

Long range predictability of winter circulation

Tim Stockdale, Franco Molteni and Laura Ferranti
ECMWF

Outline

- **ECMWF System 4**
- **Predicting the Arctic Oscillation and other modes**
- **Atmospheric initial conditions**
- **Conclusions and cautions**

System 4 configuration

● Real time forecasts:

- **51 member ensemble forecast to 7 months**
- SST and atmos. perturbations added to each member

- **15 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 1981-2010 (**30 years**)

- 15 member ensemble every month
- 15 members extended to 13 months once per quarter

- **51 members** for Feb/May/Aug/Nov starts

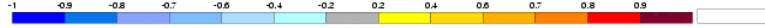
ENSO forecasts are good

1981-1995

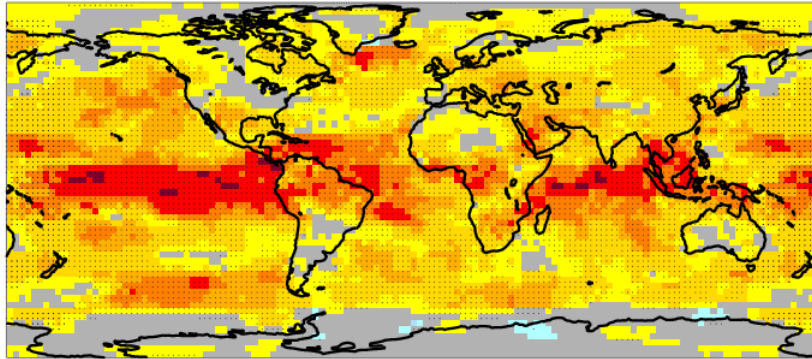
1996-2010

So are deterministic scores in the tropics

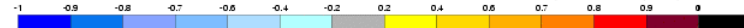
Anomaly Correlation Coefficient for ECMWF with 15 ensemble members
Near-surface air temperature
Hindcast period 1981-2010 with start in February average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (1000 samples)



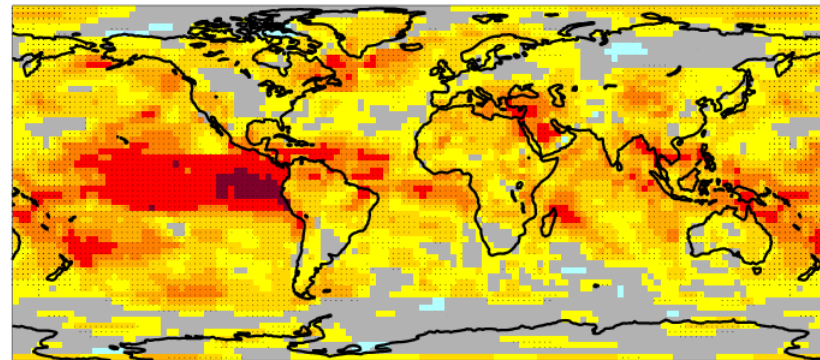
MAM



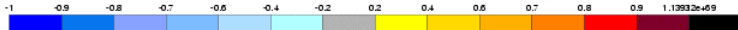
Anomaly Correlation Coefficient for ECMWF with 15 ensemble members
Near-surface air temperature
Hindcast period 1981-2010 with start in May average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (1000 samples)



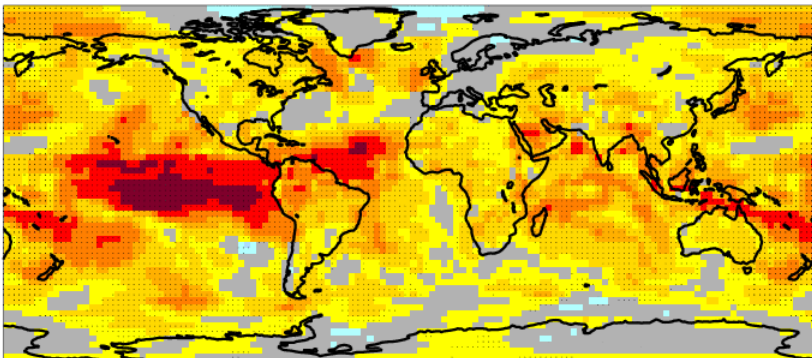
JJA



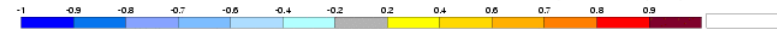
Anomaly Correlation Coefficient for ECMWF with 15 ensemble members
Near-surface air temperature
Hindcast period 1981-2010 with start in August average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (1000 samples)



