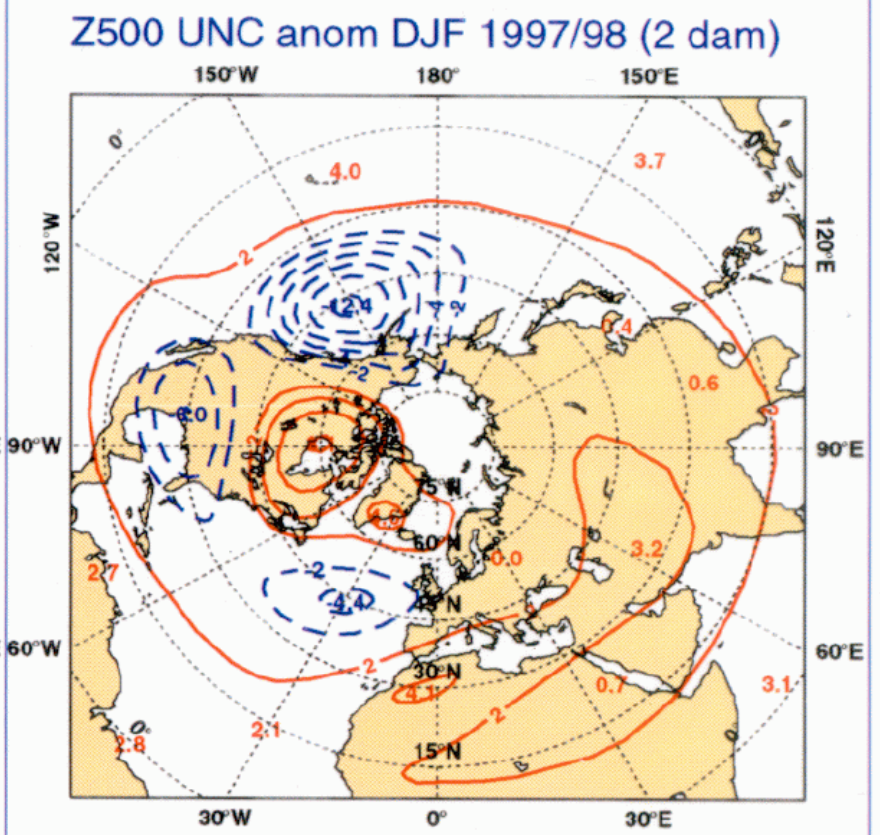


Models have significant BIAS, which is estimated from previous integrations and removed.

Earlier model versions had quite large bias, but this did not destroy their forecast skill

Despite the large SST bias in S1, anomalies in the coupled forecasts could be remarkably similar to those obtained using observed (unbiased) SSTs



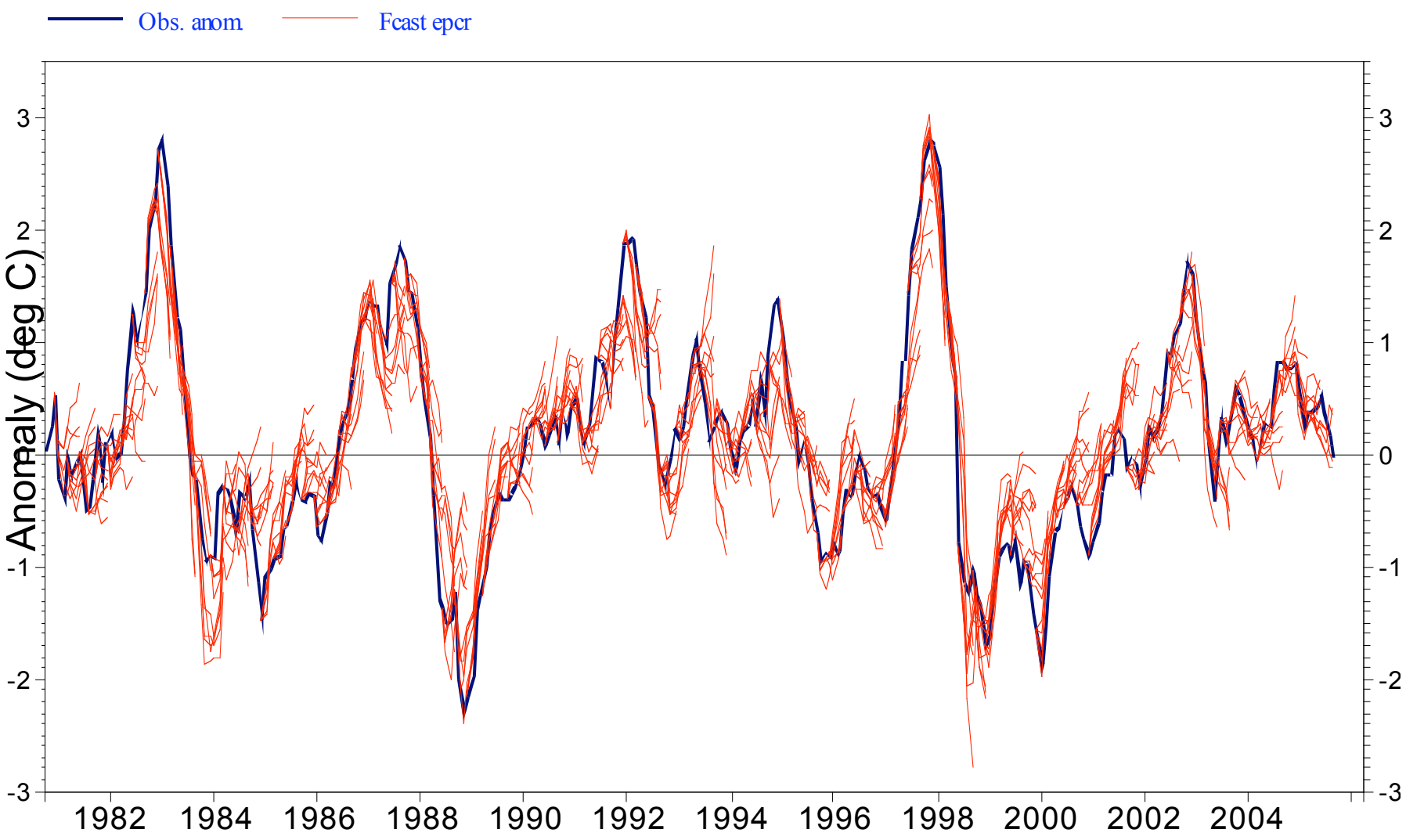


ECMWF

System 3

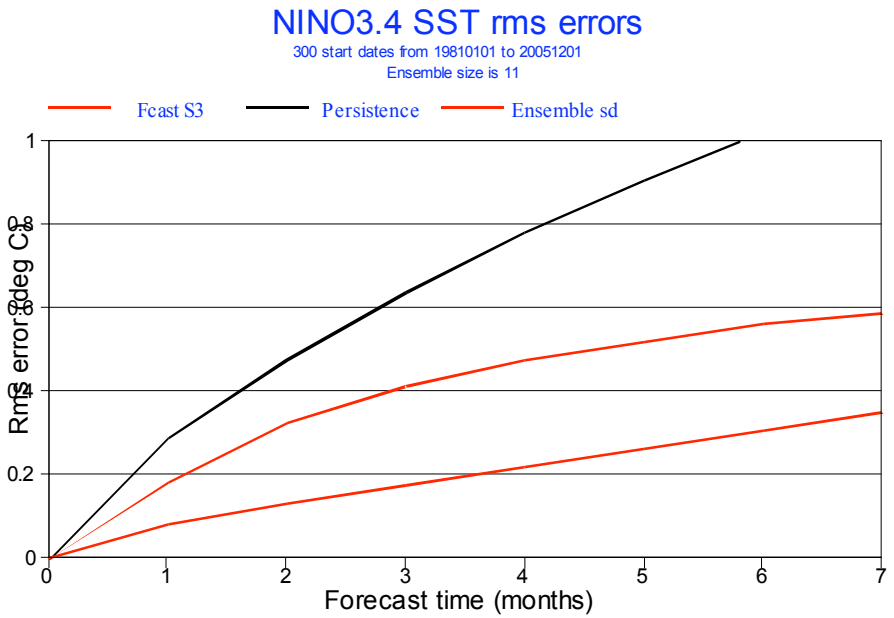
NINO3.4 SST forecast anomalies

ECMWF forecasts at month 6
Ensemble size is 5 SST obs:

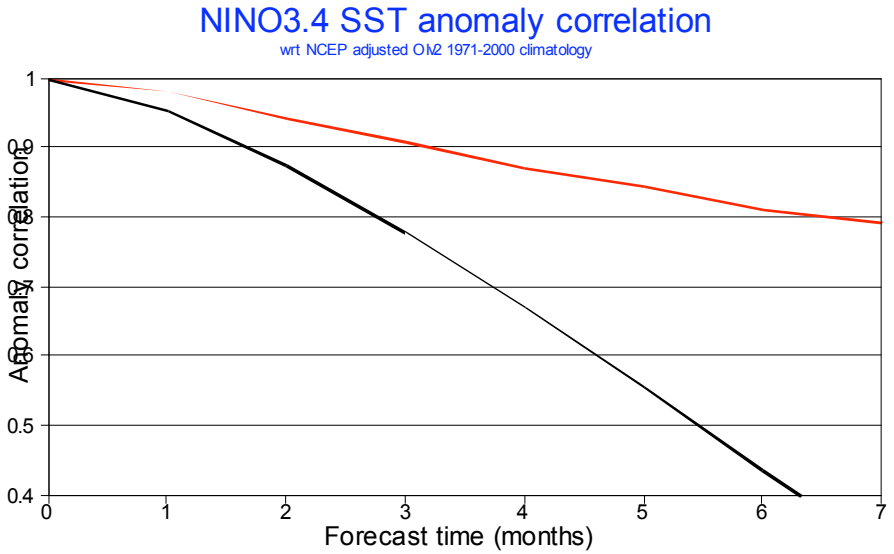


ECMWF

Forecast performance

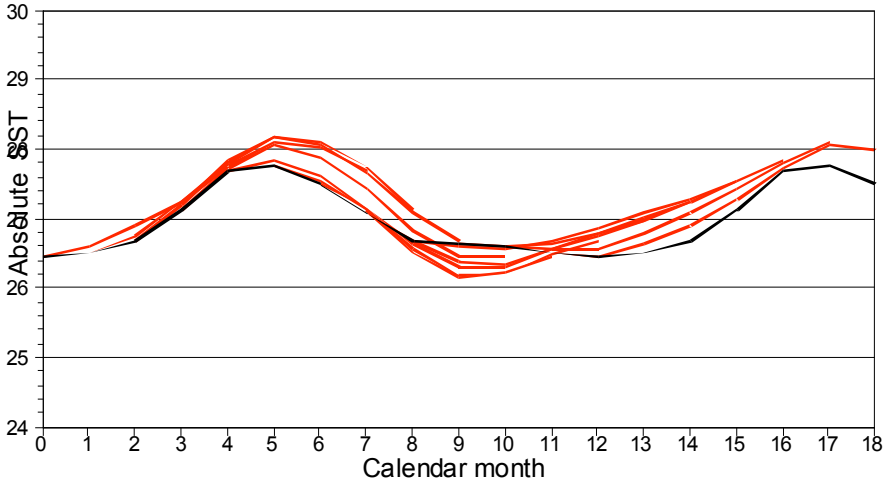


Actual rms errors > model estimate of "perfect model" errors



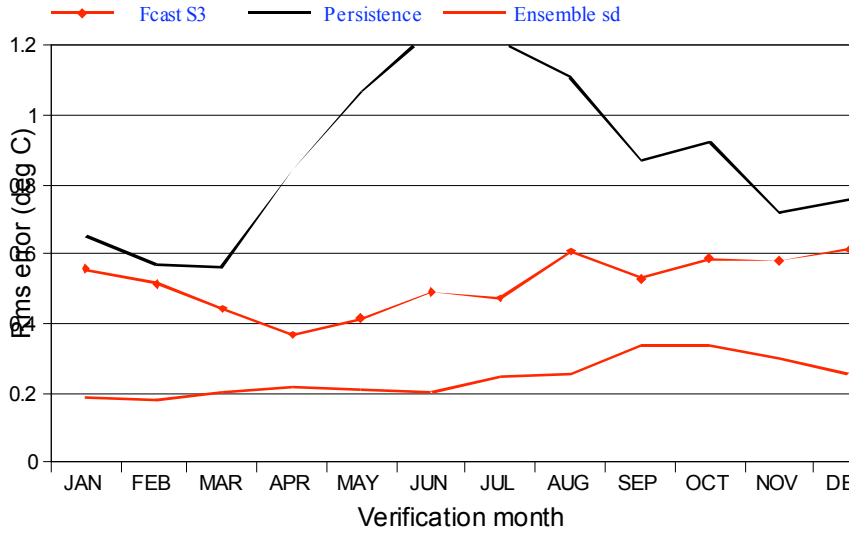
ECMWF

NINO3.4 mean absolute SST



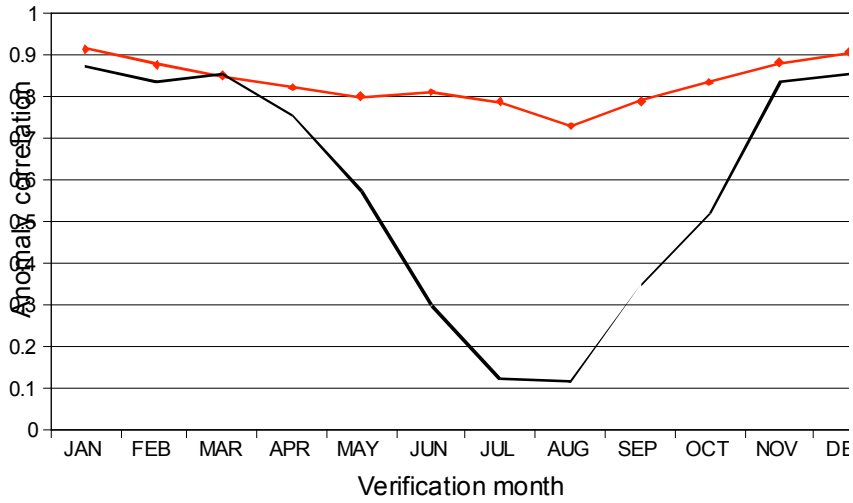
NINO3.4 SST rms errors at 5 months

300 start dates from 19810101 to 20051201
Ensemble size is 11

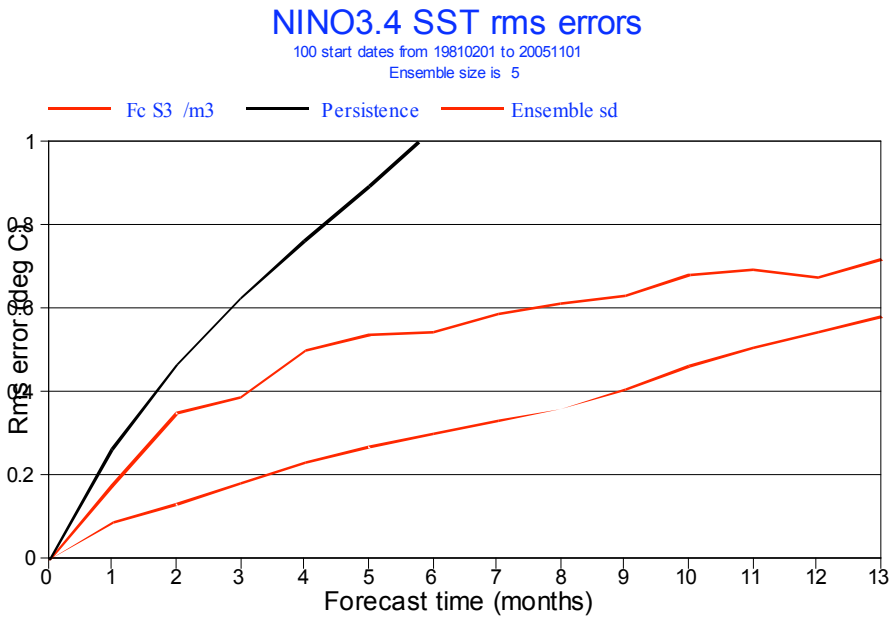


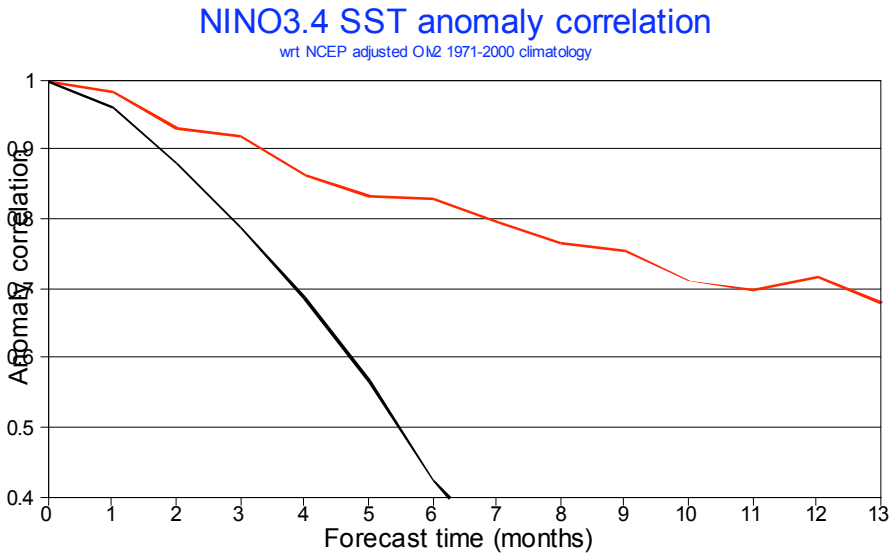
NINO3.4 SST anomaly correlation at 5 months

