Seasonal prediction over Europe

F. J. Doblas-Reyes, ICREA & IC3, Barcelona, Spain

With substantial contributions from L. Rodrigues, J. García (IC3), A. Weisheimer, T.N. Palmer and the Seasonal Forecast Section (ECMWF), G.J. van Oldenborg, B. van den Hurk, Climate Explorer (KNMI)
Contents

- Seasonal forecasting: empirical and dynamical
- Operational seasonal forecast system: ECMWF System 3
- Systematic errors
- Skill: correlation
- Sources of predictability and error reduction
- Conclusions

- Focus on one-month lead seasonal forecasts of temperature and precipitation, except for JFM
Sources of seasonal predictability

• Important:
  o ENSO - biggest single signal
  o Other tropical ocean SST - difficult
  o Climate change - important in mid-latitudes
  o Local land surface conditions - soil moisture, snow
  o Atmospheric composition - difficult

• Other factors:
  o Volcanic eruptions - important for large events
  o Mid-latitude ocean temperatures - still somewhat controversial
  o Remote soil moisture/snow cover - not well established
  o Sea-ice anomalies - at least local effects
  o Stratospheric influences - various possibilities
  o Remote tropical atmospheric teleconnections

• Unknown or Unexpected
Methods of seasonal forecasting

• **Empirical forecasting**
  - Use past observational record and statistical methods
  - Works with reality instead of error-prone numerical models ✔
  - Limited number of past cases ✗
  - A non-stationary climate is problematic ✗
  - Can be used as a benchmark ✔

• **Two-tier forecast systems**
  - First predict SST anomalies (ENSO or global; dynamical or statistical)
  - Use ensemble of atmosphere GCMs to predict global response
  - Systematic model error is an issue

• **Single-tier GCM forecasts**
  - Include comprehensive range of sources of predictability ✔
  - Predict joint evolution of ocean and atmosphere flow ✔
  - Includes a large range of physical processes ✔
  - Includes uncertainty sources, important for prob. Forecasts ✔
  - Systematic model error is an issue! ✗
Simple empirical model: persistence


Northern Europe (40-75°N, 10°W-40°E)

Southern Europe (35-45°N, 5°W-30°E)
Temperature skill: persistence

Correlation of GHCN temperature of one-month lead anomaly persistence over 1981-2005. Only values statistically significant with 80% confidence are plotted.
To produce dynamical forecasts

- Build a coupled model
- Prepare initial conditions
- Initialize coupled system
  - The aim is to start the system close to reality. Accurate SST is particularly important, plus ocean sub-surface. Usually, worry about “imbalances” a posteriori.
- Run an ensemble forecast
  - Explicitly generate an ensemble on the e.g. 1st of each month, with perturbations to represent the uncertainty in the initial conditions; run forecasts for several months.
- Produce probability forecasts from the ensemble
- Apply calibration and combination if significant improvement is found
ECMWF’s System 3 seasonal forecasts

- **Real-time forecasts:**
  - 41-member ensemble forecast to 7 months
  - Five-member ocean analysis
  - SST and atmos. perturbations added to each member
  - Initial conditions are valid for 0 GMT on the 1st of a month

  - 11 member ensemble forecast to 13 months
  - Designed to give an ‘outlook’ for ENSO
  - Only once per quarter (Feb, May, Aug and Nov starts)

- **Re-forecasts from 1981-2005 (25 years)**
  - 11-member ensemble every month
  - 5 members to 13 months once per quarter
  - The observations have only 1 member, so large ensembles are much less helpful than large numbers of cases.
Main systematic errors in dynamical climate forecasts:

- Differences between the model climatological pdf (computed for a lead time from all start dates and ensemble members) and the reference climatological pdf (for the corresponding times of the reference dataset): systematic errors in mean and variability.

- Conditional biases in the forecast pdf: errors in conditional probabilities implying that probability forecasts are not trustworthy. This type of systematic error is best assessed using the reliability diagram.