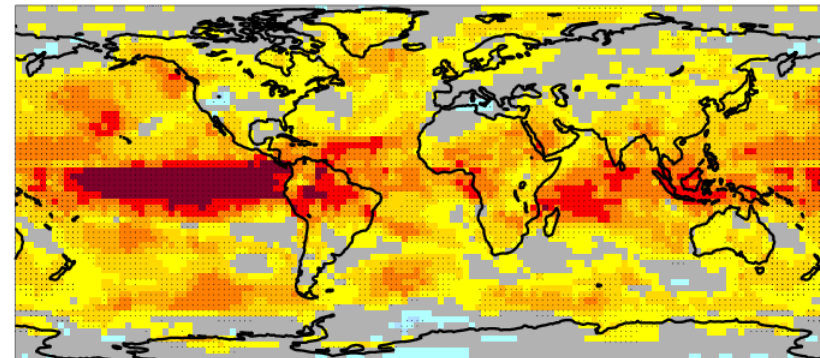
SON



Anomaly Correlation Coefficient for ECMWF with 15 ensemble members
Near-surface air temperature
Hindcast period 1981-2010 with start in November average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (1000 samples)



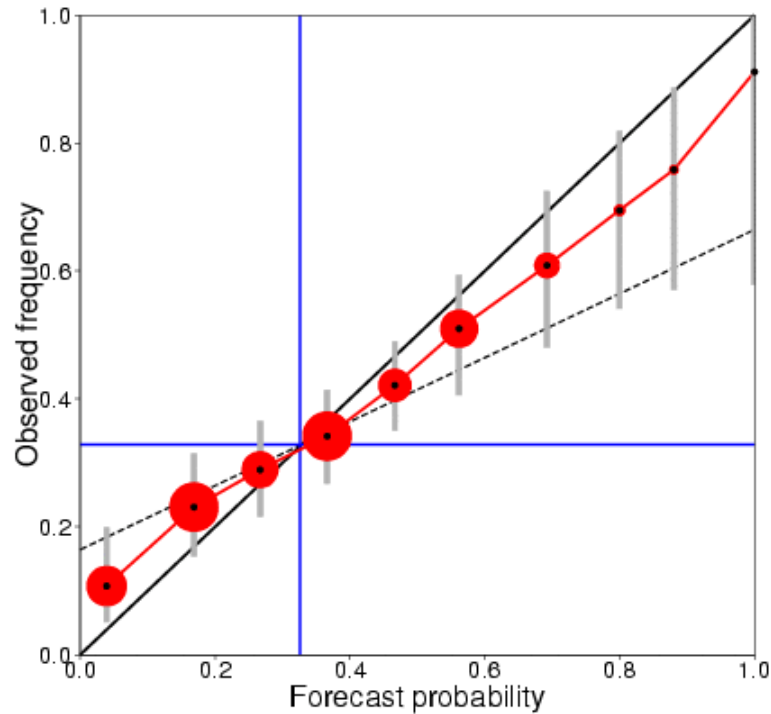
DJF



So are probabilistic scores

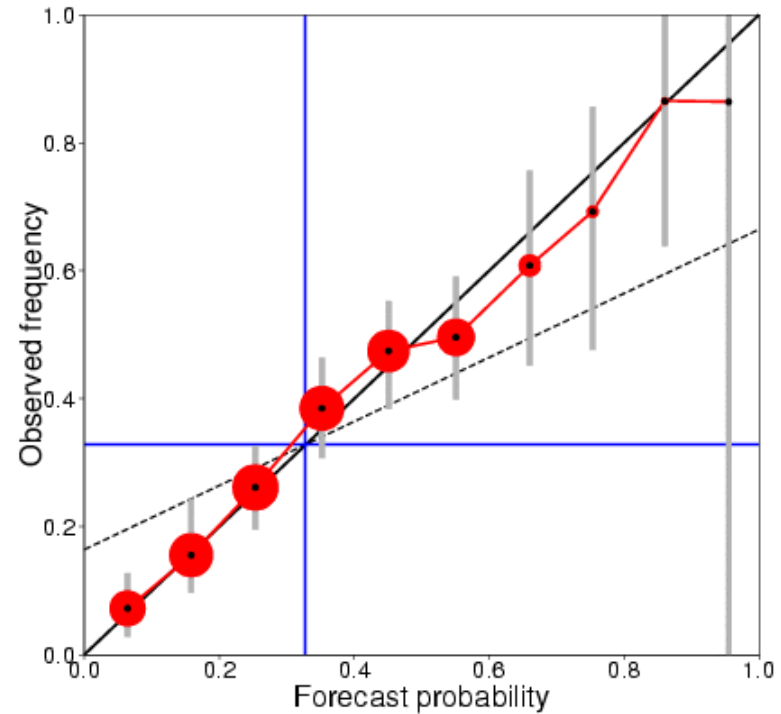
15 members

JJA Europe T2m > upper tercile
Re-forecasts from 1 May, 1981-2010
Reliability score: 0.987
ROC skill score: 0.38



51 members

JJA Europe T2m > upper tercile
Re-forecasts from 1 May, 1981-2010
Reliability score: 0.996
ROC skill score: 0.43