wrt NCEP adjusted OI2 1971-2000 climatology



ECMWF

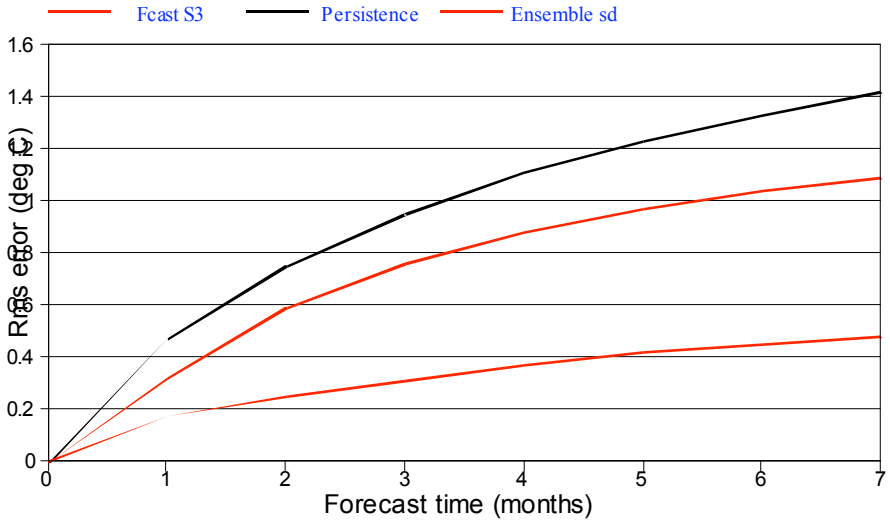




At longer leads, model spread starts to catch u

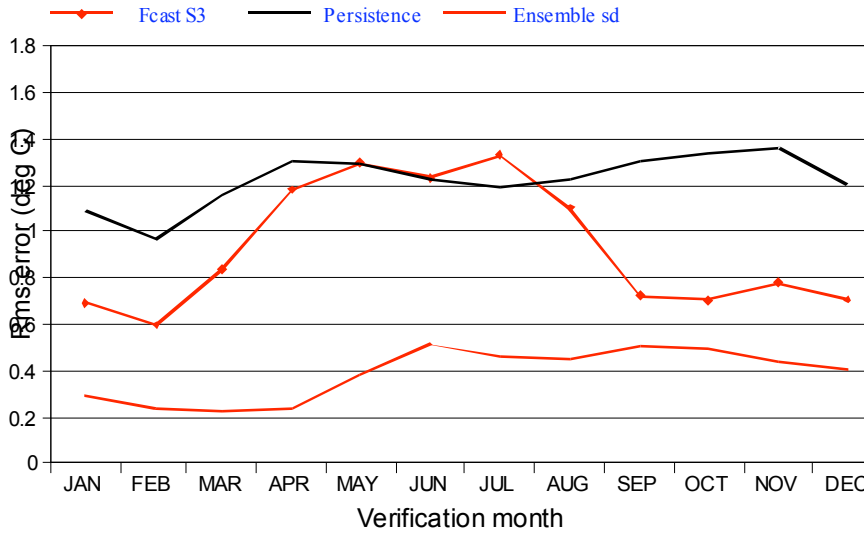
NINO1+2 SST rms errors

300 start dates from 19810101 to 20051201
Ensemble size is 11

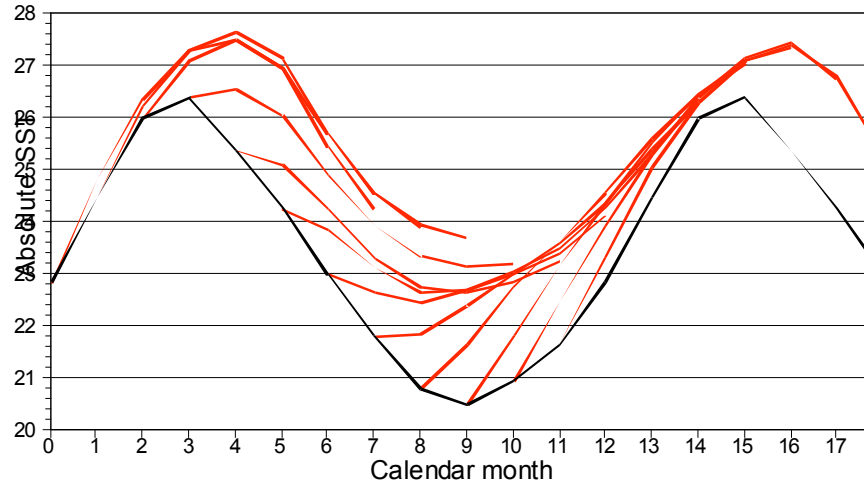


NINO1+2 SST rms errors at 5 months

300 start dates from 19810101 to 20051201
Ensemble size is 11

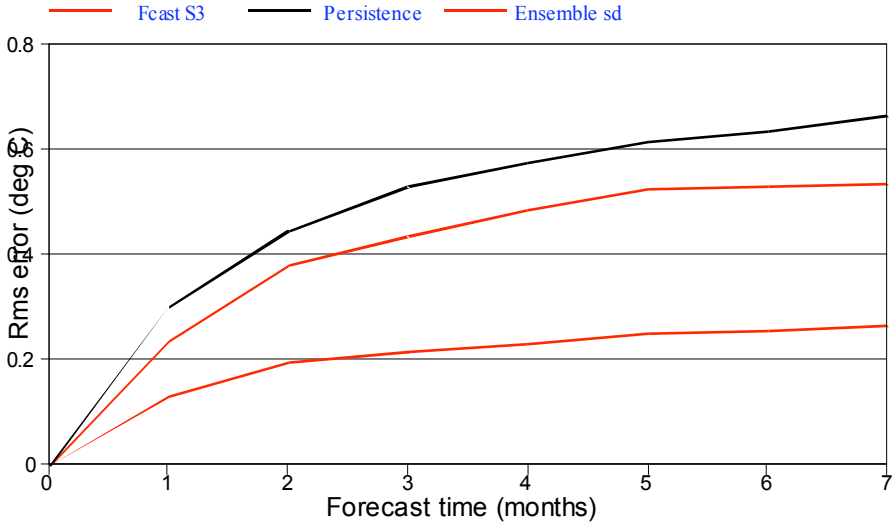


NINO1+2 mean absolute SST



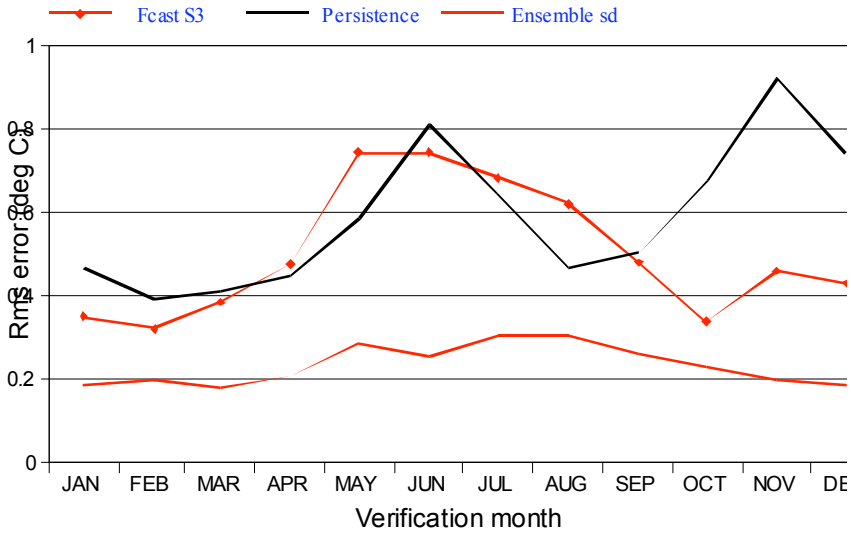
ATL3 SST rms errors

300 start dates from 19810101 to 20051201
Ensemble size is 11

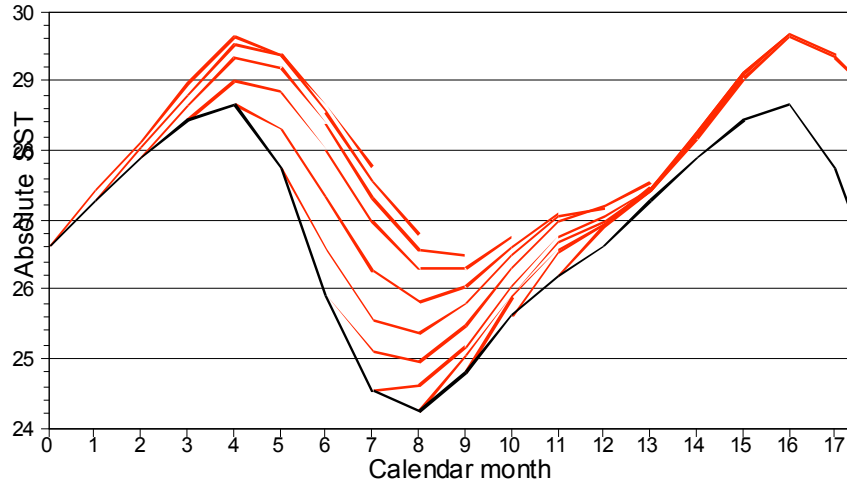


ATL3 SST rms errors at 5 months

300 start dates from 19810101 to 20051201
Ensemble size is 11



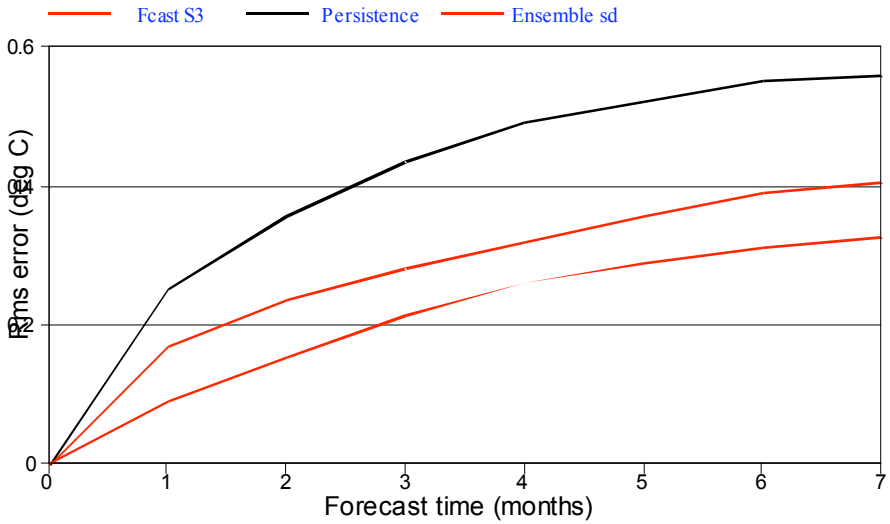
ATL3 mean absolute SST



ECMWF

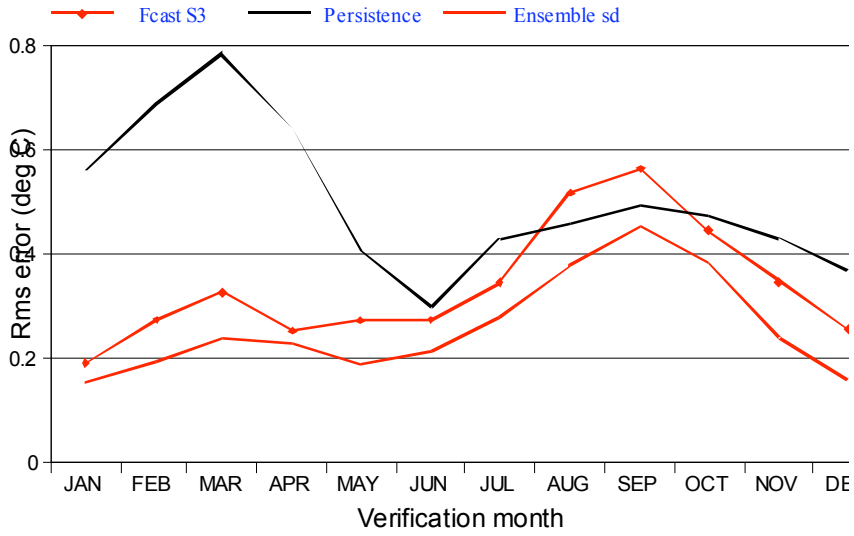
IND2 SST rms errors

300 start dates from 19810101 to 20051201
Ensemble size is 11

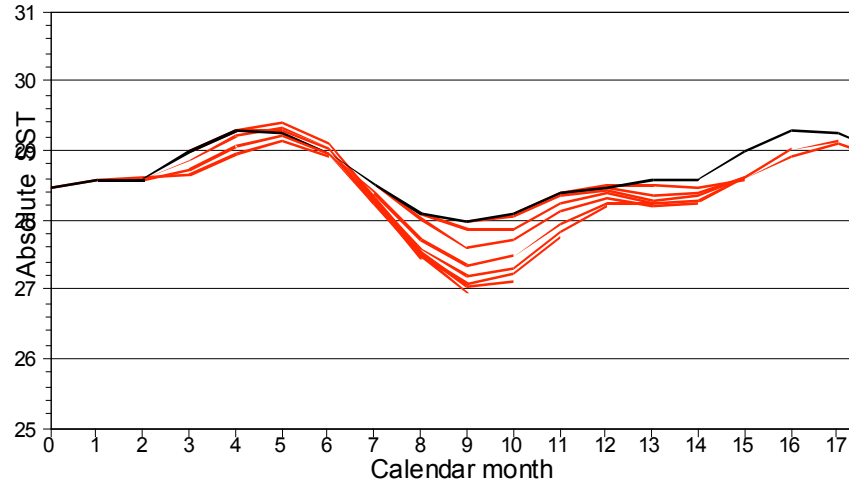


IND2 SST rms errors at 5 months

300 start dates from 19810101 to 20051201
Ensemble size is 11



IND2 mean absolute SST



ECMWF

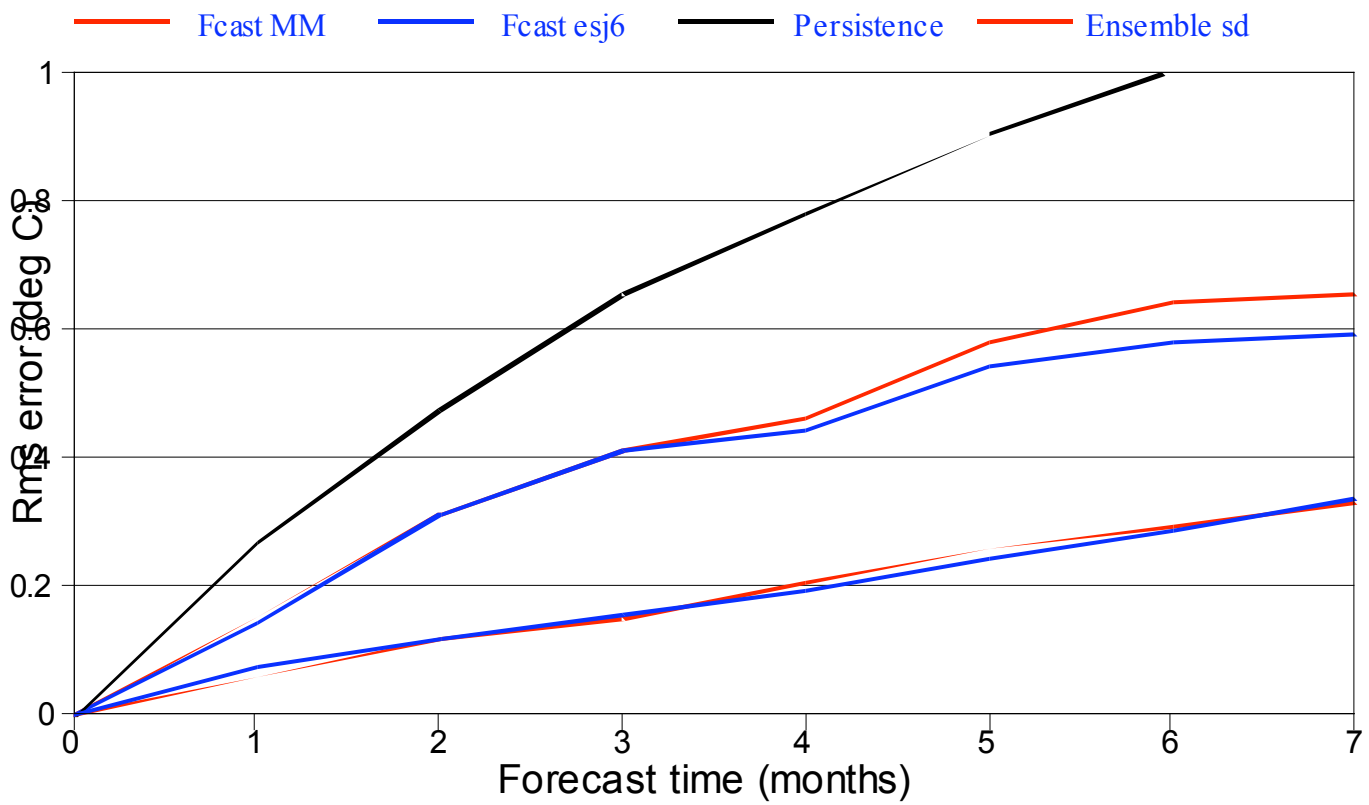
Model error in ENSO forecasts

- Model error appears to dominate ENSO forecasts
- But is initial condition error underestimated?

Red: **different analysis systems**

Blue: **ECMWF ensemble analysis**

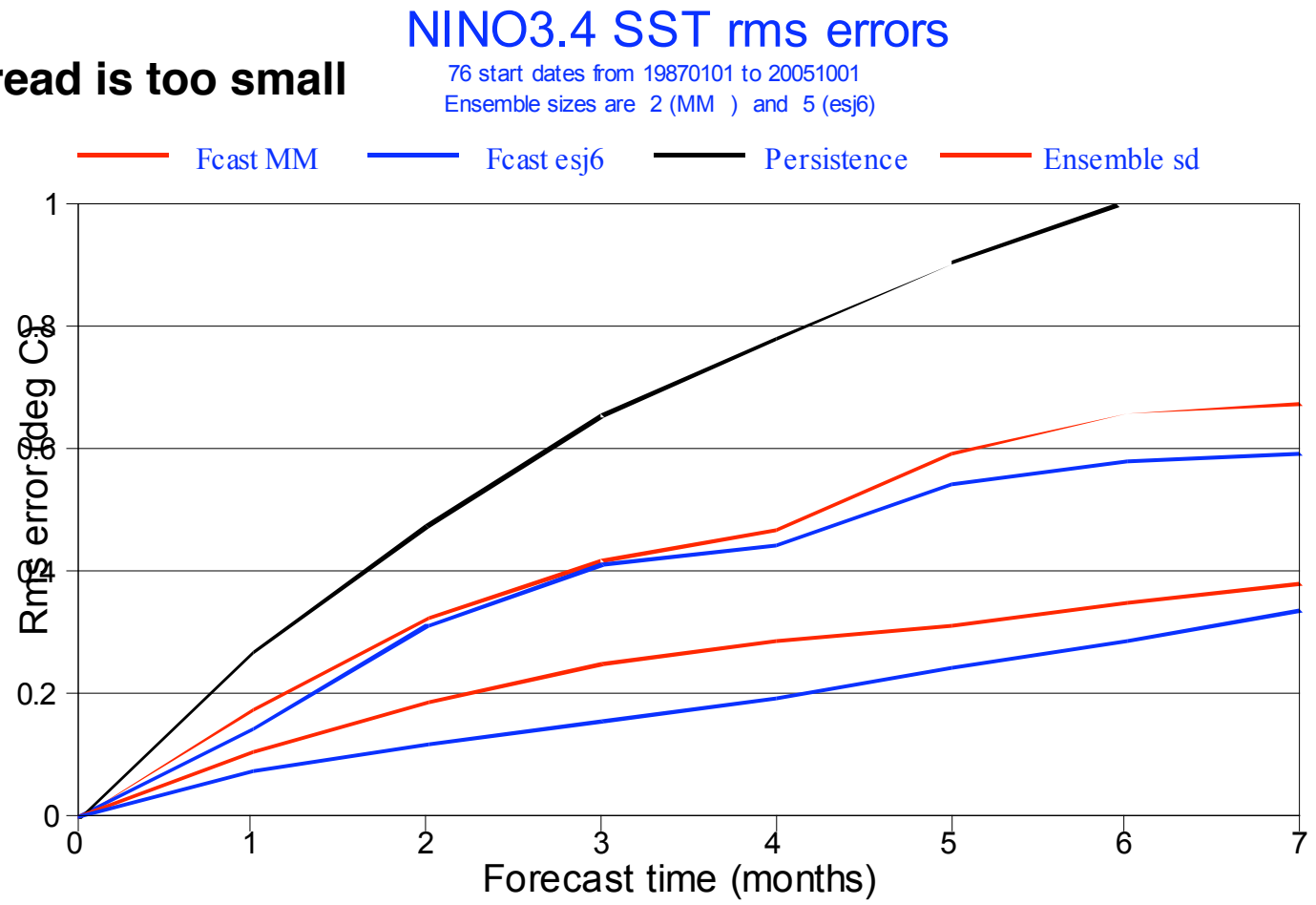
Suggests spread is about right



Red: **different analyses with very different mean states**

Blue: **ECMWF ensemble analysis**

Suggests spread is too small

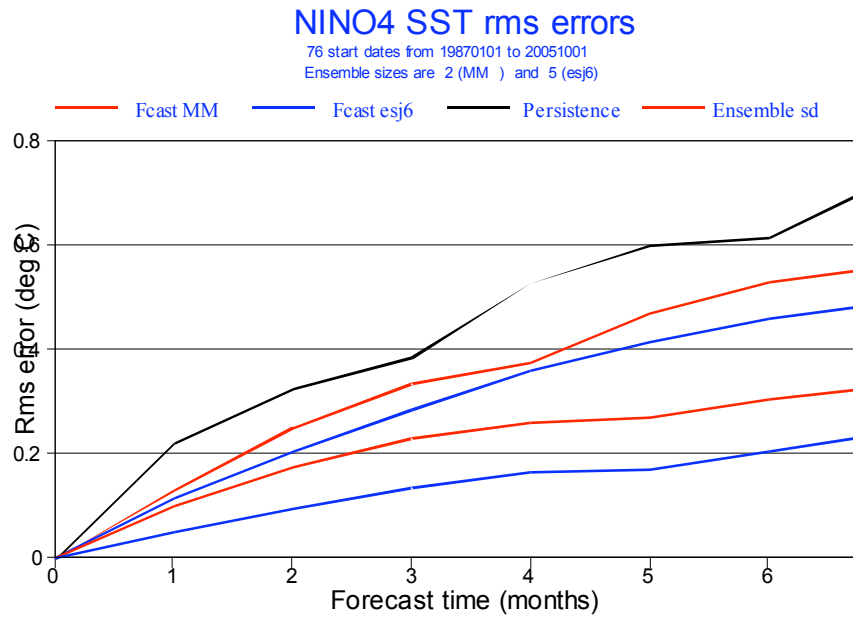


So what is going on?

Different analysis gives very different forecasts in W Pacific, but has little impact on forecast differences in E Pacific.

Suggests that impact of analysis in this case is dominated by mean state differences, not perturbations themselves