### Differences in climatological pdfs

<table>
<thead>
<tr>
<th>Reference pdf</th>
<th>Model pdf</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean bias</strong></td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td></td>
</tr>
</tbody>
</table>

### Different variabilities

- **T=1**
- **T=2**
- **T=3**

### Actual occurrences

- Threshold

- t=1
- t=2
- t=3

---

ECMWF Seminar – Seasonal prediction over Europe

Reading, 7 September 2010
Systematic error: climatological pdf

Climatological PDF of DJF T2m (°C) for ERA-40/OPS and ECMWF System 3, 1st of November start date computed over 1981-2005

For deterministic forecasts, compute anomalies with respect to the corresponding mean.

For probabilistic forecasts, compute probabilities with respect to hindcast and reference thresholds (terciles).

Central Europe (50°N, 10°E)

Equatorial Pacific (0, 180°)
Mean error: SSTs

System 3 one-month lead SST (K) bias wrt HadISST1 over 1981-2005.

Typical of CMIP3 models

DJF

MAM

JJA

SON

INSTITUT CATALÀ DE CIÈNCIES DEL CLIMA
Mean error: MSLP

System 3 one-month lead mean sea level pressure (hPa) bias wrt NCEP/NCAR R1 over 1981-2005.

DJF

MAM

JJA

SON
Mean error: T2m

System 3 one-month lead near-surface air temperature (K) bias wrt GHCN over 1981-2005.
Mean error: precipitation

System 3 one-month lead precipitation (mm/day) bias wrt GPCC over 1981-2005.
Seasonal re-forecasts for Europe

System 3 temperature re-forecasts for Southern (left) and Northern Europe (right) over 1981-2005. The green box-and-whisker show the ensemble range, the blue dot the ensemble mean and the red dot the ERA40/ERAInt value.

- **JJA**
  - Ratio sd: 1.35
  - Corr: 0.62

- **DJF**
  - Ratio sd: 0.68
  - Corr: 0.24

- **JJA**
  - Ratio sd: 1.06
  - Corr: -0.12

- **DJF**
  - Ratio sd: 0.46
  - Corr: 0.14
Seasonal re-forecasts for Europe

System 3 precipitation re-forecasts for Southern (left) and Northern Europe (right) over 1981-2005. The green box-and-whisker show the ensemble range, the blue dot the ensemble mean and the red dot the GPCP value.
Constructing a probability forecast from a nine-member ensemble

A threshold relative to the model climate (e.g. percentile)
From ensembles to probability forecasts

Constructing a probability forecast from a nine-member ensemble

A threshold relative to the model climate (e.g. percentile)
Constructing a probability forecast from a multi-model ensemble
Probabilistic prediction

One-month lead DJF 2009-10 System 3 seasonal forecasts: tercile summary
References: what actually happened


T2m (K)

Prec (mm/day)

MSLP (hPa)
Probabilistic prediction

One-month lead DJF 2009-10 IRI (flexible format) temperature forecasts for anom. above the upper tercile
December start date forecasts: JFM

JFM 2010 mean sea level pressure seasonal anomalies for (left) NCEP/NCAR R1 (hPa) and (right) tercile summary for the one-month lead System 3 forecasts wrt 1981-2005.
Temperature skill: System 3

Correlation of System 3 seasonal forecasts of temperature wrt GHCN over 1981-2005. Only values statistically significant with 80% confidence are plotted.
Temperature skill: Potential predictability

Precipitation skill: System 3

Correlation of System 3 seasonal forecasts of precipitation wrt GPCC over 1981-2005. Only values statistically significant with 80% confidence are plotted.
Precipitation skill: System 3

Sources of predictability and error

- ENSO and tropical Atlantic
- Extratropical SSTs
- Trends and anthropogenic warming
- Model inadequacy
- Model improvement
- Soil moisture
- Snow
- Stratospheric processes
- Volcanic aerosol
Global skill: System 3

Correlation of System 3 seasonal forecasts of temperature (top) and precipitation (bottom) wrt GHCN and GPCC over 1981-2005. Only values significant with 80% conf. plotted.
ENSO teleconnections: observations

Regression of NCEP/NCAR R1 mean sea level pressure on HadISST1 Niño3.4 time series (hPa/K) over 1981-2005.
Regression of System 3 seasonal temperature on predicted Niño3.4 SST time series (hPa/K) over 1981-2005.
Regression of GHCN seasonal temperature on HadISST1 Niño3.4 time series (K/K) over 1981-2005.
Regression of System 3 seasonal temperature on predicted Niño3.4 SST time series (K/K) over 1981-2005.
Tropical SSTs: the Atlantic case