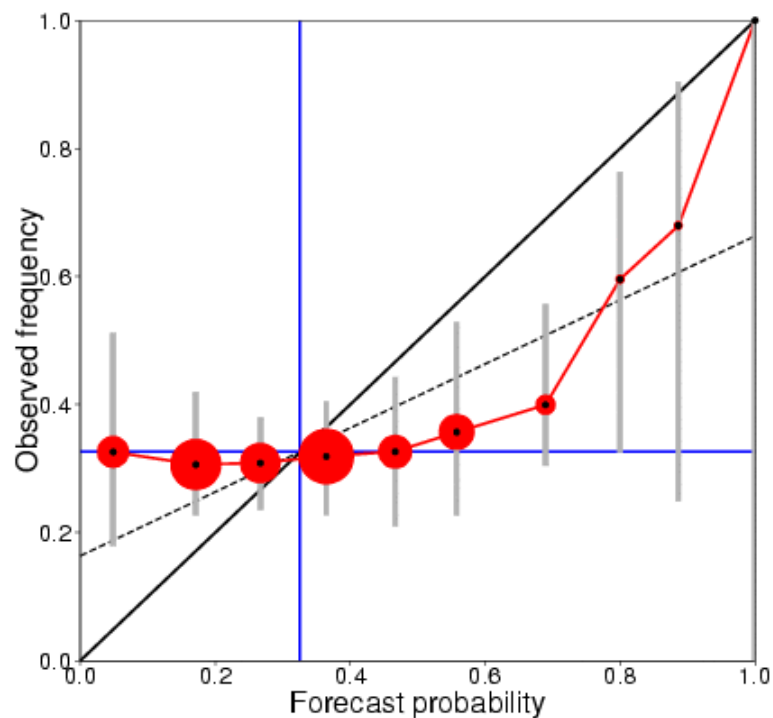


(Figures from Susanna Corti)

Ensemble size important for low-signal areas

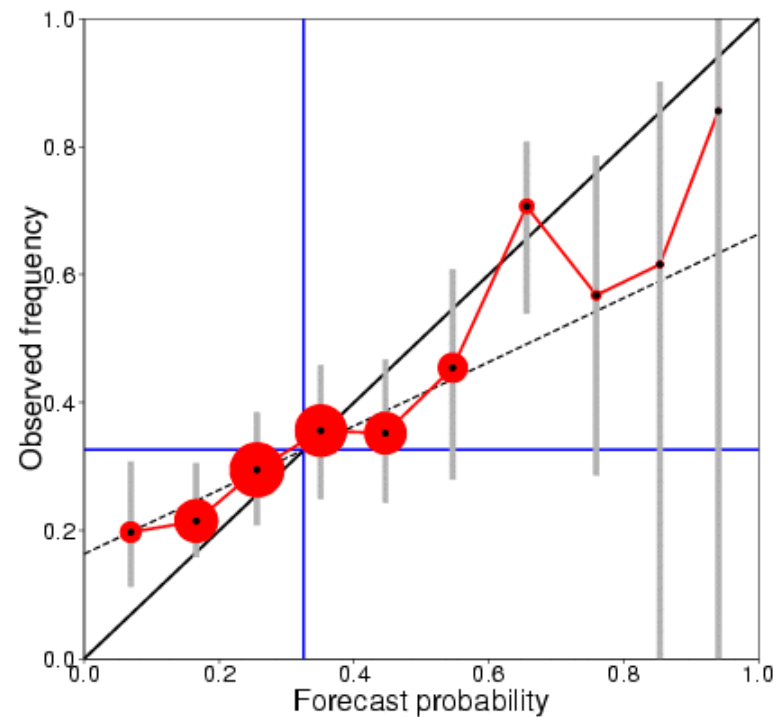
15 members

DJF Europe T2m > upper tercile
Re-forecasts from 1 Nov, 1981-2010
Reliability score: 0.902
ROC skill score: 0.06



51 members

DJF Europe T2m > upper tercile
Re-forecasts from 1 Nov, 1981-2010
Reliability score: 0.981
ROC skill score: 0.22



(Figures from Susanna Corti)

Stratosphere is also OK

Arctic Oscillation

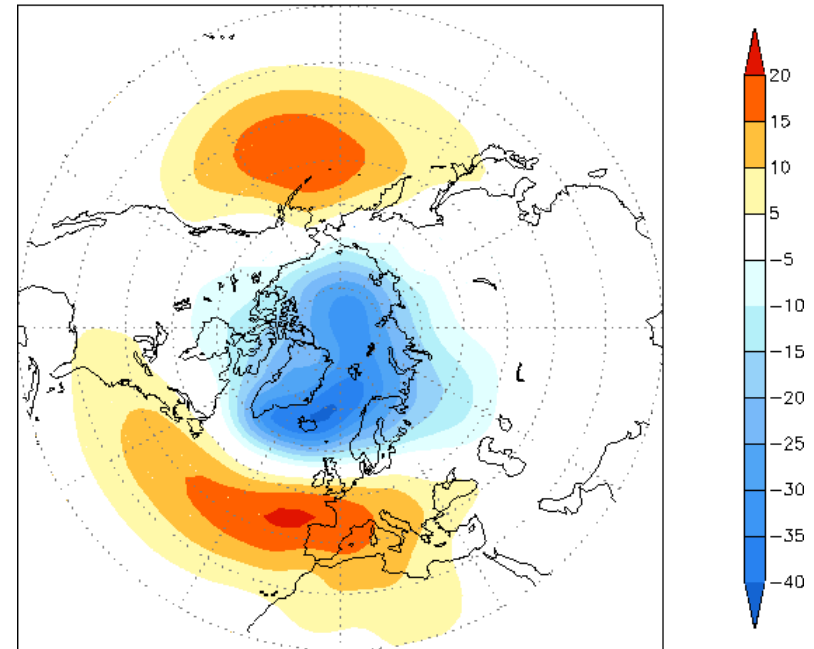
Calculated as first EOF of monthly mean MSLP anomalies, poleward of 20N.

