NARCCAP UQ and Stat Stuff

Stephan R. Sain
Geophysical Statistics Project
Institute for Mathematics Applied to Geosciences
National Center for Atmospheric Research
Boulder, CO

Tammy Greasby, Matt Heaton (NCAR); Reinhard Furrer, Steve Geinitz (Zurich); Cari Kaufman (Berkeley); Dan Cooley, Grant Weller (CSU)

Supported by NSF DMS/ATM.
Outline and Goals

• Understanding sources of variation – functional ANOVA.
  – Implications for design of future experiments, “completing” the table, etc.

• Delivering climate change information and uncertainty – temperature profiles.

• Combining information and model weighting.

• Others – extremes, conveying uncertainty, stat methods, etc.
### The NARCCAP Design

<table>
<thead>
<tr>
<th></th>
<th>Phase I NCEP</th>
<th>Phase II CGCM3</th>
<th>Phase II HADCM3</th>
<th>Phase II CCSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCM</td>
<td>finished</td>
<td>finished</td>
<td>planned</td>
<td>finished</td>
</tr>
<tr>
<td>ECP2</td>
<td>finished</td>
<td>finished</td>
<td>finished running</td>
<td>finished</td>
</tr>
<tr>
<td>HRM3</td>
<td>finished</td>
<td>finished</td>
<td>finished running</td>
<td>finished</td>
</tr>
<tr>
<td>MM5I</td>
<td>finished</td>
<td>finished</td>
<td>finished running</td>
<td>finished</td>
</tr>
<tr>
<td>RCM3</td>
<td>finished</td>
<td>finished</td>
<td>finished running</td>
<td>finished</td>
</tr>
<tr>
<td>WRFG</td>
<td>finished</td>
<td>finished</td>
<td>finished running</td>
<td>finished</td>
</tr>
</tbody>
</table>

- **Phase I**: 1980-2000
- **Phase II**: 1971-2000 (Current), 2041-2070 (Future)
- All future runs use the A2 scenario
- Focus on seasonal summaries
A $2^3$ example

CRCM

WRFG

Current

Future

CCSM

CGCM

CCSM

CGCM

30-year average summer (JJA) temperature.

Reinhard Furrer, Steve Geinitz (Zurich), Kari Kaufman (Berkeley)
Does Model 2 respond to the forcing in the same way that Model 1 does ($\theta_4 = 0$)? Or does it respond in a way that is systematically different ($\theta_4 \neq 0$)?
A $2^3$ example

Baseline

RCM

GCM

Scenario

RCM*GCM

RCM*Scenario

GCM*Scenario

RCM*GCM*Scenario
<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCM</td>
<td>1</td>
<td>5.51</td>
<td>5.51</td>
<td>2.01</td>
<td>0.16</td>
</tr>
<tr>
<td>RCM</td>
<td>1</td>
<td>493.37</td>
<td>493.37</td>
<td>180.27</td>
<td>&lt;2.2e-16</td>
</tr>
<tr>
<td>Scenario</td>
<td>1</td>
<td>387.73</td>
<td>387.73</td>
<td>141.67</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>GCM:RCM</td>
<td>1</td>
<td>47.43</td>
<td>47.43</td>
<td>17.33</td>
<td>4.42e-05</td>
</tr>
<tr>
<td>GCM:Scenario</td>
<td>1</td>
<td>53.92</td>
<td>53.92</td>
<td>19.70</td>
<td>1.4e-05</td>
</tr>
<tr>
<td>RCM:Scenario</td>
<td>1</td>
<td>2.89</td>
<td>2.89</td>
<td>1.06</td>
<td>0.30</td>
</tr>
<tr>
<td>GCM:RCM:Scenario</td>
<td>1</td>
<td>0.08</td>
<td>0.08</td>
<td>0.03</td>
<td>0.86</td>
</tr>
<tr>
<td>Residuals</td>
<td>232</td>
<td>634.94</td>
<td>2.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An ANOVA table for a single grid near the S. Dakota, N. Dakota, Minnesota border.

Interpretation complicated by spatial dependence, large numbers of tests, etc.

Issues can be addressed by treating the “effects” as spatial processes in a functional ANOVA framework embedded in a Bayesian hierarchical model.
Finite-sample variances via a Bayesian hierarchical spatial model
Annual Temperature Profiles

For each grid box, fit a periodic spline w/temporal trend to both current and future runs for a particular model. Allow coefficients to “borrow strength” through a multivariate spatial model on the coefficients of the spline.

Tammy Greasby (NCAR)
Posterior mean length of increase (days) in summer-like temperatures. Other measures (e.g., growing-degree days, heating/cooling-degree days, etc.) can also be examined.
Onset of spring-like temperatures.
The Bayesian Paradigm

Postulate a model (pdf) for data that depends on some parameters:

\[ Y_1, \ldots, Y_n \sim \pi(Y_1, \ldots, Y_n|\theta). \]

⇒ This forms the *likelihood*.

Postulate a model (pdf) for the parameters:

\[ \theta \sim \pi(\theta) \]

⇒ This forms the *prior*.