Leading two pairs of the MCA between NCEP/NCAR R1 DJF mean sea level pressure (left) and the leading October North Atlantic HadISST1 SSTs (right) over 1981-2005. The two pairs explain 31% and 16% of the covariance.
Tropical SSTs: the Atlantic case

Leading two pairs of the MCA between System 3 one-month lead DJF MSLP (left) and the leading October North Atlantic HadISST1 SSTs (right) over 1981-2005. The two pairs explain 24% and 17% of the covariance.
Extratropical SSTs: the North Atlantic

Leading pair of the MCA between NCEP/NCAR R1 DJF mean sea level pressure (left) and the leading May North Atlantic HadISST1 SSTs (right) over 1981-2005 (after Rodwell and Folland, 2002). The pair explains 23% of the covariance.
Extratropical SSTs: the North Atlantic

Leading pairs of the MCA between System 3 one-month lead DJF MSLP (left) and the leading May North Atlantic HadISST1 SSTs (right) over 1981-2005. The pairs explain 14% and 12% of the covariance, respectively.
Impact of the trend: temperature

Correlation of System 3 seasonal forecasts of detrended temperature wrt GHCN over 1981-2005. Only values statistically significant with 80% confidence are plotted.
Temperature: global trend

Global-mean near-surface air temperature for System 3. The green box-and-whisker plots show the ensemble range, the blue dot the ensemble mean and the red dot the ERA40/ERAInt value.

**JJA**
Ratio sd: 0.59  
Corr: 0.69

**DJF**
Ratio sd: 1.26  
Corr: 0.44
Temperature: impact of the trend

Regression of CRUTEM3/HadSST2 temperature with the HadCRUT3 global-mean temperature over 1981-2005.
Temperature: impact of the trend

Regression of System 3 temperature seasonal forecasts on the global-mean temperature over 1981-2005. Only values statistically significant with 80% confidence are plotted.
Model inadequacy: ENSEMBLES

- Common experimental set-up with realistic initialization using **multi-model, perturbed parameters and stochastic physics**
  - Seasonal and annual (14 months) hindcasts
  - Two streams
    - Stream 1: 1991-2001, May and Nov start dates, nine members
    - Stream 2: 1960-2005, 4 start dates per year for seasonal/annual (9 members)
Reliability diagram

Forecast reliability measures conditional biases of forecast probabilities; example for event “values above the median”

Overconfident

Perfect reliability
Multi-model improvement

Attribute diagrams for one-month lead seasonal (DJF) temperature over Southern Europe for System 3 (left, 9 members) and the ENSEMBLES Stream 2 multi-model (right, 45 members) over the period 1981-2005 verified against ERA40. Brier and ROC skill scores and 95% confidence intervals (in brackets) computed using a bootstrap method, are shown on top of each panel.

**System 3**

<table>
<thead>
<tr>
<th>Above upper tercile</th>
<th>Below lower tercile</th>
</tr>
</thead>
<tbody>
<tr>
<td>-.078 (-.287,.093)</td>
<td>-.140 (-.348,.022)</td>
</tr>
<tr>
<td>.148 (-.073,.379)</td>
<td>.109 (-.095,.325)</td>
</tr>
</tbody>
</table>

**Multi-model**

<table>
<thead>
<tr>
<th>Above upper tercile</th>
<th>Below lower tercile</th>
</tr>
</thead>
<tbody>
<tr>
<td>.054 (-.063,.140)</td>
<td>.056 (-.101,.093)</td>
</tr>
<tr>
<td>.256 (-.004,.464)</td>
<td>.326 (-.119,.516)</td>
</tr>
</tbody>
</table>
Attribute diagrams for one-month lead seasonal (JJA) precipitation over Southern Europe for System 3 (left, 9 members) and the ENSEMBLES Stream 2 multi-model (right, 45 members) over the period 1981-2005 verified against GPCP. Brier and ROC skill scores and 95% confidence intervals (in brackets) computed using a bootstrap method, are shown on top of each panel.