Use same method as CPC, but using ERA interim analysis, 1981-2010.

Model and analysis time-series both obtained by projection onto **observed** EOF.

Correlation of our observed time-series with CPC is 0.996.

Leading EOF (19%) shown as regression map of 1000mb height (m)

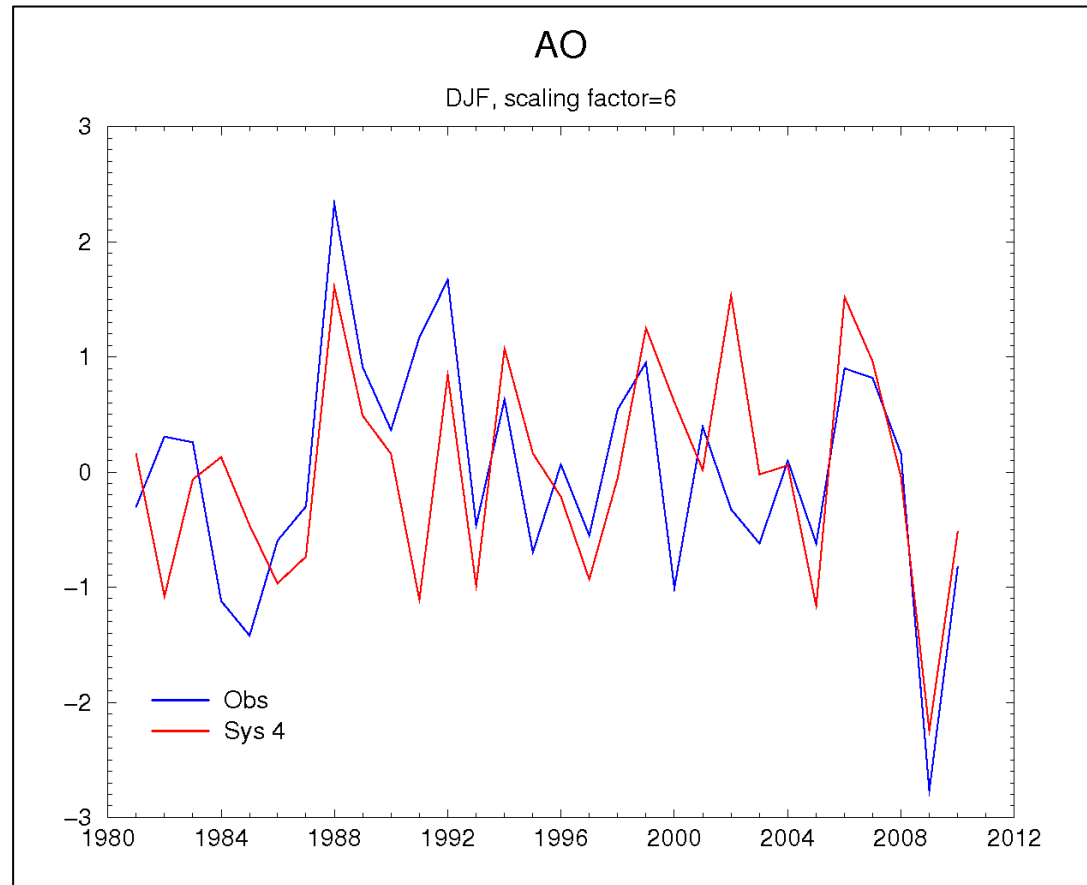


EOF (from CPC)

AO re-forecast skill

Correlation (30y) = 0.608

**26 years (no volcanoes)
Correlation = 0.73**



Surprising because model AO is very noisy

Statistical analysis

Unbiased variance estimates: Obs/Tot/Int/Ext: 1.0000 0.8390 0.8316 0.0074

Model/obs stddev ratio: 0.9159

Model/obs stddev ratio interval: 0.693 1.129

← model variability consistent with obs

Bootstrap over nens, pval for ratio=1: 0.7960

```
=====
SNR actual                                   :    0.0941
SNR jackknife over nens                   :    0.0202    0.1029    0.1857
=====
```

```
=====
ACC actual                                   :    0.6085
ACC basic bootstrap over nens             :    0.5568    0.7121    0.8144
ACC basic bootstrap over nyears:         :    0.2052    0.6069    0.8326
=====
```

← 95% interval due to ensemble size

← bigger uncertainty range here

ACP from internal sampling: -0.2947 0.0583 0.4010

Mean ACC for nens-1: 0.6049

p val of measured acc if model perfect: 0.9996

← only a 0.0004 chance we could get this correlation

- **Model skill for these years is relatively high**
- **Model predictability limit must be wrong (because we exceed it so much)**

Other teleconnection patterns

	ACC	S/N	ACP	P-val
PNA (EOF)	0.696	0.64	0.54	0.065
NAO (EOF)	0.465	0.13	0.10	0.017

PNA has high skill and high predictability

NAO has moderate skill, and low predictability

NAO skill is, like AO, higher than expected

Does resolution help?