Difference in mean state in these experiments >> error in analysis mean state



Forecast sensitive to ic anomalies

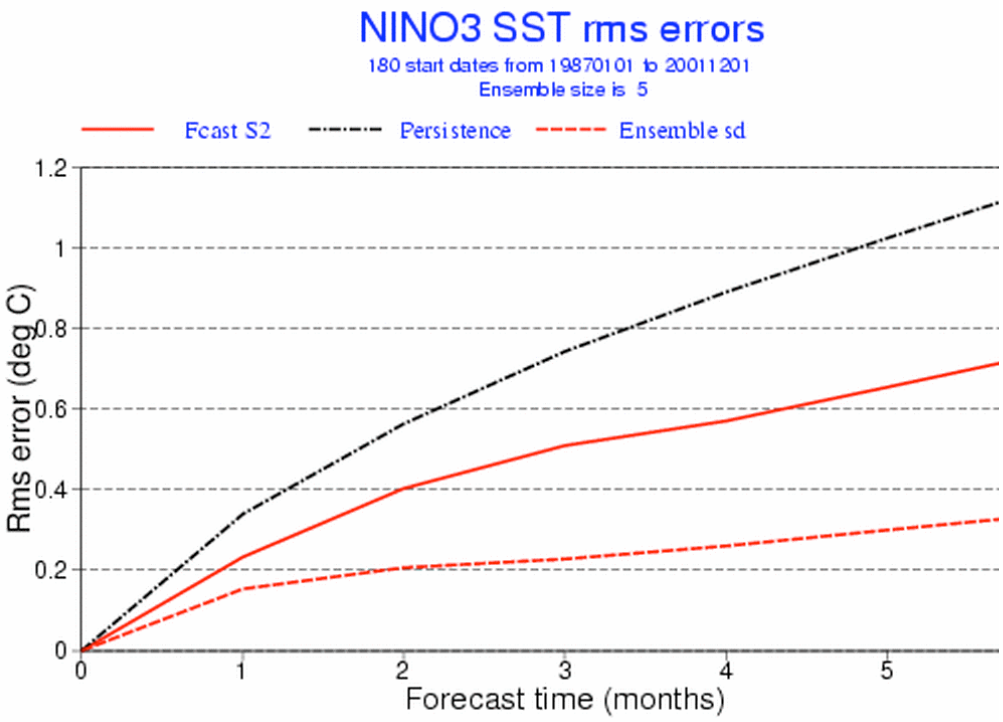
Forecast sensitive to ic mean state

Model error in ENSO forecasts

- **Model error appears to dominate ENSO forecasts**
- **But initial condition error is a little underestimated**
  - Comparison of forecasts made with different ocean analysis systems suggests effect small where ocean mean state is similar
  - Analyses with substantial differences in ocean mean state (eg wind forced vs in situ-data) give bigger forecast differences in west-central Pacific
- **Model error affects initial conditions**
  - Errors in atmosphere model physics (surface processes, boundary layer, convection ..) affect surface forcing from analyses / reanalyses
  - Errors in ocean models affect the assimilation of observed data

Impact of model error on predictability

- Model underestimates unforced wind variability (noise) – especially MJO
- So true predictability limit of rms error will be higher than the model estimate
- Effect for S2 was estimated to be quite significant



Best estimate of Nir
3 SST predictability
6 months = 0.45 (0.4
> 0.55 °C)

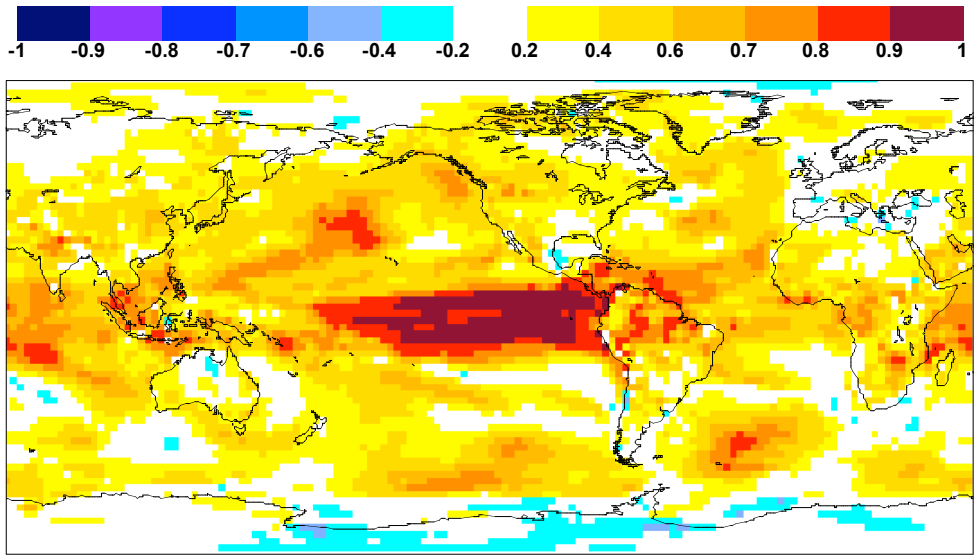
ECMWF

How good are the forecasts?

Deterministic skill: DJF ACC

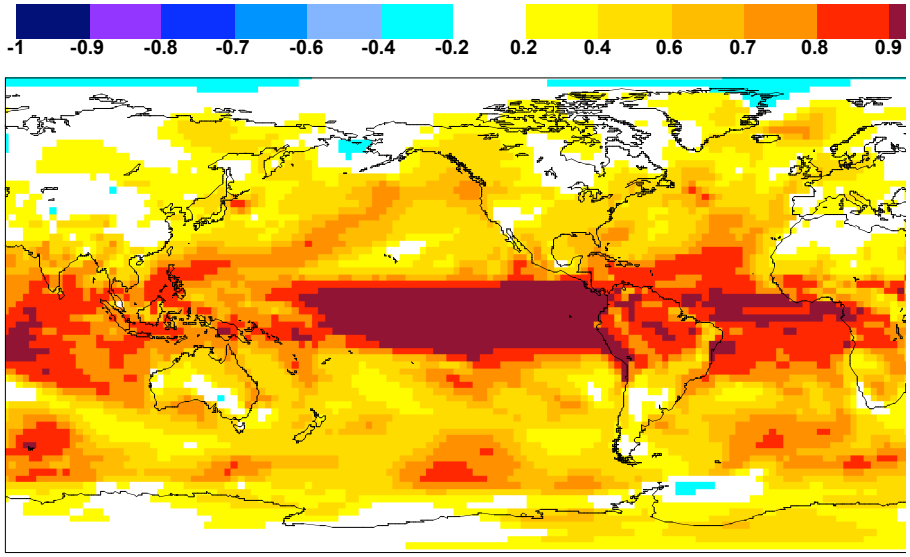
Temperature: actual forecasts

Anomaly Correlation Coefficient for CodOecmfE0001S003M001 with 11 ensemble members
Near-surface temperature
Hindcast period 1981-2003 with start in November and averaging period 2 to 4



Temperature: perfect model

Perfect-model Anomaly Correlation Coefficient for CodOecmfE0001S003M001 with 11 ensemble membe
Near-surface temperature
Hindcast period 1981-2003 with start in November and averaging period 2 to 4



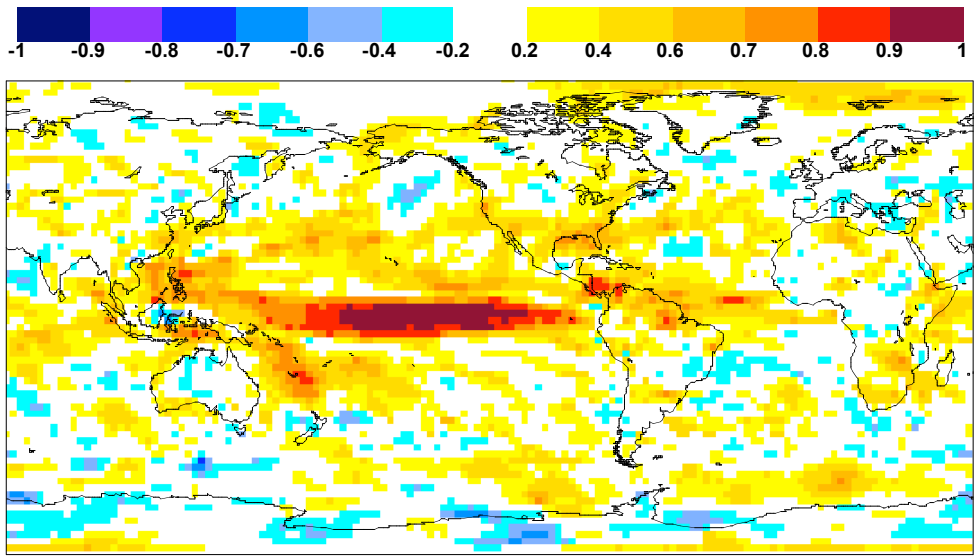
ECMWF

How good are the forecasts?