**System 3**
- Above upper tercile: -.088 (-.167, -.038)
- Below lower tercile: -.118 (-.178, -.071)
- Above upper tercile: .075 (-.027, .175)
- Below lower tercile: .081 (-.018, .140)

**Multi-model**
- Above upper tercile: .001 (-.069, .046)
- Below lower tercile: .133 (-.022, .274)
- Above upper tercile: .031 (-.082, .002)
- Below lower tercile: .051 (-.070, .174)
Commonality of systematic errors

Northern Hemisphere DJF blocking frequency for 1-month lead seasonal hindcasts for the ECMWF, Météofrance, Met Office, IfM and INGV forecast systems over the period 1981-2005 compared to ERA40/ERAInt. Results are for the Tibaldi and Molteni index (reversal of the meridional gradient of Z500). 95% intervals are shown.

ECMWF Seminar – Seasonal prediction over Europe
Reading, 7 September 2010
Dealing with model inadequacy

Debiased Brier skill score of one-month lead predictions of land temperature over the Giorgi regions for Multi-model (45 members, left columns), Perturbed parameters (9 members, central columns) and Stochastic physics (9 members, right columns) over 1991-2005. Significantly positive or negative scores are in bold.

<table>
<thead>
<tr>
<th>Region</th>
<th>JJA cold</th>
<th>JJA warm</th>
<th>DJF cold</th>
<th>DJF warm</th>
<th>JJA cold</th>
<th>JJA warm</th>
<th>DJF cold</th>
<th>DJF warm</th>
<th>JJA cold</th>
<th>JJA warm</th>
<th>DJF cold</th>
<th>DJF warm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>11.5</td>
<td>13.9</td>
<td>3.2</td>
<td>6.7</td>
<td>-0.3</td>
<td>11.0</td>
<td>0.5</td>
<td>5.2</td>
<td>7.0</td>
<td>17.3</td>
<td>11.8</td>
<td>8.0</td>
</tr>
<tr>
<td>Amazon Basin</td>
<td>0.2</td>
<td>17.1</td>
<td>4.5</td>
<td>23.4</td>
<td>-13.7</td>
<td>28.6</td>
<td>-6.3</td>
<td>11.2</td>
<td>39.1</td>
<td>14.7</td>
<td>26.6</td>
<td>16.9</td>
</tr>
<tr>
<td>Southern South America</td>
<td>9.2</td>
<td>9.0</td>
<td>1.8</td>
<td>9.9</td>
<td>2.8</td>
<td>7.2</td>
<td>2.9</td>
<td>14.7</td>
<td>16.9</td>
<td>8.8</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>Central America</td>
<td>5.9</td>
<td>11.6</td>
<td>-2.6</td>
<td>4.5</td>
<td>2.4</td>
<td>5.5</td>
<td>-3.9</td>
<td>3.3</td>
<td>12.1</td>
<td>-0.3</td>
<td>0.2</td>
<td>-3.7</td>
</tr>
<tr>
<td>Western North America</td>
<td>10.2</td>
<td>12.2</td>
<td>6.3</td>
<td>12.5</td>
<td>6.7</td>
<td>-1.2</td>
<td>3.3</td>
<td>8.9</td>
<td>28.1</td>
<td>8.0</td>
<td>6.4</td>
<td>4.7</td>
</tr>
<tr>
<td>Central North America</td>
<td>-0.2</td>
<td>-7.3</td>
<td>-3.3</td>
<td>10.4</td>
<td>-8.5</td>
<td>-12.7</td>
<td>7.2</td>
<td>13.8</td>
<td>-21.4</td>
<td>-20.3</td>
<td>-2.6</td>
<td>8.8</td>
</tr>
<tr>
<td>Eastern North America</td>
<td>4.1</td>
<td>-7.0</td>
<td>-4.5</td>