Project Minerva has run the ECMWF coupled model at different atmospheric resolutions. We have 30 years of winter forecasts, with 51 member ensembles:

	T319		T639	
	ACC	S/N	ACC	S/N
PNA (EOF)	0.68	0.69	0.69	0.73
NAO (EOF)	0.36	0.17	0.63	0.18

S/N does not seem to be affected by resolution.

NAO structure and skill is significantly (at 5% level) improved by higher atmosphere resolution.

Where does model signal come from?

● Not obvious in initial conditions

- Can see traces of La Nina, not much sign of snow ics or QBO
- 30 hPa winds at 60N seem to have some correspondence

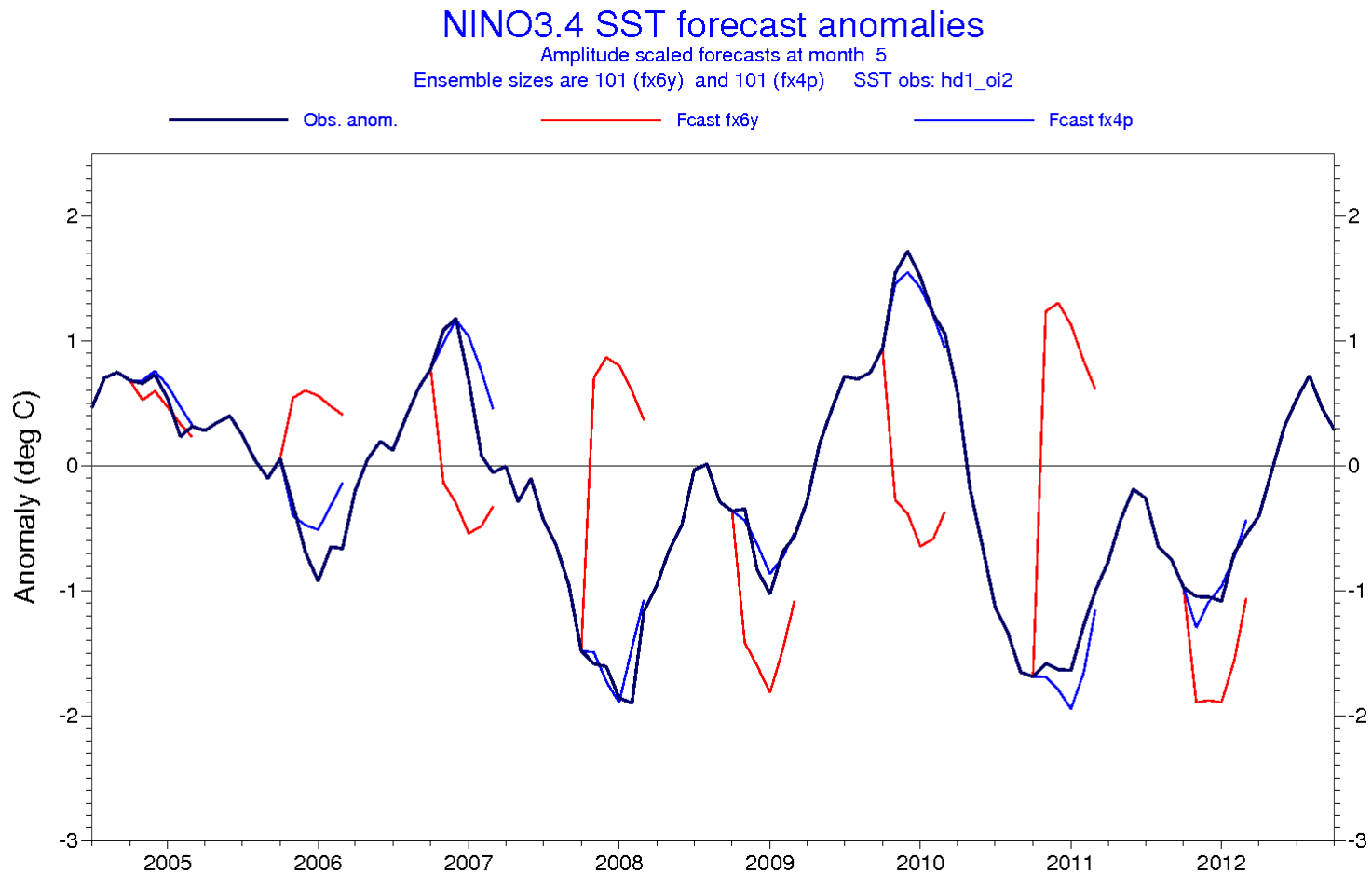
● Experiment – separate surface and atmos

- CONTROL: Atmos, land, sea-ice, ocean ics all from same year
- SHIFT: Atmos initial conditions from one year; ocean, sea-ice and land surface values from preceding year
- Six years with strong signal, **201** member ensembles for each expt.

