Simple Regression Models for Multi-Model Ensembles

David Stephenson and Phil Sansom
Exeter Climate Systems

Acknowledgements:
Giuseppe Zappa, Len Shaffrey, Chris Ferro

Workshop on Stochastic Weather Generators, 29 May – 1st June, 2012,
Roscoff, Brittany, France
How will European storms respond to climate change?

**NERC project TEMPEST:**
Testing and Evaluating Model Predictions of European Storms

**NCAS, Reading***
Len Shaffrey  
Tim Woollings  
Mike Blackburn

**Met Dept, Reading**
Brian Hoskins  
Helen Dacre

**Exeter University**
David Stephenson  
Phil Sansom

**Oxford University**
Tim Palmer

**Met Office (Partner)**
Simon Brown  
Ruth McDonald

**ECMWF (Partner)**
Thomas Jung

*Plus David Brayshaw as Co-I Researcher at Reading
Need simulations from a set of future weather generators …

i.e. a coordinated set of climate model simulations
e.g. CMIP5 Coupled Model Intercomparison Project Phase 5
CMIP5 Multi-Model Ensemble (MME)

\( y_{msr} = 30\text{-year wintertime (DJF) mean of track densities from run } r=1,\ldots,R_{ms} \)

of model \( m=1,\ldots, M=14 \) under scenario \( s=\text{Historical (1975-2004)} \)
or Future (2069-2098 RCP4.5 emissions scenario)

<table>
<thead>
<tr>
<th>Climate model m</th>
<th>No. of runs ( R_{mH} )</th>
<th>No. of runs ( R_{mF} )</th>
<th>( w_m = R_{mH} R_{mF} / (R_{mH} + R_{mF}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCC-CSM1.1</td>
<td>3</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>CanESM2</td>
<td>5</td>
<td>1</td>
<td>0.83</td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>5</td>
<td>1</td>
<td>0.83</td>
</tr>
<tr>
<td>CSIRO-Mk3.6.0</td>
<td>4</td>
<td>5</td>
<td>2.22</td>
</tr>
<tr>
<td>EC-Earth</td>
<td>3</td>
<td>2</td>
<td>1.20</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>INM-CM4</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>IPSL-CM5A-LR</td>
<td>4</td>
<td>3</td>
<td>1.71</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>MIROC5</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>3</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>MPI-ESM-LR</td>
<td>3</td>
<td>3</td>
<td>1.50</td>
</tr>
<tr>
<td>MRI-CGCM3</td>
<td>3</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>NorECM1-M</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>M=14</strong></td>
<td><strong>Sum=38</strong></td>
<td><strong>Sum=23</strong></td>
<td><strong>Sum=13.04</strong></td>
</tr>
</tbody>
</table>
How best to infer climate projections from MMEs?

Collins et al. (2012)
Quantifying future climate change,
Nature Climate Change
Approaches to statistical inference from MMEs

- **Heuristic weighting** e.g.
  \[
  \frac{1}{M} \sum_{m=1}^{M} \left( \bar{Y}_{mF.} - \bar{Y}_{mH.} \right)
  \]

- **Fixed effects regression models**
  \[
  Y_{msr} \sim N(\alpha_m + \beta_s + \gamma_{ms}, \sigma^2) \quad \text{ANOVA}
  \]
  \[
  \bar{Y}_{mF.} - \bar{Y}_{mH.} \sim N(\beta_0 + \beta_1 \bar{Y}_{mH.}, \sigma^2) \quad \text{State-dependent response}
  \]
  e.g. Bracegirdle and Stephenson, Clim Dynamics, 2012

- **Bayesian hierarchical models**
A Multi-Model Ensemble is like my fruit bowl ...

A “fruit bowl of opportunity” \( \{Y_{mrs}; \ m=1,\ldots,M; \ r=1,\ldots,R_{ms}; \ s=H,F\} \)
Note: Not a random sample from one homogeneous population (and it does not include all possible fruit!)
What does reality look like?

It could not have been drawn out of my fruit bowl!

How can we infer properties of this from the fruit in the fruitbowl?
A multi-model mean is a smoothie - a heuristically weighted average of fruits. It is not an item of real fruit! (important information has been lost by averaging)

→ We require modelling frameworks for obtaining samples of real fruit from the posterior distribution p(Y|X) (weather generators?)
Should we use everything in the fruit bowl?

Should we select subsets? How should we weight the fruits?