<td>10.1</td>
<td>-9.9</td>
<td>-14.7</td>
<td>32.2</td>
<td>8.2</td>
<td>-13.4</td>
<td>-10.9</td>
<td>-11.3</td>
<td>4.0</td>
</tr>
<tr>
<td>Alaska</td>
<td>-0.8</td>
<td>-0.9</td>
<td>-0.6</td>
<td>0.6</td>
<td>-0.4</td>
<td>-2.9</td>
<td>6.5</td>
<td>4.8</td>
<td>0.5</td>
<td>12.2</td>
<td>-20.3</td>
<td>-1.0</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>18.0</td>
<td>12.8</td>
<td>5.8</td>
<td>4.3</td>
<td>18.3</td>
<td>15.5</td>
<td>-17.5</td>
<td>-14.5</td>
<td>22.7</td>
<td>12.2</td>
<td>6.2</td>
<td>2.6</td>
</tr>
<tr>
<td>Northern Europe</td>
<td>-3.3</td>
<td>0.2</td>
<td>4.9</td>
<td>0.5</td>
<td>1.1</td>
<td>4.6</td>
<td>-0.6</td>
<td>-4.0</td>
<td>4.6</td>
<td>6.3</td>
<td>1.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Western Africa</td>
<td>7.8</td>
<td>7.9</td>
<td>10.5</td>
<td>20.5</td>
<td>11.8</td>
<td>9.5</td>
<td>9.6</td>
<td>14.8</td>
<td>78.6</td>
<td>-2.2</td>
<td>14.0</td>
<td>15.6</td>
</tr>
<tr>
<td>Eastern Africa</td>
<td>9.4</td>
<td>7.3</td>
<td>-7.7</td>
<td>0.9</td>
<td>-19.5</td>
<td>-7.1</td>
<td>-3.9</td>
<td>-5.4</td>
<td>-9.7</td>
<td>-3.1</td>
<td>-3.7</td>
<td>8.2</td>
</tr>
<tr>
<td>Southern Africa</td>
<td>14.0</td>
<td>4.7</td>
<td>1.7</td>
<td>10.6</td>
<td>3.2</td>
<td>10.2</td>
<td>-1.7</td>
<td>2.7</td>
<td>0.0</td>
<td>7.7</td>
<td>6.0</td>
<td>13.6</td>
</tr>
<tr>
<td>Sahel</td>
<td>12.9</td>
<td>7.2</td>
<td>11.5</td>
<td>15.4</td>
<td>9.9</td>
<td>13.1</td>
<td>6.6</td>
<td>15.7</td>
<td>16.3</td>
<td>10.1</td>
<td>13.9</td>
<td>14.7</td>
</tr>
<tr>
<td>South East Asia</td>
<td>8.6</td>
<td>12.4</td>
<td>11.6</td>
<td>13.4</td>
<td>9.3</td>
<td>4.2</td>
<td>13.9</td>
<td>6.1</td>
<td>9.6</td>
<td>9.6</td>
<td>3.8</td>
<td>1.6</td>
</tr>
<tr>
<td>East Asia</td>
<td>10.6</td>
<td>10.2</td>
<td>0.3</td>
<td>5.8</td>
<td>8.3</td>
<td>10.5</td>
<td>-4.2</td>
<td>10.1</td>
<td>6.4</td>
<td>14.1</td>
<td>3.1</td>
<td>0.4</td>
</tr>
<tr>
<td>South Asia</td>
<td>8.7</td>
<td>13.3</td>
<td>14.4</td>
<td>10.6</td>
<td>4.3</td>
<td>9.2</td>
<td>0.1</td>
<td>9.3</td>
<td>129.1</td>
<td>15.7</td>
<td>13.6</td>
<td>18.1</td>
</tr>
<tr>
<td>Central Asia</td>
<td>14.3</td>
<td>8.2</td>
<td>-2.4</td>
<td>7.1</td>
<td>14.1</td>
<td>11.8</td>
<td>-2.0</td>
<td>19.1</td>
<td>21.1</td>
<td>10.1</td>
<td>-8.5</td>
<td>6.5</td>
</tr>
<tr>
<td>Tibet</td>
<td>16.9</td>
<td>16.1</td>
<td>-0.1</td>
<td>4.1</td>
<td>7.8</td>
<td>7.2</td>
<td>-10.4</td>
<td>3.8</td>
<td>83.3</td>
<td>15.7</td>
<td>56.5</td>
<td>7.6</td>
</tr>
<tr>
<td>North Asia</td>
<td>7.3</td>
<td>3.9</td>
<td>4.2</td>
<td>8.5</td>
<td>6.2</td>
<td>8.4</td>
<td>-1.5</td>
<td>12.6</td>
<td>4.2</td>
<td>1.6</td>
<td>-1.9</td>
<td>1.2</td>
</tr>
</tbody>
</table>