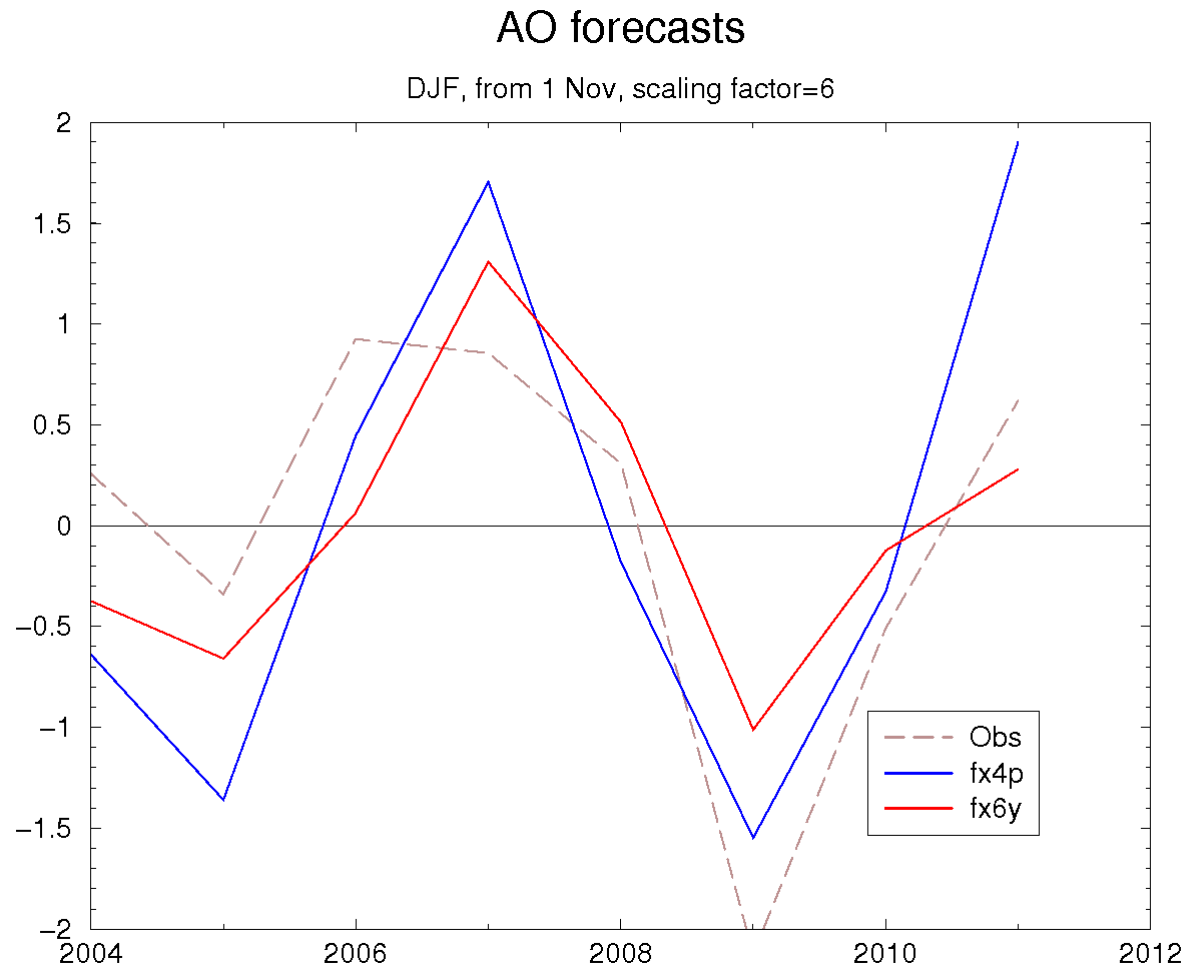
Does the model AO signal follow the SST forcing (plus sea-ice, snow cover etc) ...

.... or the free atmosphere initial conditions?