Deterministic skill: DJF ACC

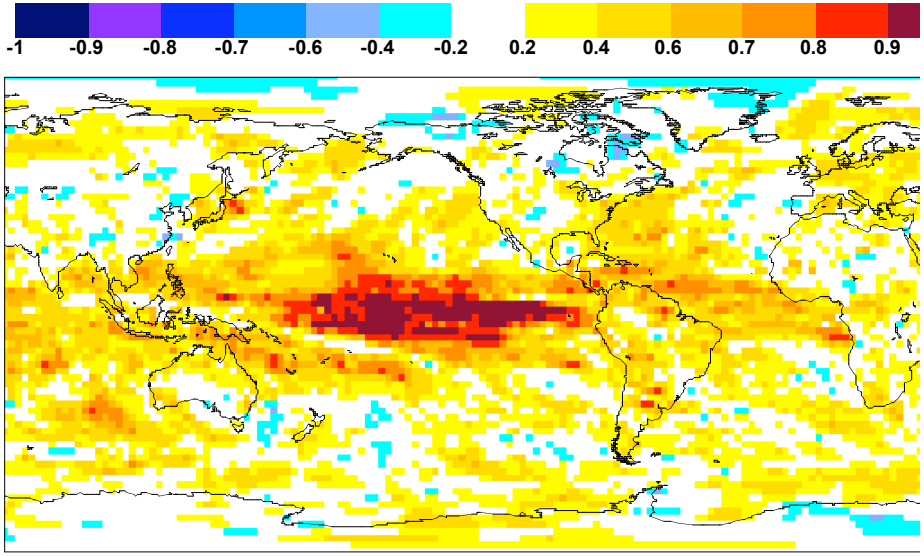
Precip: actual forecasts

Anomaly Correlation Coefficient for CodOecmfE0001S003M001 with 11 ensemble members
Precipitation
Hindcast period 1981-2003 with start in November and averaging period 2 to 4



Precip: perfect model

Perfect-model Anomaly Correlation Coefficient for CodOecmfE0001S003M001 with 11 ensemble members
Precipitation
Hindcast period 1981-2003 with start in November and averaging period 2 to 4

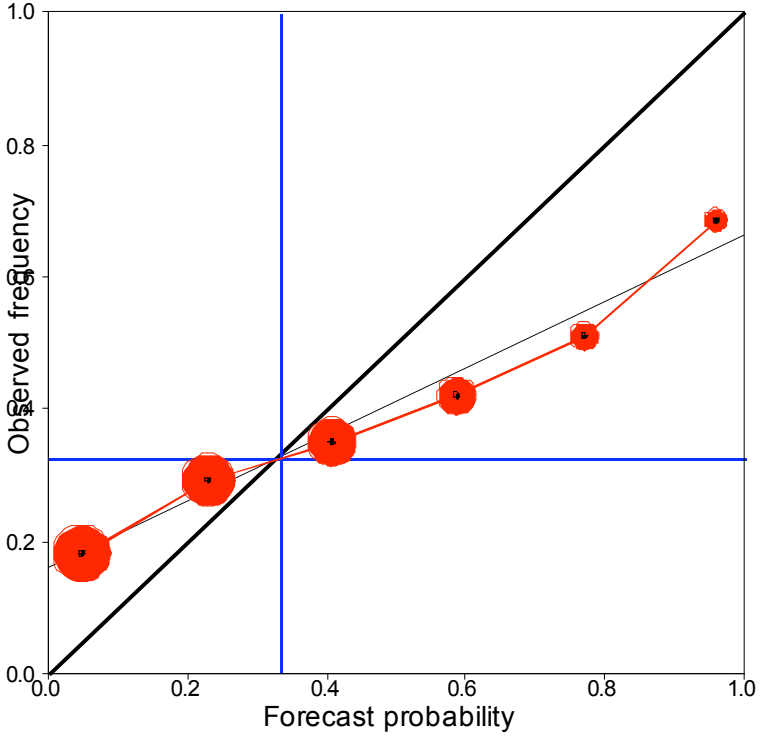


ECMWF

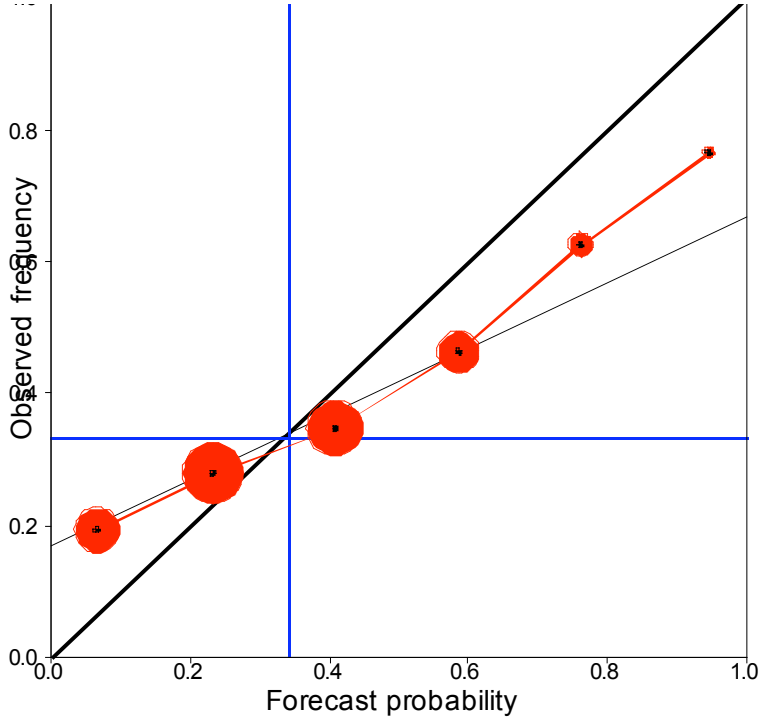
How good are the forecasts?

Probabilistic skill: Reliability diagrams

Tropical precip < lower tercile, JJA



NH extratrop temp > upper tercile, DJF

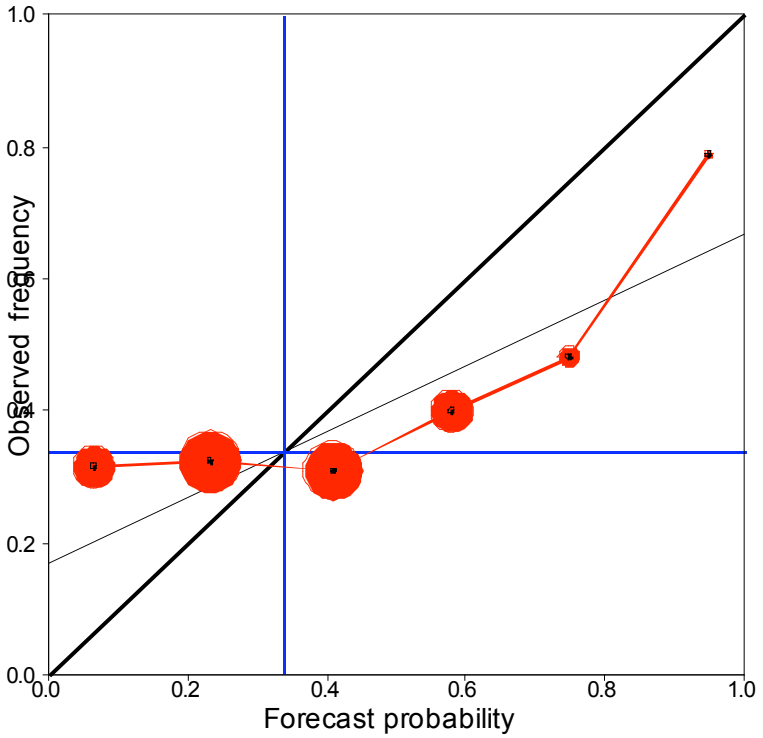


ECMWF

How good are the forecasts?

Probabilistic skill: Reliability diagrams

Europe: Temp > upper tercile, DJF



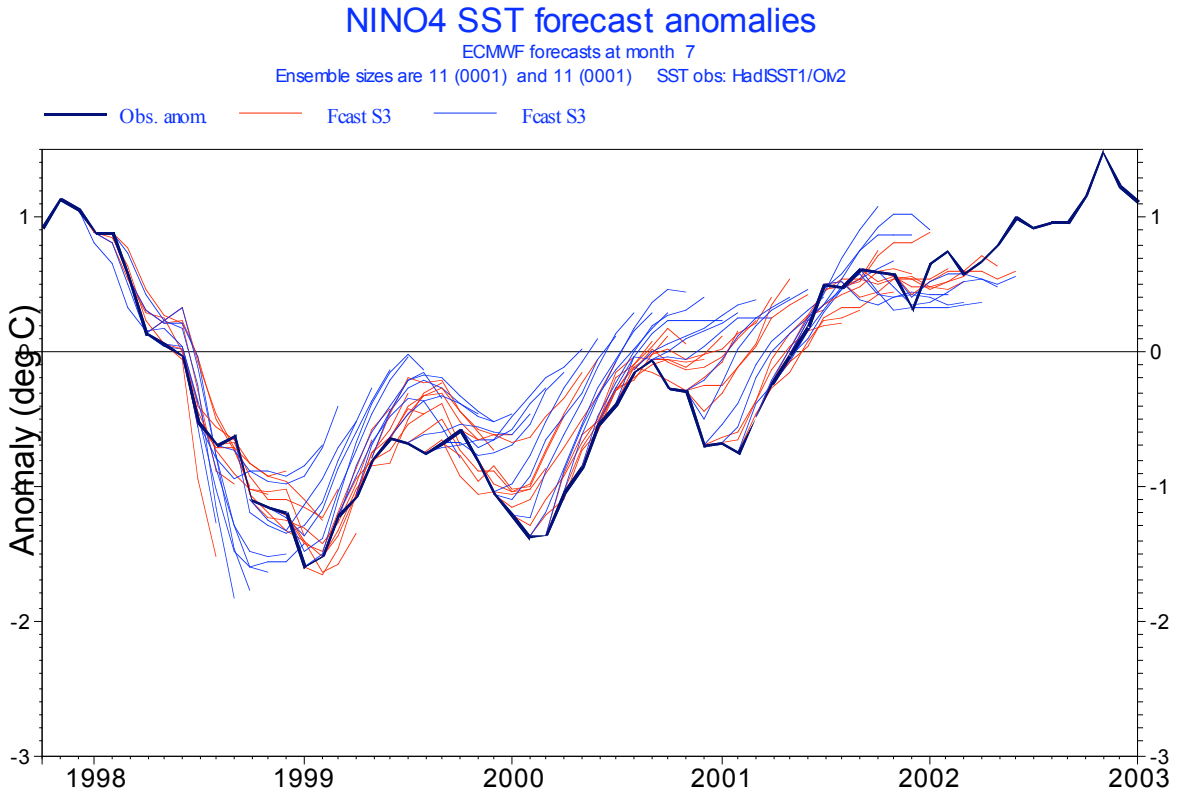
ECMWF

Side-stepping model error?

- **Correction terms in model integrations**
  - Conceptually not very nice at all, but a pragmatic forecaster might consider them.
  - Our experience is that it is possible to get small improvements in ENSO skill, but not easily, and not a big gain.
- **We have tried ….**
  - Heat flux corrections
  - Multiplicative wind stress corrections
  - Additive wind stress corrections, idealized or to ERA40
  - Ocean increments, derived from analyses

Side-stepping model error?

- **Statistical massaging**
  - Adjust amplitude of signal, or use model as input to statistical schemes



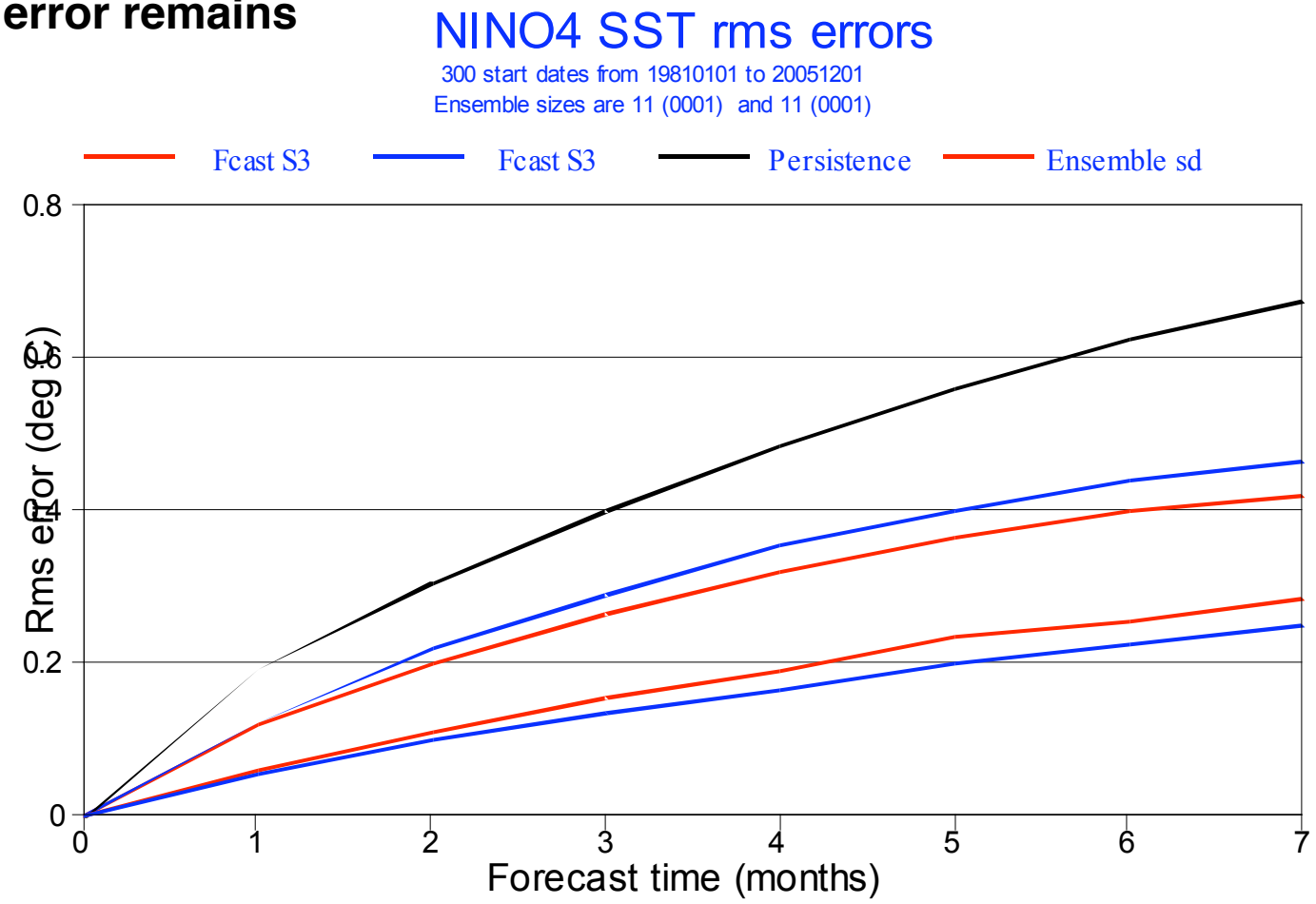
Blue: original

Red: corrected

ECMWF

Red: **statistically corrected forecast**

Improved rms errors show state-dependent forecast errors can be partially removed, but most model error remains