Inference follows by examining of the posterior distribution:

\[ \pi(\theta|Y_1, \ldots, Y_n) \propto \pi(Y_1, \ldots, Y_n|\theta)\pi(\theta) \]

posterior \( \propto \) likelihood \( \times \) prior

⇒ From *Bayes’ Theorem*. 
A Simple Model

\[ Y_1, \ldots, Y_n \sim \mathcal{N}(\mu, \sigma^2) \]

\[ \mu | \sigma^2 \sim \mathcal{N}(\mu_0, \sigma^2 / \kappa_0) \]

\[ \sigma^2 \sim \text{Inv} - \chi^2 (\nu_0, \sigma_0^2) \]

Posterior distribution for \( \mu \):

\[ p(\mu | Y_1, \ldots, Y_n) = t_{\nu_n}(\mu_n, \sigma_n^2, \kappa_n) \]

\[ \mu_n = \frac{\kappa_0 \mu_0 + n \bar{Y}}{\kappa_0 + n} \]

\[ \kappa_n = \kappa_0 + n \]

\[ \nu_n = \nu_0 + n \]

\[ \nu_n \sigma_n^2 = \nu_0 \sigma_0^2 + (n - 1)s^2 + \frac{\kappa_0 n}{\kappa_0 + n} (\bar{Y} - \mu_0)^2 \]
A Simple Model

\[ Y_1, \ldots, Y_n \sim \mathcal{N}(\mu, \sigma^2) \]

\[ \mu | \sigma^2 \sim \mathcal{N}(\mu_0, \sigma^2 / \kappa_0) \]

\[ \sigma^2 \sim \text{Inv} - \chi^2(\nu_0, \sigma_0^2) \]

Posterior predictive distribution:

1. Sample \( \sigma^2 | \{Y_i\} \) from \( \text{Inv} - \chi^2(\nu_n, \sigma_n^2) \).

2. Sample \( \mu | \sigma^2, \{Y_i\} \) from \( \mathcal{N}(\mu_n, \sigma^2 / \kappa_n) \).

3. Sample \( Y^* \) from \( \mathcal{N}(\mu, \sigma^2) \).
Model weighting

The Tebaldi model:

\[ X_0 \sim \mathcal{N}\left(\mu, \lambda_0^{-1}\right) \]
\[ X_j \sim \mathcal{N}\left(\mu, \lambda_j^{-1}\right) \]
\[ Y_j \sim \mathcal{N}\left(\nu + \beta (X_j - \mu), (\theta \lambda_j)^{-1}\right) \]

- \(X_0\) indicates an observed climate
- \(X_j\) indicates model output for the current time period \(j = 1, \ldots, 13\).
- \(Y_j\) indicates model output for the future time period.
- \(\mu\) is current mean temperature, \(\nu\) is future temperature
- \(\theta\) allows the climate model variance to change between time periods.
- \(\beta\) accounts for correlation between the current and future climate models. \(\beta = 1\) is equivalent to modeling climate change directly.
• Modify Tebaldi model to:
  – Incorporate multiple observational datasets.
  – Model precisions ($\lambda$s), which control the influence of a particular model on estimates of $\mu$, $\nu$ and $\nu - \mu$, as a function of NARCCAP design (GCM, RCM, scenario).
  – *Tammy Greasby (NCAR)*

• Start from scratch (the kitchen sink model):
  – Incorporate multiple observational datasets.
  – Incorporate GCMs, NCEP-driven RCMs, GCM-driven RCMs.
  – Incorporate “familial” relationship between GCMs and RCMs.
  – Include model-to-model correlations and “bias” terms.
  – *Matt Heaton (NCAR)*
The Kitchen Sink Model:
Pacific Southwest winter temperatures.

<table>
<thead>
<tr>
<th>Data</th>
<th>Current Weight</th>
<th>Future Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDEL</td>
<td>0.245</td>
<td>NA</td>
</tr>
<tr>
<td>CRU</td>
<td>0.224</td>
<td>NA</td>
</tr>
<tr>
<td>CRCM</td>
<td>0.006</td>
<td>NA</td>
</tr>
<tr>
<td>ECP2</td>
<td>0.020</td>
<td>NA</td>
</tr>
<tr>
<td>HRM3</td>
<td>0.045</td>
<td>NA</td>
</tr>
<tr>
<td>MM5I</td>
<td>0.021</td>
<td>NA</td>
</tr>
<tr>
<td>RCM3</td>
<td>-0.000</td>
<td>NA</td>
</tr>
<tr>
<td>WRFG</td>
<td>0.010</td>
<td>NA</td>
</tr>
<tr>
<td>CCSM</td>
<td>0.081</td>
<td>0.190</td>
</tr>
<tr>
<td>CGCM3</td>
<td>0.083</td>
<td>0.186</td>
</tr>
<tr>
<td>GFDL</td>
<td>0.066</td>
<td>0.127</td>
</tr>
<tr>
<td>HADCM3</td>
<td>0.084</td>
<td>0.204</td>
</tr>
<tr>
<td>CRCM-ccsm</td>
<td>0.013</td>
<td>0.066</td>
</tr>
<tr>
<td>CRCM-cgcm3</td>
<td>0.012</td>
<td>0.034</td>
</tr>
<tr>
<td>ECP2-gfdl</td>
<td>0.017</td>
<td>0.048</td>
</tr>
<tr>
<td>HRM3-hadcm3</td>
<td>0.025</td>
<td>0.052</td>
</tr>
<tr>
<td>MM5I-ccsm</td>
<td>0.013</td>
<td>0.021</td>
</tr>
<tr>
<td>RCM3-cgcm3</td>
<td>0.011</td>
<td>0.021</td>
</tr>
<tr>
<td>RCM3-gfdl</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>WFRG-ccsm</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>WFRG-cgcm3</td>
<td>0.014</td>
<td>0.050</td>
</tr>
</tbody>
</table>

Pr(Fut>Cur) = 0.98
Pr(Fut>Cur) = 0.76
Pr(Fut>Cur) = 0.31
Questions?

Many opportunities for visits and collaboration: ASP, RSVP, SIParCs, GSP, IMGae, Theme-of-the-Year,...

ssain@ucar.edu
http://www.image.ucar.edu/~ssain

Thank You!