ENSO: ocean ic's dominate

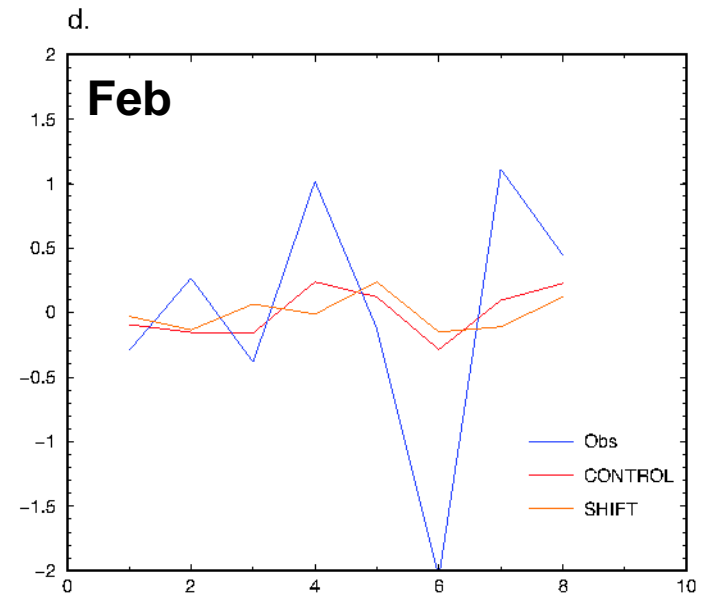
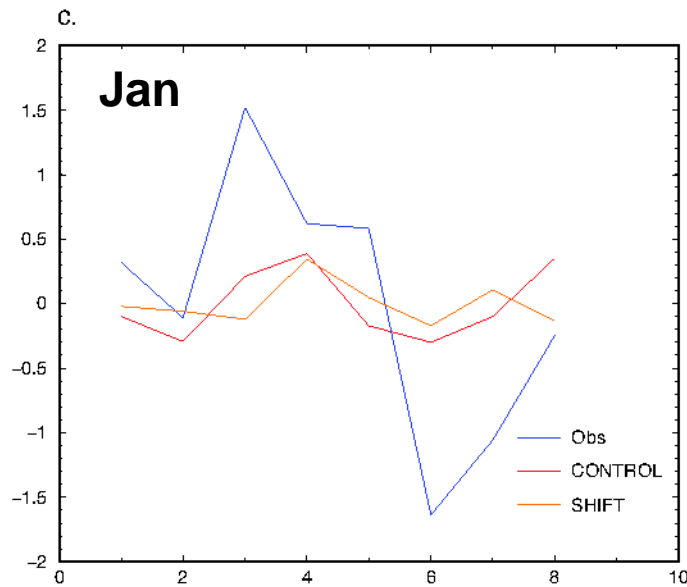
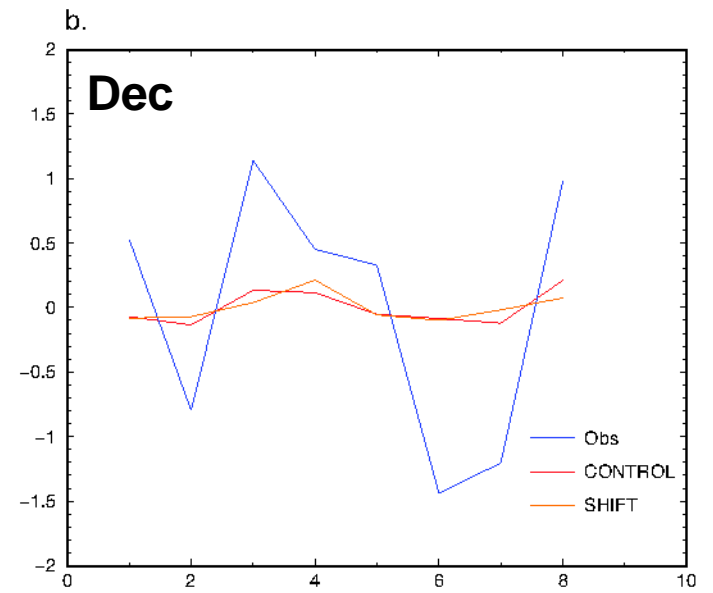
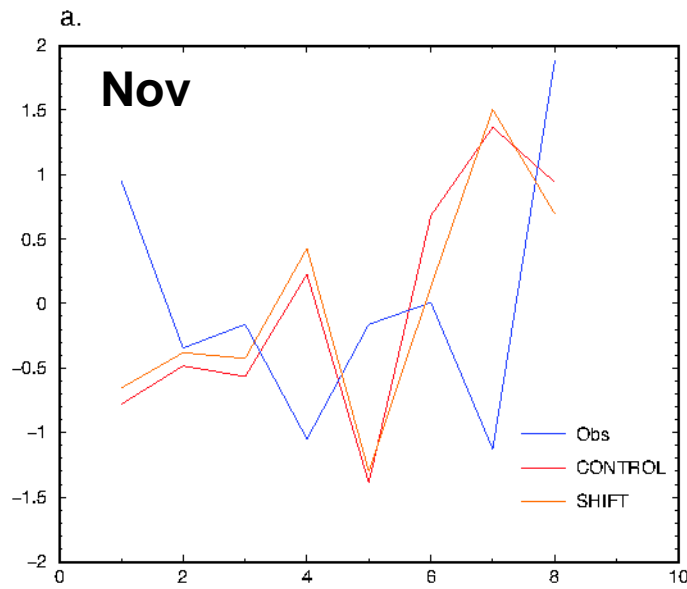