ECMWF

Side-stepping error: multi-model

ot sampling equally likely outcomes

e expect all models to be wrong, none of them to be possible solutions

re sampling model-induced errors in predicted outcomes of individual events

f these errors are independent, then a large multi-model ensemble might have a *small* error in the ensemble mean, much smaller than the inter-model spread

- **Errors are patently not independent**

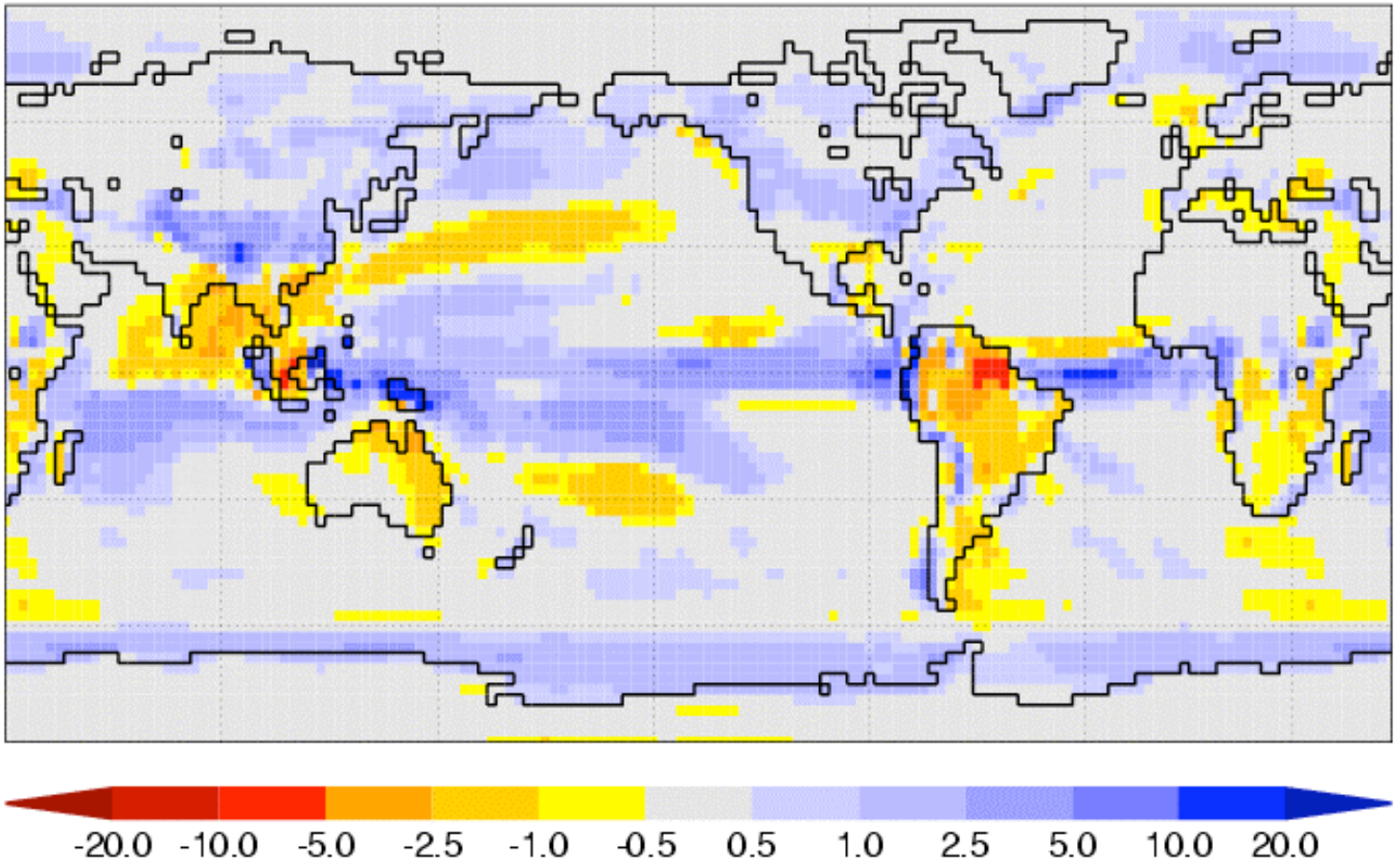
ECMWF precip bias

Total Precipitation [mm day^{-1}]

Bias: EXP(ECMWF_assim) regarding GPCP

Forecast start month and years: February / 1987-2001

FC period: months 2-4 (MAM), ens: 0-8



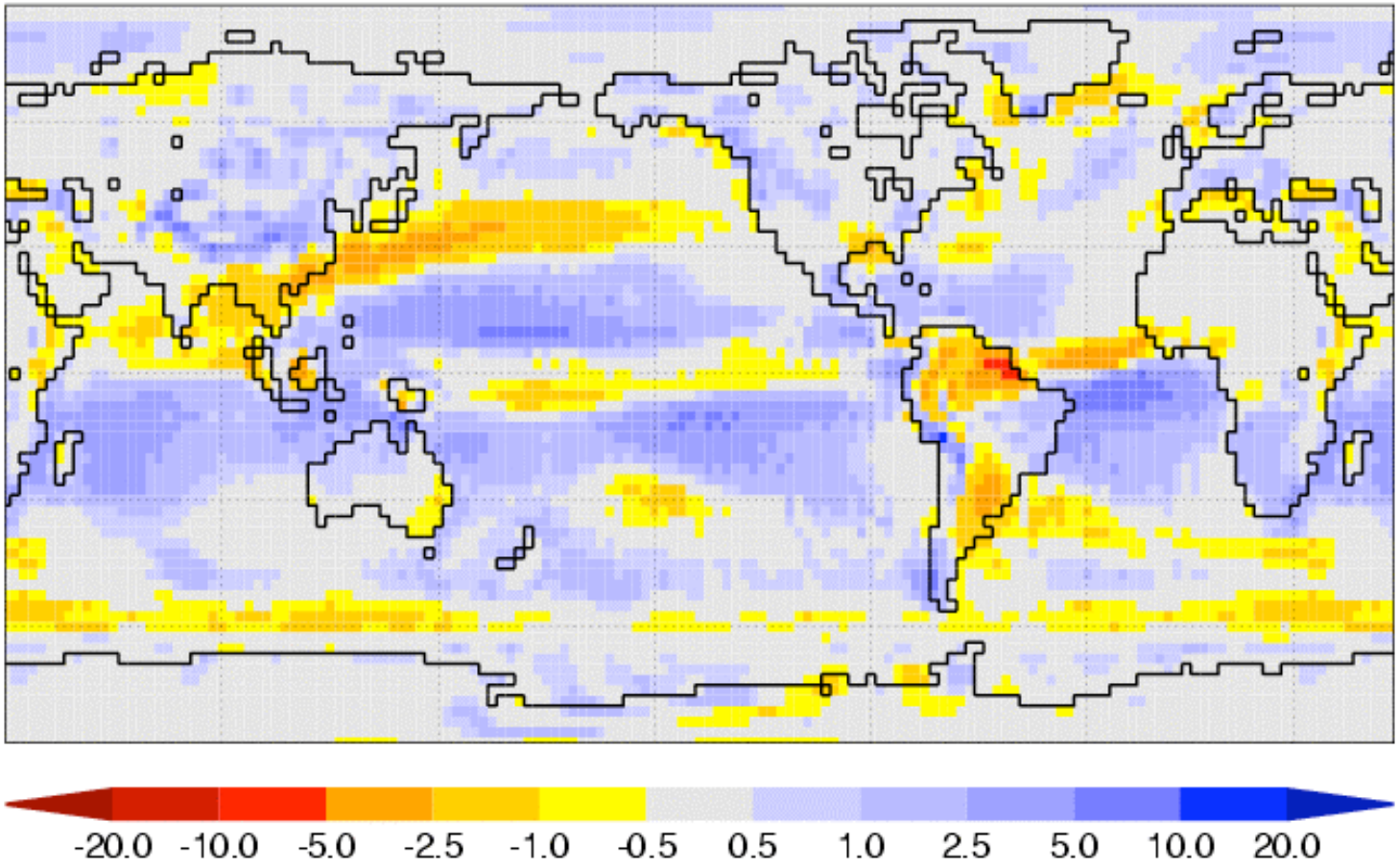
MeteoFrance precip bias

Total Precipitation [mm day^{-1}]

Bias: EXP(CNRM) regarding GPCP

Forecast start month and years: February / 1979-2001

FC period: months 2-4 (MAM), ens: 0-8



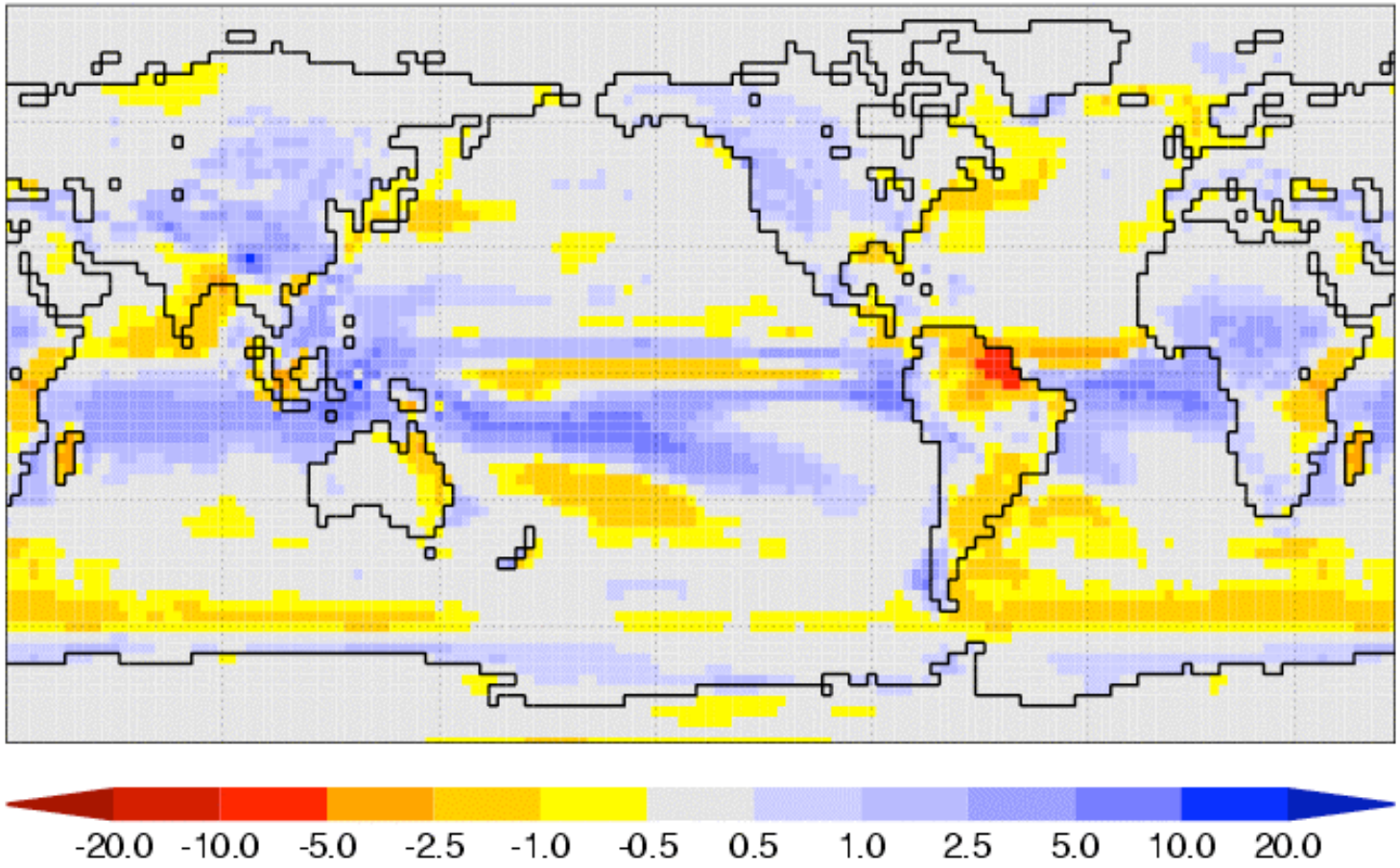
UKMO precip bias

Total Precipitation [mm day^{-1}]

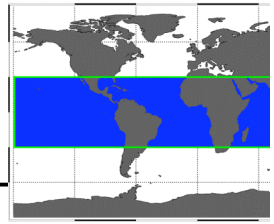
Bias: EXP(UKMO) regarding GPCP

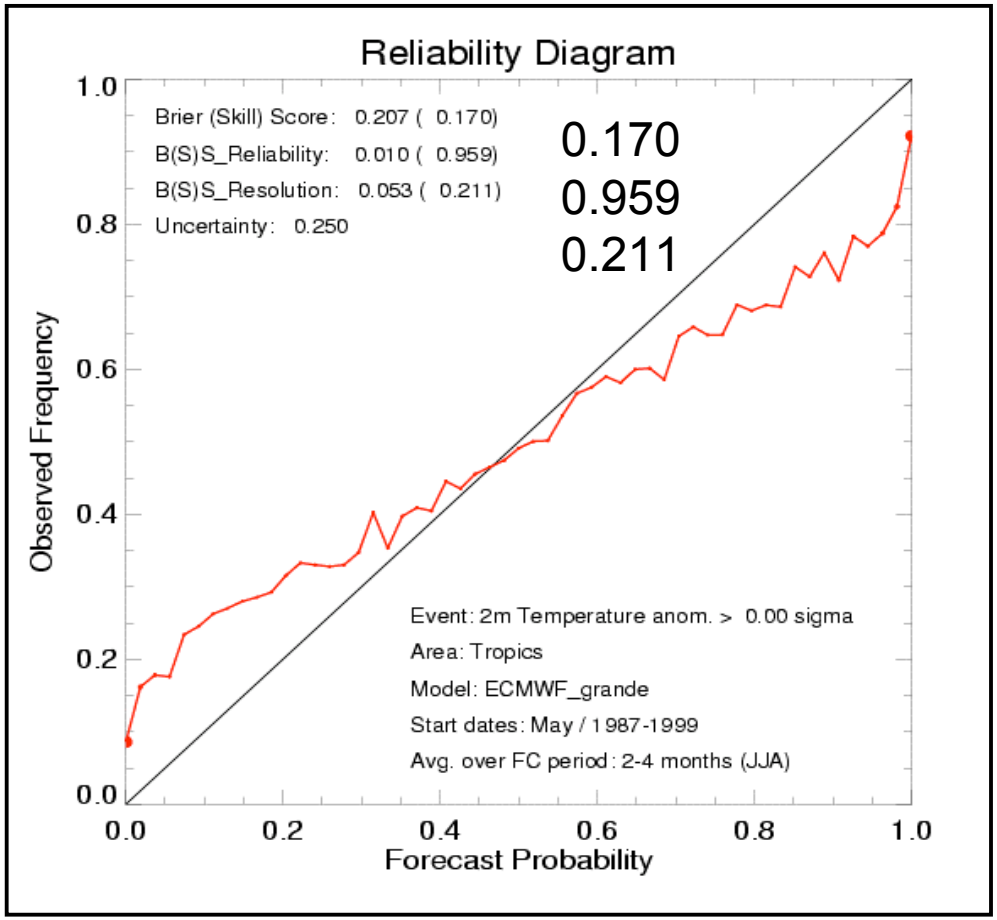
Forecast start month and years: February / 1979-2001