**multi-model** | **perturbed parameters** | **stochastic physics**
Dealing with model inadequacy

Debiased Brier skill score of one-month lead predictions of land temperature over the Giorgi regions for Multi-model (45 members, left columns), Perturbed parameters (9 members, central columns) and Stochastic physics (9 members, right columns) over 1991-2005. Significantly positive or negative scores are in bold.

<table>
<thead>
<tr>
<th>Region</th>
<th>JJA</th>
<th>DJF</th>
<th>JJA</th>
<th>DJF</th>
<th>JJA</th>
<th>DJF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dry</td>
<td>wet</td>
<td>dry</td>
<td>wet</td>
<td>dry</td>
<td>wet</td>
</tr>
<tr>
<td>Australia</td>
<td>7.6</td>
<td>7.0</td>
<td>0.9</td>
<td>3.0</td>
<td>5.1</td>
<td>8.0</td>
</tr>
<tr>
<td>Amazon Basin</td>
<td>10.3</td>
<td>10.3</td>
<td>16.0</td>
<td>14.3</td>
<td>8.8</td>
<td>5.4</td>
</tr>
<tr>
<td>Southern South America</td>
<td>6.2</td>
<td>7.1</td>
<td>4.6</td>
<td>6.0</td>
<td>13.1</td>
<td>1.6</td>
</tr>
<tr>
<td>Central America</td>
<td>9.2</td>
<td>7.8</td>
<td>23.4</td>
<td>18.9</td>
<td>12.9</td>
<td>5.2</td>
</tr>
<tr>
<td>Western North America</td>
<td>2.4</td>
<td>8.1</td>
<td>7.2</td>
<td>7.8</td>
<td>4.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Central North America</td>
<td>0.6</td>
<td>22</td>
<td>7.7</td>
<td>10.4</td>
<td>3.5</td>
<td>-5.7</td>
</tr>
<tr>
<td>Eastern North America</td>
<td>-1.9</td>
<td>-1.3</td>
<td>8.3</td>
<td>10.6</td>
<td>-9.6</td>
<td>-1.1</td>
</tr>
<tr>
<td>Alaska</td>
<td>-1.3</td>
<td>0.0</td>
<td>4.0</td>
<td>-2.2</td>
<td>2.3</td>
<td>-1.0</td>
</tr>
<tr>
<td>Greenland</td>
<td>2.6</td>
<td>0.8</td>
<td>2.3</td>
<td>2.0</td>
<td>4.4</td>
<td>-2.3</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>-1.2</td>
<td>1.2</td>
<td>-1.0</td>
<td>-1.3</td>
<td>-6.1</td>
<td>-4.4</td>
</tr>
<tr>
<td>Northern Europe</td>
<td>2.3</td>
<td>2.1</td>
<td>-3.1</td>
<td>-4.7</td>
<td>7.7</td>
<td>11.5</td>
</tr>
<tr>
<td>Western Africa</td>
<td>4.6</td>
<td>8.1</td>
<td>8.6</td>
<td>1.6</td>
<td>12.2</td>
<td>3.9</td>
</tr>
<tr>
<td>Eastern Africa</td>
<td>-2.8</td>
<td>1.8</td>
<td>3.9</td>
<td>2.5</td>
<td>-7.0</td>
<td>-7.6</td>
</tr>
<tr>
<td>Southern Africa</td>
<td>3.5</td>
<td>1.0</td>
<td>5.7</td>
<td>9.5</td>
<td>7.2</td>
<td>4.7</td>
</tr>
<tr>
<td>Sahel</td>
<td>-4.6</td>
<td>-3.6</td>
<td>-3.2</td>
<td>-1.5</td>
<td>-9.2</td>
<td>-6.7</td>
</tr>
<tr>
<td>South East Asia</td>
<td>14.3</td>
<td>9.7</td>
<td>8.8</td>
<td>8.3</td>
<td>5.5</td>
<td>4.8</td>
</tr>
<tr>
<td>East Asia</td>
<td>0.5</td>
<td>-0.5</td>
<td>4.7</td>
<td>4.6</td>
<td>5.6</td>
<td>1.4</td>
</tr>
<tr>
<td>South Asia</td>
<td>0.2</td>
<td>0.9</td>
<td>6.5</td>
<td>7.4</td>
<td>0.6</td>
<td>-2.7</td>
</tr>
<tr>
<td>Central Asia</td>
<td>-0.8</td>
<td>0.2</td>
<td>7.4</td>
<td>5.7</td>
<td>0.8</td>
<td>-3.1</td>
</tr>
<tr>
<td>Tibet</td>
<td>5.5</td>
<td>3.5</td>
<td>6.5</td>
<td>5.4</td>
<td>-3.4</td>
<td>-1.9</td>
</tr>
<tr>
<td>North Asia</td>
<td>2.4</td>
<td>2.6</td>
<td>3.1</td>
<td>0.6</td>
<td>3.3</td>
<td>2.9</td>
</tr>
</tbody>
</table>