AO: atmospheric's dominate

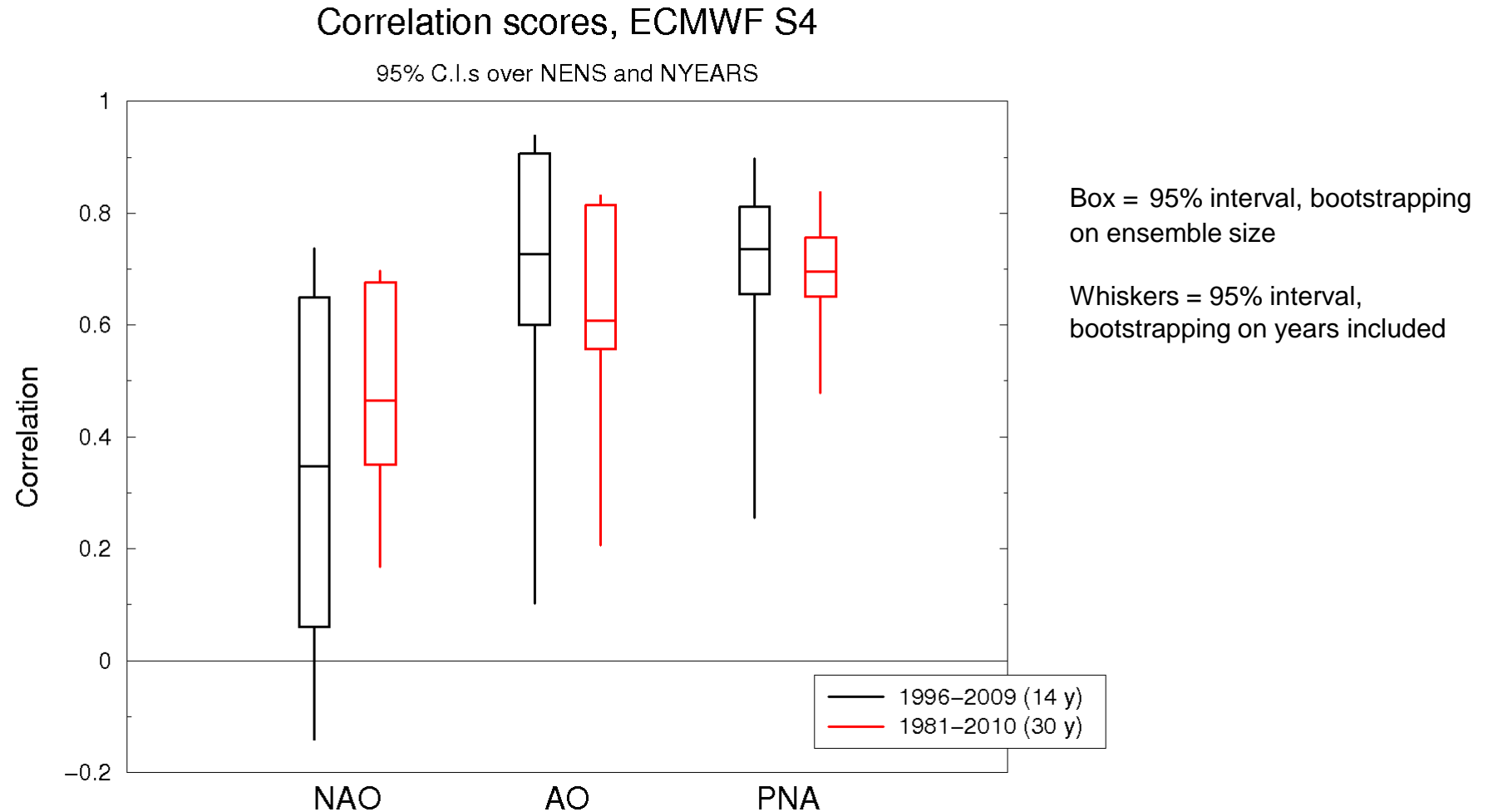


CONTROL
SHIFT

Month by month evolution

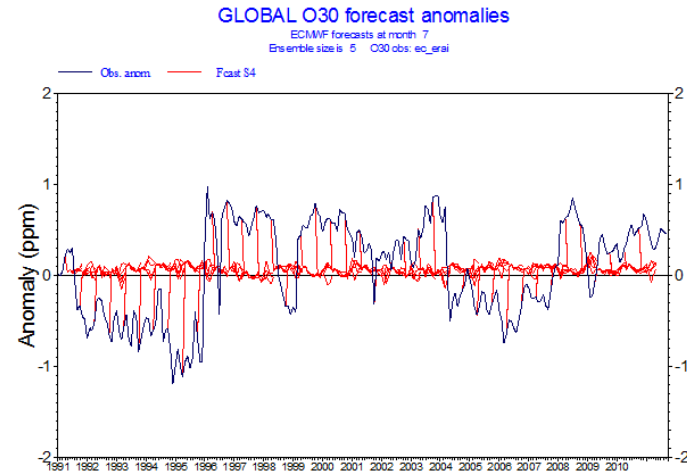


BUT: sampling errors are large!



BUT: real-time use uncertain!

- Homogeneity of re-forecasts and real-time forecasts is critical for long-range forecast systems
- Stratosphere analyses are a particular challenge
- Ensemble mean signal is small, needs scaling up: any inconsistencies will be scaled up, too
- Advise caution until mechanisms better understood



Conclusions

- **S4 has substantial skill in predicting AO phase over a 30 year period**
 - How typical this is of expected future performance is unknown
 - Amplitude of model signal is too weak
- **The real AO is more predictable than our model**
 - How much more is not known
 - Implies models can be improved
- **Importance of atmospheric initial conditions**
 - Dominate AO for recent high-signal years
 - Surface influence stronger later in season

Promising results, but still many unknowns