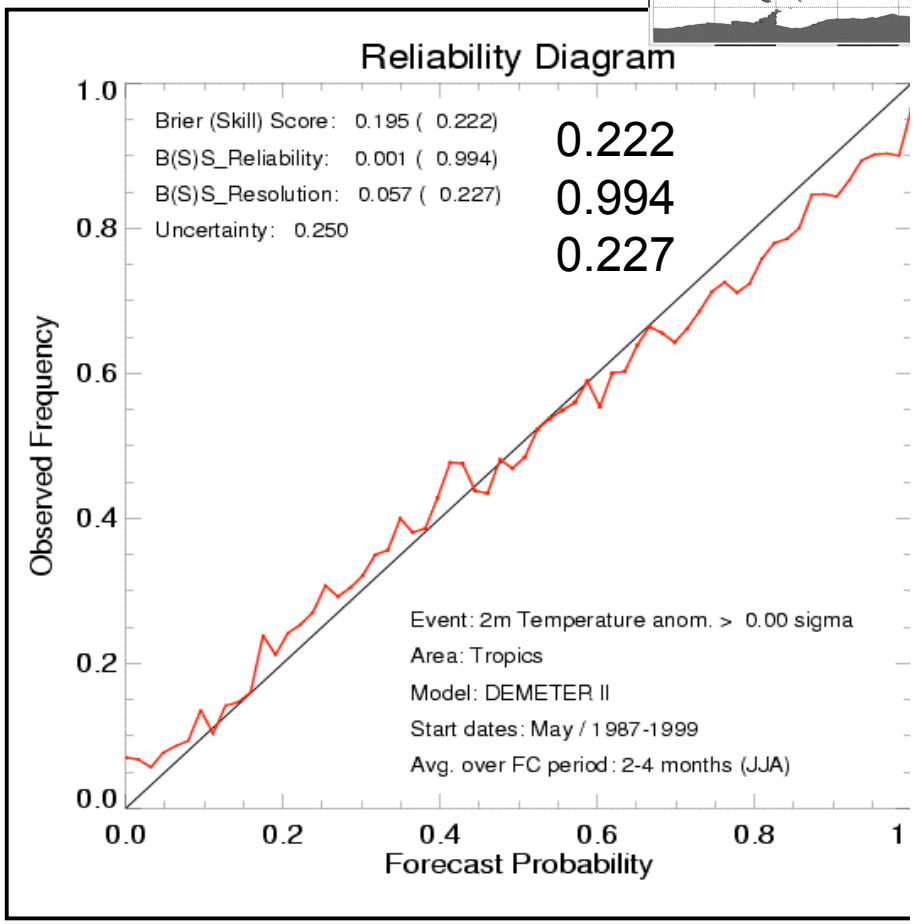
FC period: months 2-4 (MAM), ens: 0-8



Reliability diagram: 2m-Temp.>0







single-model (54 members)

multi-model

ECMWF

Side-stepping error: multi-model

ot sampling equally likely outcomes

e expect all models to be wrong, none of them to be possible solutions

re sampling model-induced errors in predicted outcomes of individual events

f these errors are independent, then a large multi-model ensemble might have a *small* error in the ensemble mean, much smaller than the inter-model spread

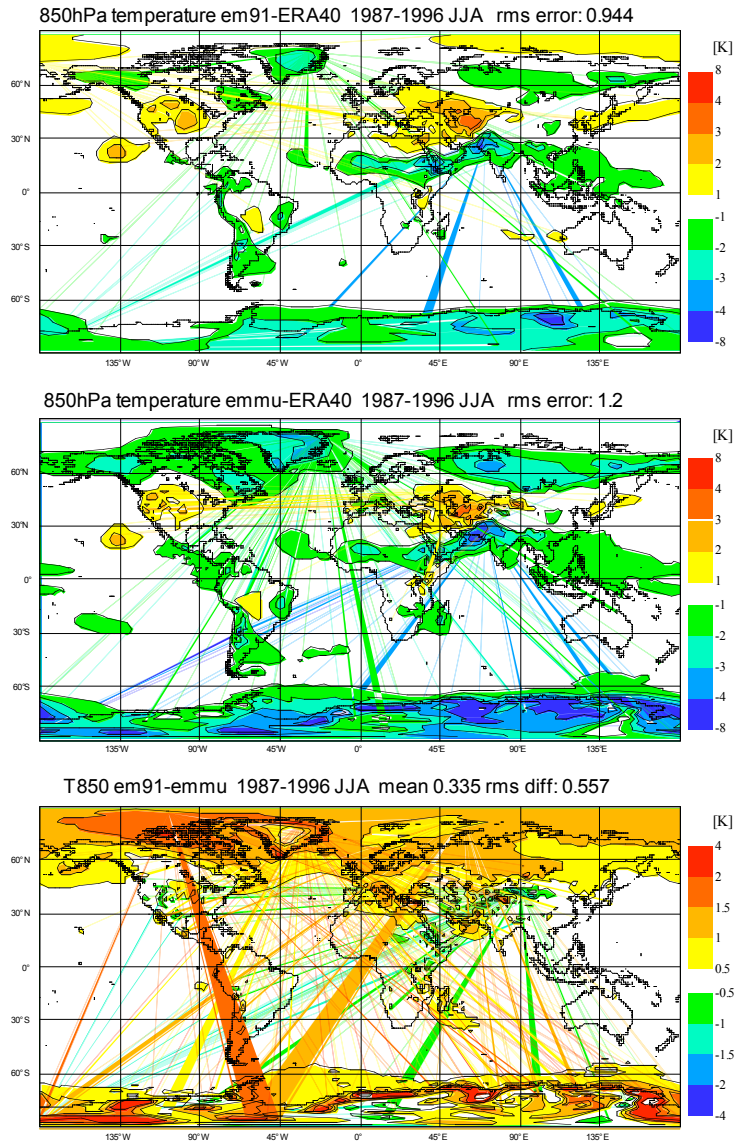
- **Errors are patently not independent**
- **Multi-model approach is worthwhile ….**

… but no substitute for better models

Specific comments on model error

- Resolution
- Timestep
- Interactions in the West-Central Pacific
- Coupling of surface currents

Model biases and resolution



T159 vs T95

Resolution matters! (or it did!!)

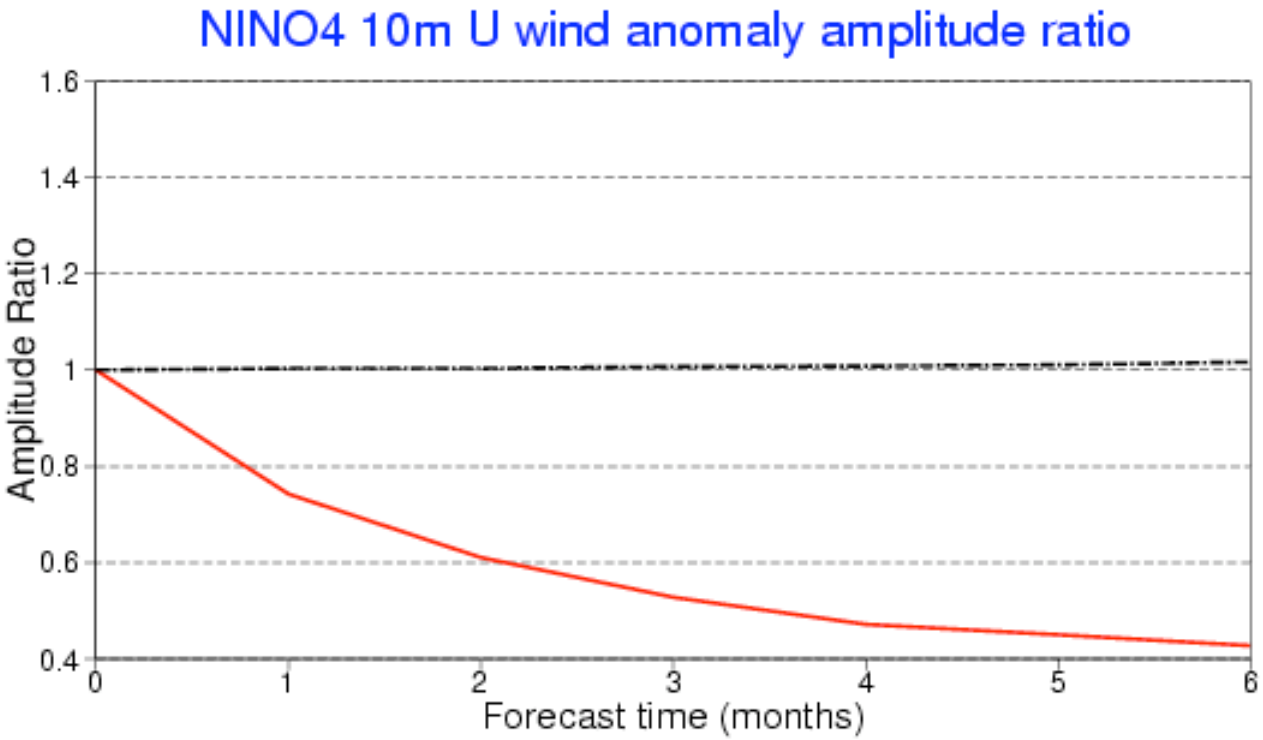
Horizontal resolution affects temperature structure and equatorial winds

Vertical resolution affects surface fluxes and SST, and forecast skill in East Pacific (no help in west)

Time-step dependence was a real issue, but was reduced some time ago

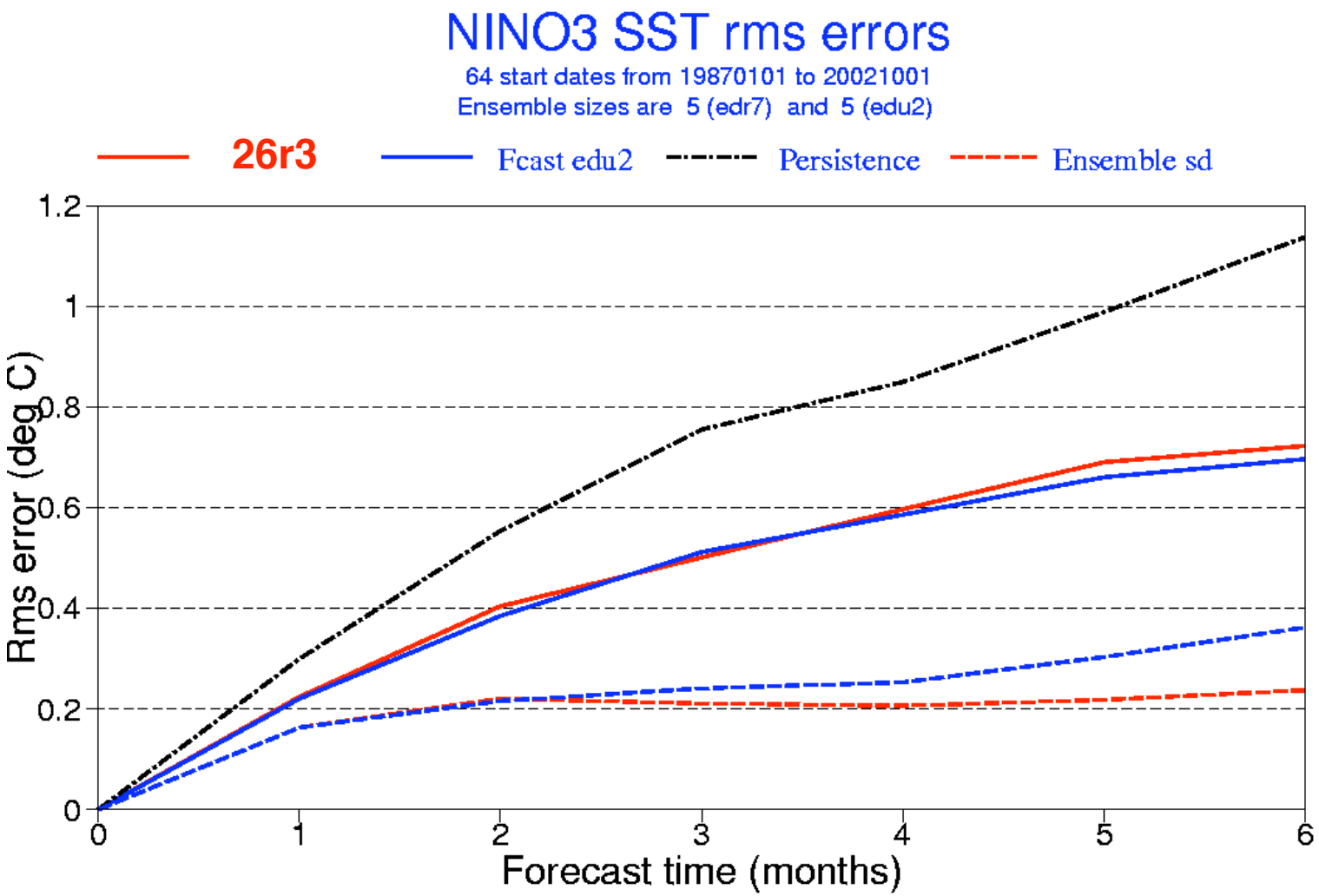
Interactions in the west-central Pacific

Variability of zonal wind in System 2 was strongly damped, even in the early months when SST is close to observed



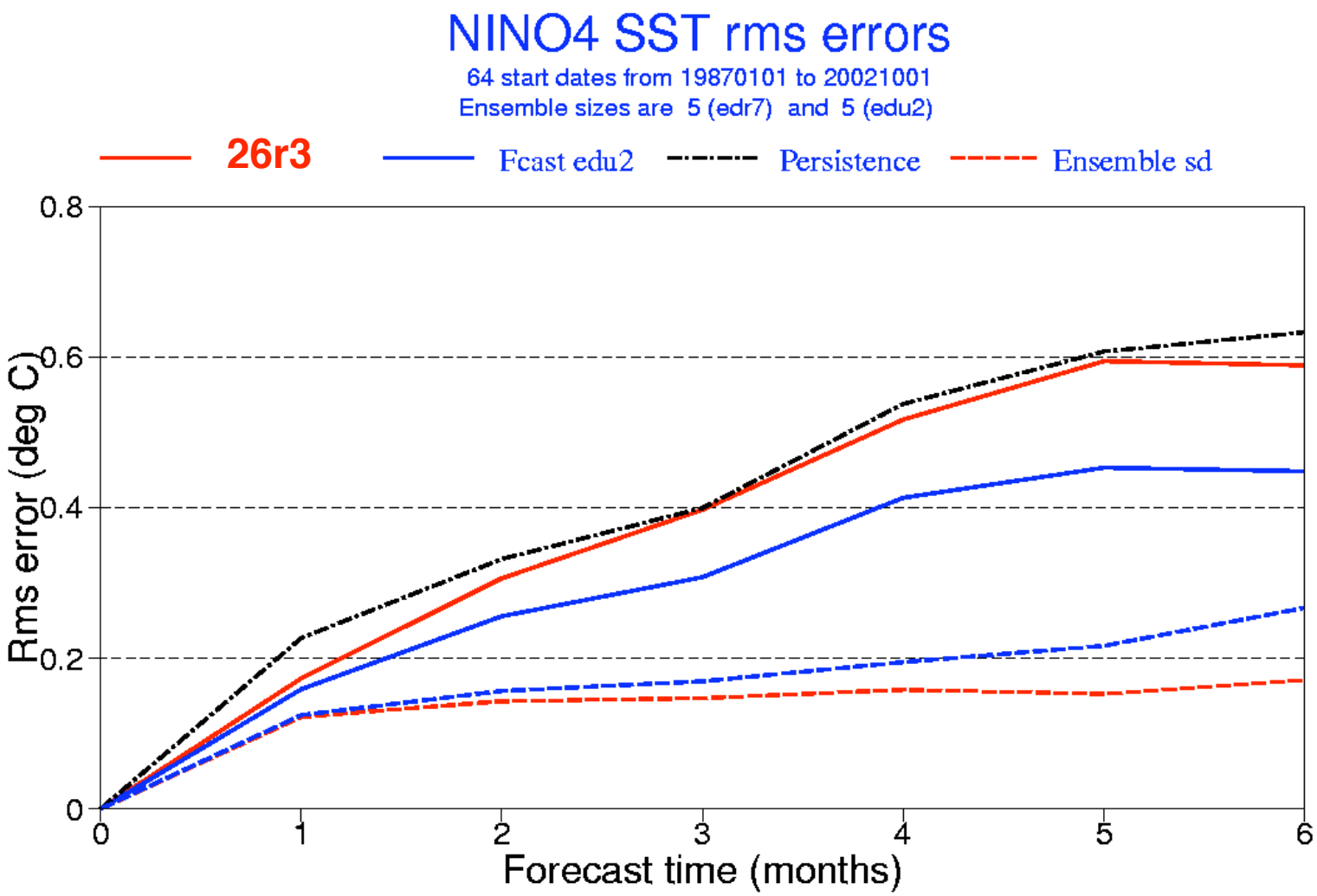
ECMWF

Then along came Cycle 26r3 ....

