Predictability of the Arctic Oscillation

Tim Stockdale, Franco Molteni and Laura Ferranti

ECMWF
Outline

- Intro: ECMWF long-range forecasting
- Predicting the Arctic Oscillation
- Atmospheric initial conditions
- Conclusions and questions
Seasonal prediction at ECMWF

- Started in the 1990’s
- Strategy: fully coupled global GCMs
- Real-time forecasts since early 1997
  - Forecasts issued publicly from December 1997
- Now using “System 4”
  - Lifetime of systems has been about 5 years each

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
</table>
System 4 seasonal forecast model

- **IFS (atmosphere)**
  - T$_L^{255L91}$ Cy36r4, 80km grid (operational in Dec 2010)
  - Full stratosphere, enhanced stratospheric physics
  - ERA interim / operational atmosphere initial conditions

- **NEMO (ocean)**
  - Global ocean model, 1x1 resolution, 0.3 meridional near equator
  - NEMOVAR (3D-Var) analyses.

- **Sea ice**
  - No physical model of sea-ice
  - Specified values sampled from previous five years
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 ….

**NINO3.4 SST rms errors**
- 180 start dates from 19810101 to 19951201, amplitude scaled
- Ensemble size is 15
- 95% confidence interval for 0001, for given set of start dates

**NINO3.4 SST anomaly correlation**
- wrt NCEP adjusted Oi2 1971-2000 climatology

---

**1981-1995**

**1996-2010**
So are deterministic scores in the tropics ....
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

System 3

System 4

Anomaly (m/s)
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.
AO re-forecast skill

Correlation (30y) = 0.608

26 years (no volcanoes)
Correlation = 0.73

Surprising because model AO is very noisy ....
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 ➤ 95% interval due to ensemble size
ACC basic bootstrap over nyears: 0.2052 0.6069 0.8326 ➤ 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 rather high
- Model predictability limit must be wrong (because we exceed it so much)
### Other teleconnection patterns

<table>
<thead>
<tr>
<th></th>
<th>ACC</th>
<th>S/N</th>
<th>ACP</th>
<th>P-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNA (EOF)</td>
<td>0.696</td>
<td>0.64</td>
<td>0.54</td>
<td>0.065</td>
</tr>
<tr>
<td>NAO (EOF)</td>
<td>0.465</td>
<td>0.13</td>
<td>0.10</td>
<td>0.017</td>
</tr>
</tbody>
</table>

- **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:

<table>
<thead>
<tr>
<th></th>
<th>T319</th>
<th>T639</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S/N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNA (EOF)</td>
<td>0.68</td>
<td>0.69</td>
</tr>
<tr>
<td>NAO (EOF)</td>
<td>0.36</td>
<td>0.17</td>
</tr>
</tbody>
</table>

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

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

Dec and Feb AO forecasts

fx4p, 201 members

-2 -1.5 -1 -0.5 0 0.5 1 1.5 2
0 2 4 6 8 10
2006/07 2010/11

Dec obs
Dec fx4p
Feb
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?
AO forecasts

DJF, from 1 Nov, scaling factor=6
True for AO, not generally!
**Time/height evolution**

<table>
<thead>
<tr>
<th>Period/level</th>
<th>MSLP AO</th>
<th>Z50 AO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACC</td>
<td>S/N</td>
</tr>
<tr>
<td>November</td>
<td>-0.04</td>
<td>0.66</td>
</tr>
<tr>
<td>December</td>
<td>0.79</td>
<td>0.20</td>
</tr>
<tr>
<td>January</td>
<td>0.50</td>
<td>0.23</td>
</tr>
<tr>
<td>February</td>
<td>0.78</td>
<td>0.19</td>
</tr>
<tr>
<td>DJF</td>
<td>0.75</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Based on 2006-2011, with 201 member ensemble
Conclusions (1)

● S4 has substantial skill in predicting AO phase, over a 30 year period
  ○ How typical this is of expected future performance is unknown

● The real AO is more predictable than our model
  ○ How much more is not known

● Our model has sub-seasonal resolution
  ○ Some of which is skilful; limits of predictability are not known
Conclusions (2)

- Model AO signal is dominated by atmospheric initial conditions on 1 November
  - True at least for recent high-signal years
  - Surface influence stronger later in season (e.g. by Feb)
  - Not ruled out that atmospheric signal on 1st November may have come from surface during e.g. October (SST, snow cover)

- Model signal appears to start in stratosphere

- Amplitude of model signal at surface is too weak
  - Maybe even by a factor of five
Discussion

- Role of atmospheric initial conditions a surprise
  - Applies to the AO, not winter circulation in general
  - Presumably due to the AO being a linked stratosphere/troposphere mode
  - But AO is the leading mode of variability, so important

  - Would be valuable to examine this in other models, other start dates

- Why is the model signal weak?
  - Evidence *suggests* mechanism of downward stratosphere to troposphere coupling for AO is too weak
  - Literature suggests mechanisms not well understood
    (Song and Robinson, 2004; Gerber et al, BAMS, 2012)
  - Model issues include: vertical diffusion, non-orographic GWD, numerics
NH winter forecasts: vertical diffusion

T850 Anom. correlation f(u)(101)-ERA-Int 1981-2010DJF
Global z-mean acc: 0.819 NH: 0.881 TR: 0.662 SH: 0.308

Z500 Anom. correlation f(u)(101)-ERA-Int 1981-2010DJF
Global z-mean acc: 0.666 NH: 0.643 TR: 0.829 SH: 0.403

0.319

Fisher z transform diff f(u)(101)-f(u)(101) 1981-2010DJF
sigma: 0.272 mean: 0.0035

Z500 Anom. correlation f(u)(101)-ERA-Int 1981-2010DJF
Global z-mean acc: 0.664 NH: 0.519 TR: 0.852 SH: 0.130

0.371
Model/observed variability

Ratio of SD (model/reference) for ECMWF S4 with 51 ensemble members
500 hPa geopotential height
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)

Ensemble spread / r.m.s. error

Ratio Spread(sd)/RMSE for ECMWF S4 with 51 ensemble members
500 hPa geopotential height
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)

NH stddev ratio: 1.064
p val for observed stddev: 0.0785
NH stddev ratio 95% interval: 0.979 - 1.149