“All fruit are wrong, but some are tasty” - Granny Smith
Heuristic weighting

The climate change response is often estimated using:

\[
\frac{1}{M} \sum_{m=1}^{M} (\bar{y}_{mF.} - \bar{y}_{mH.})
\]

This descriptive approach:
- assumes all climate models are equal (equal weight per climate model rather than equal weight per run)
- is not resistant to outlier model runs and can not identify overly influential runs
- does not quantify uncertainty (e.g. confidence intervals)
Multi-model mean track density (storms/month)

Mean of historical runs

Mean of future runs

Difference

→ Subtle change in mean density of storms
Two-way ANOVA regression model

\[ Y_{msr} = \alpha_m + \beta_s + \gamma_{ms} + \varepsilon_{msr} \]
\[ \varepsilon_{msr} \sim iid N\left(0, \sigma^2\right) \]

where \( \beta_{mH} = 0 \) and \( \gamma_{mH} = 0 \) for all \( m \) so that:
- \( \alpha_m \) is the mean historical climate of model \( m \);
- \( \beta_s \) is the climate change response of model \( m \);
- \( \gamma_mF \) is the model-dependent climate change response;
- \( \varepsilon_{msr} \) is the natural variability or 'weather noise'.

The point estimate from the ANOVA model is identical to the multi-model mean response using equal weights for all climate models:

\[ \Rightarrow \hat{\beta}_F = \frac{1}{M} \sum_{m=1}^{M} (\bar{y}_{mF} - \bar{y}_{mH}) \]

But if the response is very model-dependent, then the real world response might be very different to the MME response!
The climate change signal is much smaller than spread between models. Model spread is greater than natural variability between runs.

38 Historical  1975-2004 runs
23 Future      2069-2098 RCP4.5 runs
From 14 state-of-the-art climate models
Fitted values from full framework

Mean response identical to climate model mean
Can obtain standard errors as well as means
Residual diagnostics from full model

Dashed lines show 0.5% and 99.5% quantiles of N(0,1)

→ MIROC-ESM is a clear outlier! So throw it out of the MME
Simpler ANOVA regression model

\[ Y_{msr} = \alpha_m + \beta_s + \varepsilon_{msr} \]

\[ \varepsilon_{msr} \sim iid N\left(0, \sigma^2\right) \]

Can use an F-test to select simpler model with no interactions:

- \( H_0 : \gamma_{mF} = 0 \) for all \( m \)  
  **Simpler model with \( M+2 \) parameters**

- \( H_a : \gamma_{mF} \neq 0 \)  
  **Full model with \( 2M+1 \) parameters**

\[ \Rightarrow \hat{\beta}_F = \left( \sum_{m=1}^{M} W_m \right)^{-1} \sum_{m=1}^{M} W_m (\overline{Y}_{mF} - \overline{Y}_{mH}) \]

\[ W_m = \frac{R_{mH} R_{mF}}{R_{mH} + R_{mF}} \]

Weights depend on the number of historical and future runs of each model
Fitted values from simpler framework

→ No-interaction framework reproduces CMIP5 data well
Mean climate change response from both models

Full model

Simpler model

Full minus simple

→ The two frameworks give similar results for storm densities
Normality of residuals: Anderson-Darling p-values

Assumption of normal residuals well satisfied for both models
Is the climate change response model-dependent?

F-test of $H_0 : \gamma_{mF} = 0$ for all $m$ against $H_1 :$ at least one $\gamma_{mF} \neq 0$

Small p-values ($< 0.05$) indicate suggest model-dependent climate change

→ Simpler no-interaction model is hard to reject
→ Storm-track response is not overly model-dependent
Where is the climate change response $\beta'_F$ significant?

T-test of $H_0 : \beta'_F = 0$ against $H_1 : \beta'_F \neq 0$
Small p-values ($< 0.05$) indicate significant climate change

$\rightarrow$ Mediterranean and northern Scandinavia significant change at 5% level
Climate change signal-to-noise F-ratio

→ the estimated climate change signal is generally much smaller than the estimated natural variability in the climate model run

GOOD NEWS for humanity – BAD NEWS for storm track scientists!
Summary

Mean track density in historical runs

Mean climate change response

Statistical significance: p-value

→ Slight increase in the centre of the stormtrack
→ Decrease on the edges especially over the Eastern Mediterranean.
Conclusions

• The multi-model mean estimate of climate change response is equivalent to the mean response estimated from a 2-way linear ANOVA model with model-scenario interactions;

• The ANOVA model can be used to identify outlying climate runs (e.g. the MIROC-ESM climate model runs).

• For storm track density, a simpler more parsimonious model with no interaction term provides a good description of the CMIP5 data – the climate change response from this model correspond to weights that depend on the number of runs of each climate model;

• The climate change response in storm track density is small compared to natural variability and model spread and not very model-dependent;

• Projected storm track response for end of 21st century:
  • Eastern Mediterranean: slight decrease in frequency and intensity
  • Northern Scandinavia: slight decrease in frequency and intensity
  • UK and central Europe: slight increase in storm frequency and intensity
Thank you for your attention

Professor David B. Stephenson
Exeter Climate Systems
Mathematics Research Institute

Any questions or you want to visit us then please contact:
d.b.stephenson@exeter.ac.uk
Additional slides for questions etc. ...
Interaction plots