**PRECIPITATION**

Dealing with model inadequacy

Debiased Brier skill score of one-month lead predictions of land temperature over the Giorgi regions for Multi-model (45 members, left columns), Perturbed parameters (9 members, central columns) and Stochastic physics (9 members, right columns) over 1991-2005. Significantly positive or negative scores are in bold.
Model improvement


From Antje Weisheimer

2m temp
model
obs

precipitation

2m temperature over Southern Europe (land)

forecast
(50 ensemble members)
observed anomaly
analysis climate
model climate

INSTITUT CATALÀ DE CIÊNCIES DEL CLIMA
Model improvement

Seasonal prediction with improved ECMWF system (changes in radiation, soil scheme and convection) for summer 2003 with May start date. Anomalies wrt period 1991-2005.

From Antje Weisheimer
Modes of variability: NAO

Leading EOF of SLP over the region 20°-85°N, 90°W-90°E for NCEP/NCAR R1 (top row) and one-month lead System 3 re-forecasts (bottom row). Variance percentage in brackets.

DJF MAM JJA SON

45% 34% 33% 31%

36% 40% 25% 23%
Modes of variability: NAO

System 3 NAO predictions over 1981-2005. The green box-and-whisker show the ensemble range, the blue dot the ensemble mean and the red dot the ERA40/ERAInt value.

- DJF: Corr: 0.33
- MAM: Corr: 0.06
- JJA: Corr: 0.03
- SON: Corr: -0.01
NAO predictions

ECMWF System 2 and DEMETER NAO DJF forecasts (November start date, 1-month lead, 1987-2001).

RPSS$_D$ for tercile categories

Predicted NAO
Müller et al. (2005)
Sources of predictability: snow cover

Correlation for System 3 MAM temperature over 1981-2005 wrt to GHCN temperature (after Shongwe et al., 2007).

- ROC area for anomalies above the upper quartile
- ROC area for anomalies below the lower quartile
Sources of predictability: snow cover

Correlation between February NOAA and February System 3 snow cover (1981-2005)

Correlation between February NOAA snow cover and MAM GHCN temp. (1981-2005)
Sources of predictability: soil moisture

GLACE2 multi-model experiment: 10-member ensembles, 10-start dates (1st April to 15th August) per year over 1986-1995; Series 1 (2) with realistic (unrealistic) soil moisture initialization.