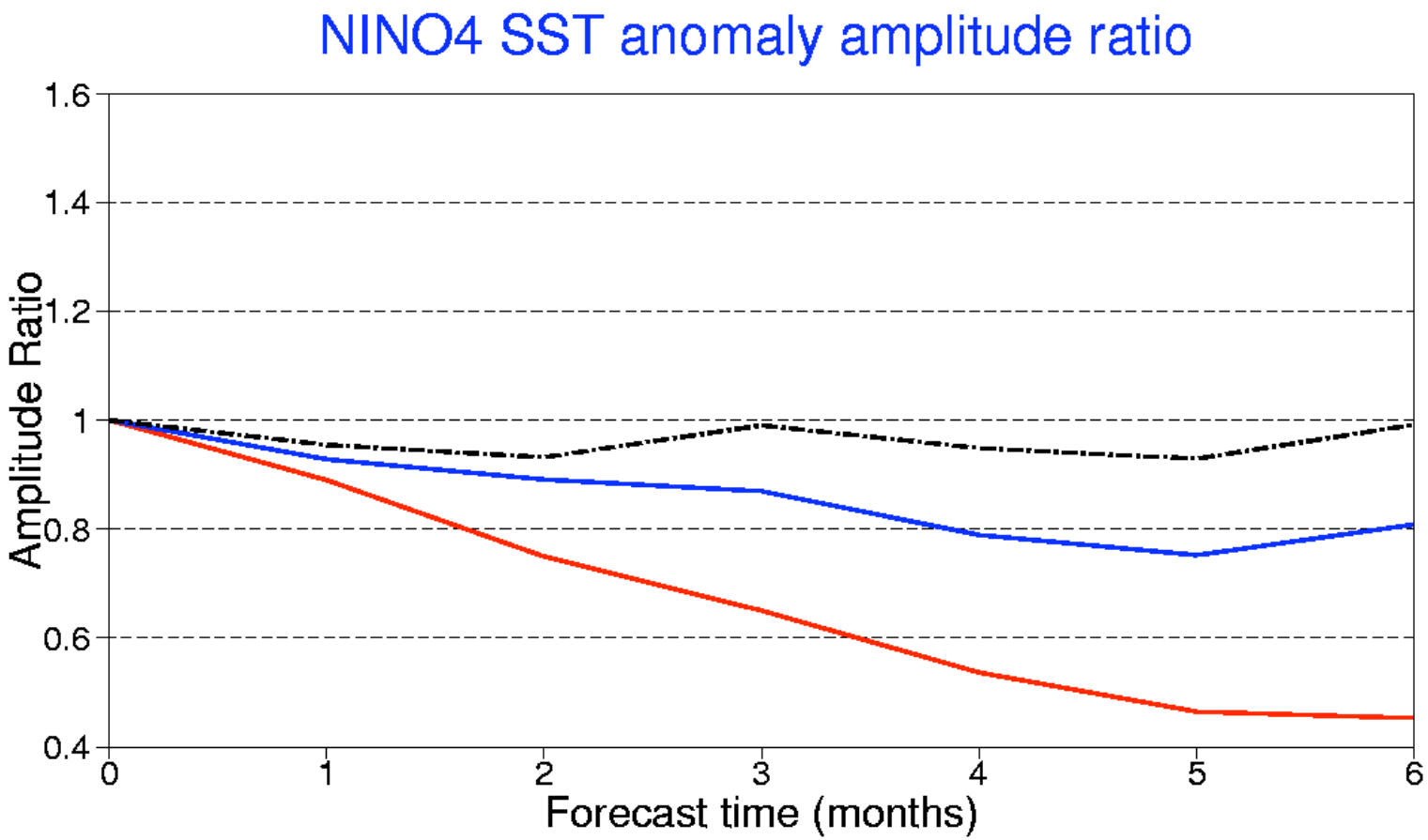


ECMWF

… with catastrophic performance in Nino 4

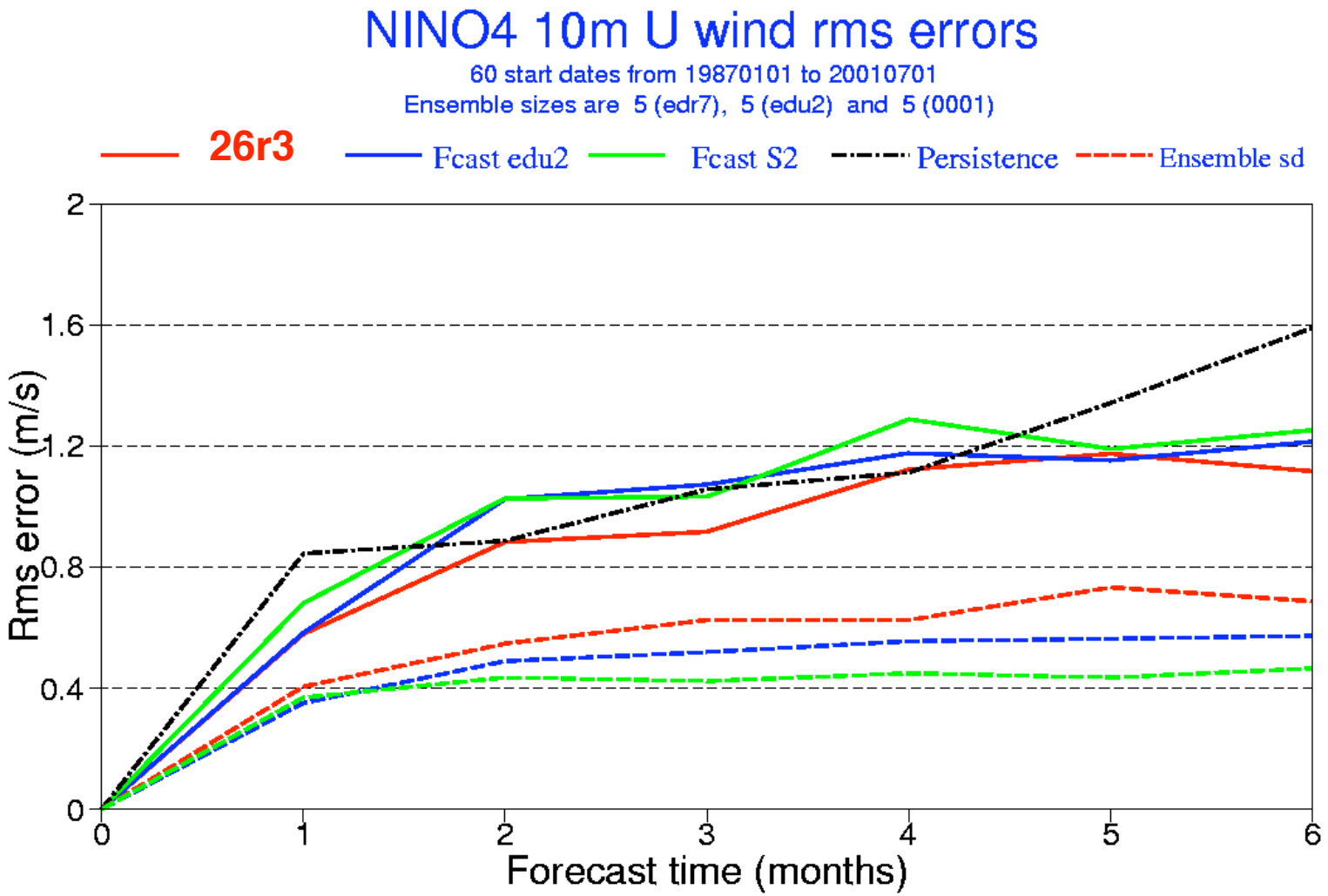


ECMWF


NINO4 SST anomaly amplitude ratio
Amplitude Ratio
1.6
1.4
1.2
1
0.8
0.6
0.4
0
1
2
3
4
5
6
Forecast time (months)

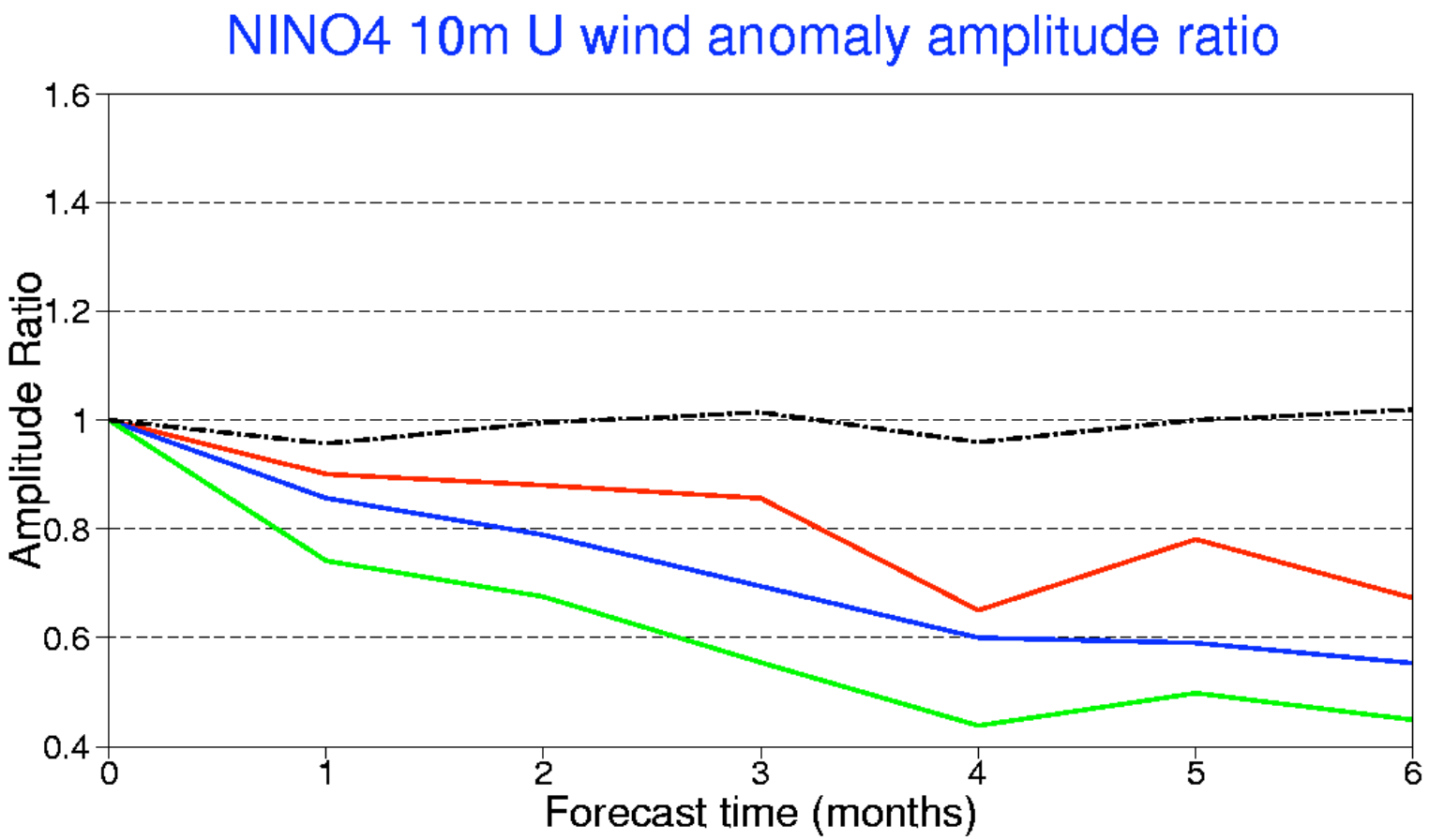
26r3

But wind errors in 26r3 are improved ...



ECMWF

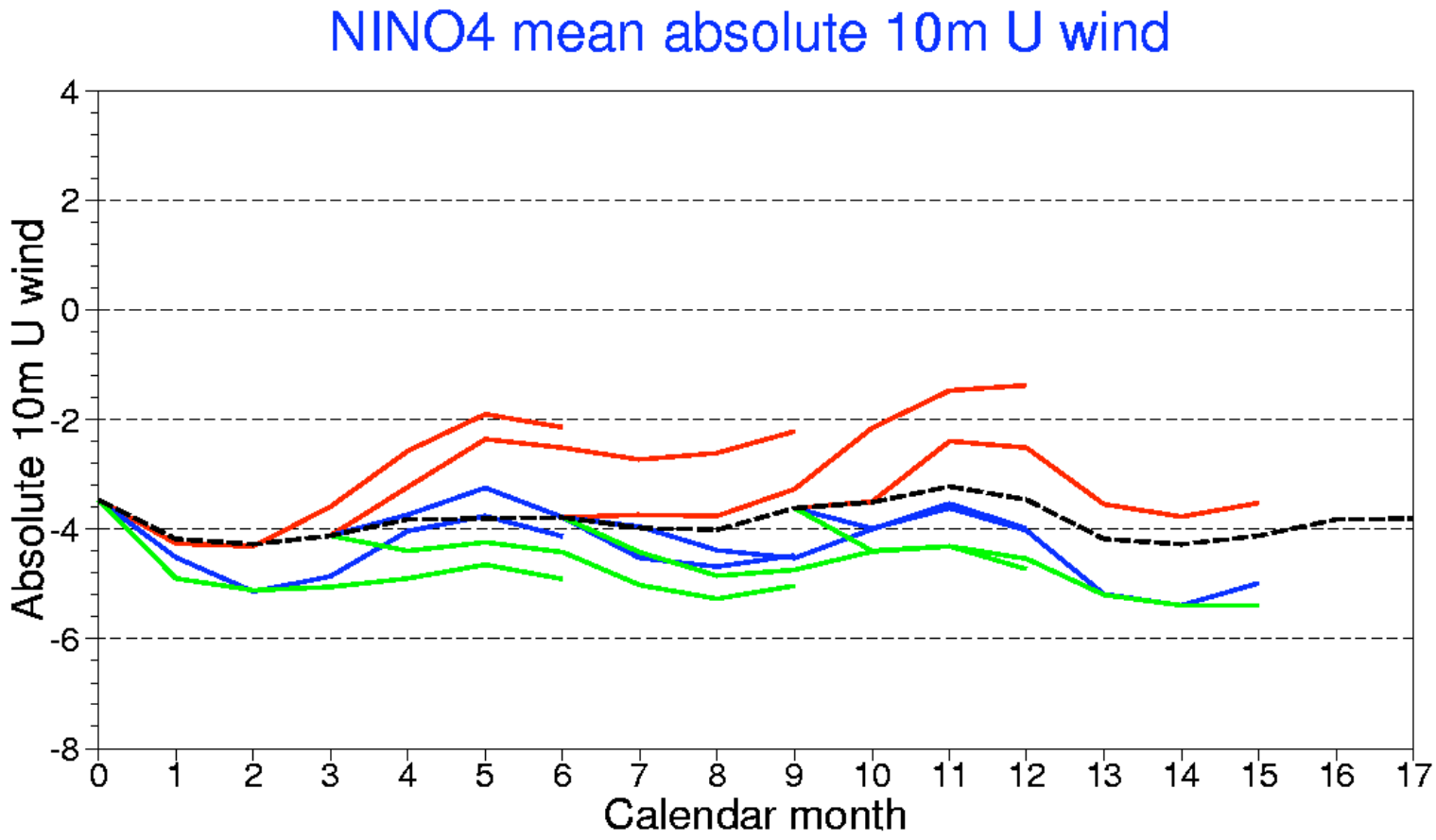
… as is the variability



26r3

ECMWF

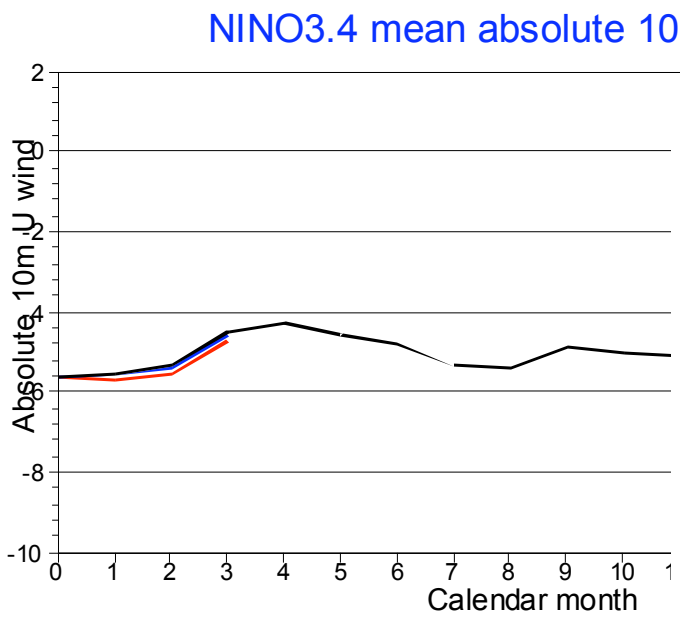
The explanation is the mean state ...