<table>
<thead>
<tr>
<th>acronym</th>
<th>model resolution</th>
<th>remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCCMA</td>
<td>2.8° × 2.8°</td>
<td></td>
</tr>
<tr>
<td>COLA</td>
<td>1.9° × 1.9°</td>
<td>version 3.2</td>
</tr>
<tr>
<td>COLA_CAM</td>
<td>1.4° × 1.4°</td>
<td>NCAR CAM 3.5</td>
</tr>
<tr>
<td>ECHAM</td>
<td>1.9° × 1.9°</td>
<td>version 5; initial soil moisture series 1 derived from different land surface model simulations</td>
</tr>
<tr>
<td>ECMWF</td>
<td>1.1° × 1.1° (*)</td>
<td>Integrated Forecasting System (IFS), ocean-atmosphere coupled</td>
</tr>
<tr>
<td>FSU</td>
<td>1.9° × 1.9°</td>
<td>Soil initialization from data assimilation suite</td>
</tr>
<tr>
<td>KNMI</td>
<td>1.1° × 1.1° (*)</td>
<td>as ECMWF, with prescribed sea surface temperatures</td>
</tr>
<tr>
<td>NCAR</td>
<td>2.8° × 2.8°</td>
<td>CAM 3.0</td>
</tr>
<tr>
<td>NCEP</td>
<td>0.9° × 0.9°</td>
<td>GFS/Noah</td>
</tr>
<tr>
<td>NSIPP</td>
<td>2.5° × 2°</td>
<td>GMAO forecasting system</td>
</tr>
</tbody>
</table>
Sources of predictability: soil moisture

GLACE2 multi-model R2 difference between Series 1 and Series 2. Grid points with statistically significant differences with 98% confidence level are dotted.

16-30 day average

31-45 day average

46-60 day average

van der Hurk et al. (2010)
Sources of predictability: soil moisture

Difference between summer (JJA) 1992 and 1987 anomalies of three ARPEGE SST-forced ensembles and observations.

Uninitialized, interactive SM

SM relaxed to GSWP2

Initialized, interactive SM

ISBA GSWP2 evap, GPCP precip and CRU temp

Conil et al. (2009)
Sources of predictability: stratosphere

Five-day running mean of standardized geopotential height polar cap anomalies from NCEP/NCAR R2 for 2009-10 winter. Red (blue) for positive (negative) anomalies and solid white contours for each 0.5 standard deviation.
Sources of predictability: stratosphere

Daily evolution of the zonal mean of the 50-hPa zonal wind for the ENSEMBLES Stream 1 hindcasts with November start date (the horizontal axis runs from the 1\textsuperscript{st} of November to the 31\textsuperscript{st} of May). Results are averaged over 1991-2001. Black lines are for the ERA40 climatology (with shaded region covering $\pm \sigma$).

Every forecast system underestimates the mean wind. The same happens with the intraseasonal variability that prevents stratospheric sudden warming events from happening with the correct frequency and amplitude.

Maycock et al. (2009)
Sources of predictability: stratosphere

(Left) Correlation between DJF 50-hPa geopotential height and predicted Niño3.4 SST time series with the November start date. (Right) Difference in DJF 46-hPa geopotential height between El Niño and non-El Niño events in HadGAM1.

Ineson and Scaife (2009)
Sources of predictability: volcanoes

Annual-mean air temperature anomalies (wrt to 1950-2009) for 17 stations over Catalonia (northeast Spain).

- El Chichón: April 1982
- Pinatubo: June 1991
Sources of predictability: volcanoes

Mean sea level pressure anomalies wrt 1981-2001 for ERA40 and two sets of one-month lead EC-Earth seasonal hindcasts.

**ERA40**

- With Pinatubo volcanic aerosol
- Without Pinatubo volcanic aerosol
Sources of predictability: volcanoes

Temperature anomalies wrt 1981-2001 for ERA40 and two sets of one-month lead EC-Earth seasonal hindcasts.

ERA40

With Pinatubo volcanic aerosol

Without Pinatubo volcanic aerosol
Summary

- Substantial systematic error, including lack of reliability, is still a fundamental problem in dynamical forecasting and forces \textit{a posteriori} corrections to obtain useful predictions. Don’t take model probabilities as true probabilities.
- Initial conditions are still an important issue.
- Estimating robust forecast quality is difficult, but there are windows of opportunity for reliable skilful predictions, and there is always the anthropogenic warming.
- There is a potential coming from methods that deal with model inadequacy (e.g. multi-model ensembles).
- Many more processes to be analyzed: sea ice, anthropogenic aerosols, ...
Some final thoughts

- In the end we need trustworthy models but model development is a slow process.
- Users will require calibration and can provide feedback on the presentation of forecast information.
- Seasonal forecasting over Europe would benefit from a coordinated effort to improve the forecast systems and to combine climate information from different sources.