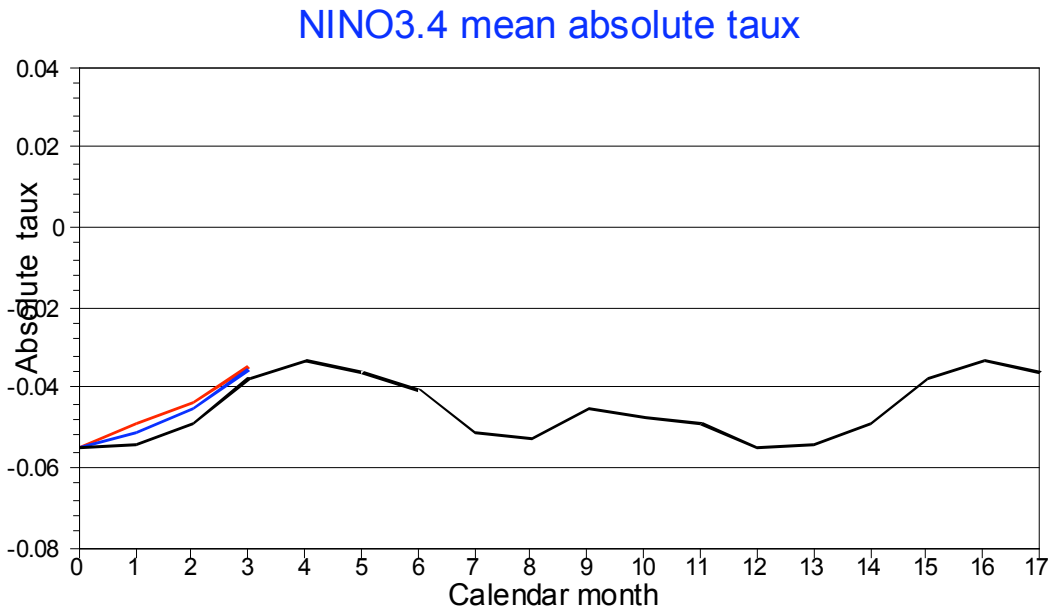


26r3

ECMWF

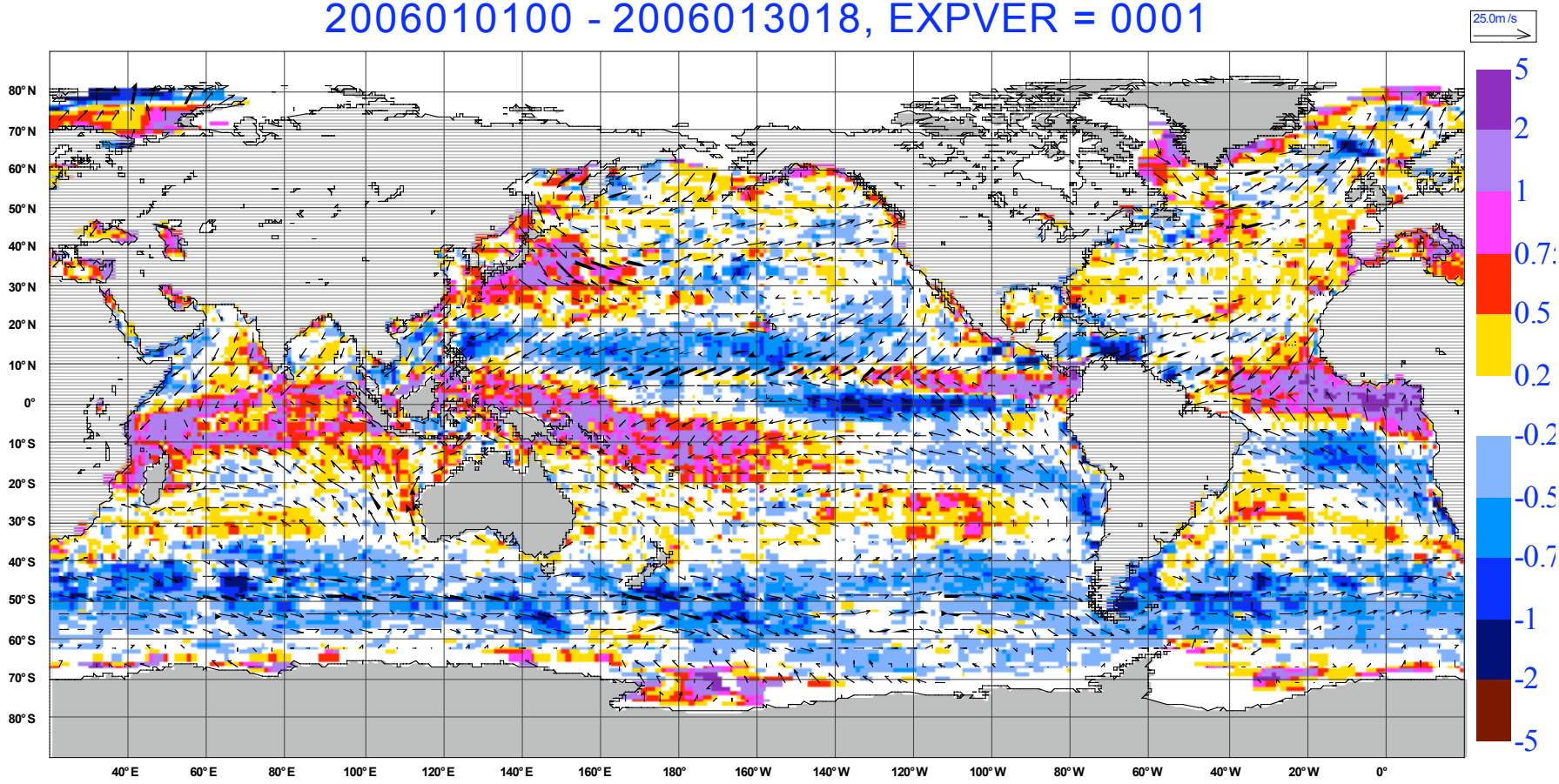
Coupled surface currents: U10 and Taux





QuickSCAT bias wrt ECMWF first guess

Wind speed bias (m/s) of QuikSCAT vs ECMWF FGAT for all flows
Globe 0.05 N.Hem 0.16 Tropics 0.17 S.Hem -0.13 MIN -2.93 MAX 3.8
2006010100 - 2006013018, EXPVER = 0001



ECMWF

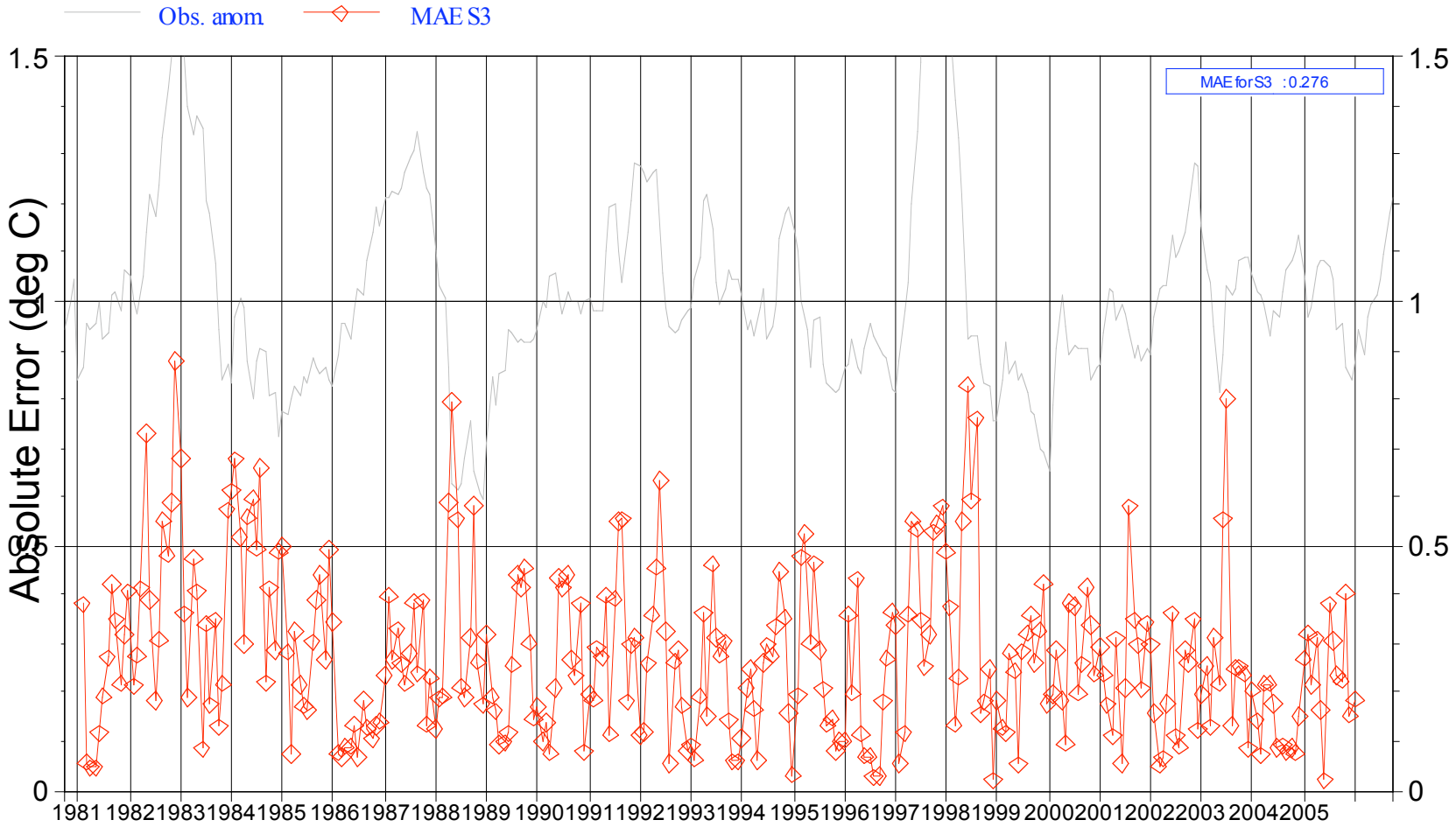
Conclusions

- **Model error dominates seasonal forecasts**
  - It is the biggest contributor to forecast error
  - Explains why it is still hard to show benefit of observing system with today's models
  - Seasonal forecast systems are mature enough to provide a metric
- **Model error is difficult to side-step**
  - It can be ameliorated, but only partially
- **Models are improving, slowly**
  - Could it be faster with more focus on seasonal forecast issues??
- **If our models are inadequate after 6 months ….**

  …. can we trust them decades into the future??

NINO3 SST absolute error scores

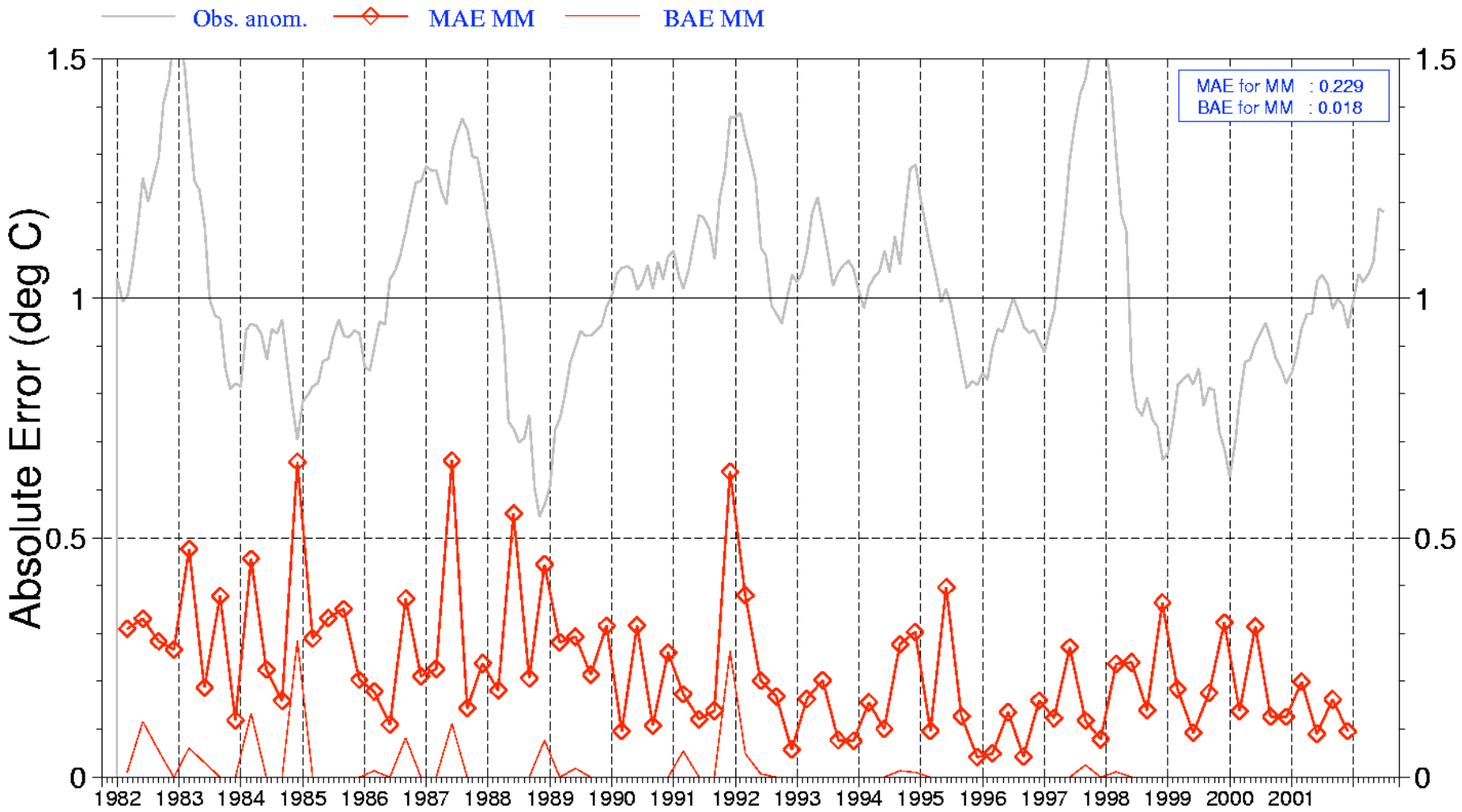
ECMWF forecasts (mean during 3 months, plotted at centre of verification period)
Ensemble size is 11 SST obs: HadISST1/OIv2



ECMWF

NINO3.4 SST absolute error scores

ECMWF forecasts (mean during 3 months, plotted at centre of verification period)
Ensemble size is 18 SST obs: NCEP Olv2



ECMWF