The key principles in dealing with multiple observational constraints and imperfect models, and implications for constraining ECS

David Sexton
23-27th March 2015
WCRP Grand Challenge Workshop: Earth's Climate Sensitivities
Equilibrium climate sensitivity (ECS) and transient climate response (TCR) are useful metrics summarizing the global climate system. ECS is defined as the equilibrium change in annual mean global surface temperature following a doubling of the atmospheric CO$_2$ concentration, while TCR determines the warming expected at a given time past the initial doubling.

Newer studies of constraints based on the observed warming since 1880 have revised the 1.5°C to 4.5°C range, with the CMIP5 model mean at 3.2°C, similar to CMIP3. Those estimates are scaled by the RF of 2W m$^{-2}$ for equilibrium climate sensitivity, based on Figure 10.20b plus climatological constraints.

Critical methods are sensitive to assumptions of independence of the variance of radiative forcing over a period of 70 years because the experts base their opinion on the same studies as we do; therefore, the elicitation for PDFs of climate sensitivity exist.

Methods are sensitive to assumptions of independence of the variance of radiative forcing over a period of 70 years.
Some things we have discussed this week that affect constraints

- Two-timescale model to capture state dependence and evolving lambda
- Forcings and fast adjustments
- Volcanoes “lower sensitivity” and “forcing 30% too high”
- Model’s indirect aerosol effect weaker
- What can explain the residuals from D&A between 1910-1940?
- What caused the grand hiatus 1945-1975?
- What caused the hiatus 1998 onwards?
- For a decent quantification of uncertainty, framework must be able to handle these issues.
Aim

• Brief overview of the Bayesian framework
• Show that the statistical model aims at constraints which want to get the physics/dynamics right:
  • accounting for model imperfections
  • using multiple constraints
• That way try to avoid over-confident constraints.
Monte Carlo sampling of metrics and prediction variables

Relationship of Weights (based on likelihood function) versus Climate Sensitivity for 1000000 parameter combinations
Comparing models with observations

- Skill of model is likelihood of model data given a set of observations

\[
\log L_o(m) = -c - \frac{n}{2} \log |V| - \frac{1}{2} (m - o)^T V^{-1} (m - o)
\]

\( V = \text{observational} + \text{simulator/emulator uncertainty} + \text{discrepancy error} + \text{(model error)} \)
Structural model errors  
(Goldstein and Rougier 2004)

- Model not perfect so there are processes in real system but not in our model that could alter model response by an uncertain amount.

- Discrepancy captures the uncertainty about what would happen to model outputs if all model errors were magically corrected.

- Bjorn led a discussion on Tuesday effectively about discrepancy for fixed-lambda model.

\[ y = f(x^*) + d \]

- True climate
- Model output of best choice of parameter values \( x^* \)
- Discrepancy \( d=0 \) for perfect model
Emergent relationships and model error

Figure 9.45 | (a) Snow-albedo feedback

Error bars that are added to figure are purely for schematic purposes

Hall and Qu 2006
Emergent relationships and model error

(a) Snow-albedo feedback

Error bars that are added to figure are purely for schematic purposes

Hall and Qu 2006
Importance of considering the uncertainties

\[ V = \text{observational} + \text{simulator/emulator} + \text{discrepancy uncertainty} + \text{error} \] (model error)

- The effectiveness of potential constraints can be compromised in a number of ways:
  - Poor quality measurements
  - Poor sampling (short time series, poor spatial coverage) - either model/emulator or observations
  - Small ensemble size for PPE
  - Large internal variability - either model/emulator or observations
  - Using variables that are relatively poorly represented compared with other models.
  - Biases in models
Effect of historical discrepancy on weighting

Discrepancy included                 excluded

Estimated from sample size of 50000

The terms in $V$ are there to guard against overly sharp likelihoods and overly-confident predictions
Why is it important to account for uncertainties?

Aldrin et al (2012):
- Simple Energy Balance Model
- Comprehensive set of radiative forcing
- Three metrics (NH and SH temperature and ocean heat content)
- Sophisticated Bayesian analysis
- Sensitivity tests
Figure 14: Prior (black, grey) and posterior (red, pink) distribution for each radiative forcing component with mean (black, red) and 95% credible interval (grey, pink).
Fig. 10. For the continuous parameters, the evolution of the ratio of the posterior probability to the prior probability for 20 intervals across the range of each parameter, as more eigenvectors (denoted by the colours—see key in VF1 panel) are included in the weighting function. Curves that lie close to 1 across the range are not much affected by the likelihood weighting. Parameter values for the standard version of HadSM3 are indicated by the red dash on the x-axis.

D. M. H. Sexton et al.: Multivariate probabilistic projections
Multiple constraints - no model is the best at every variable...
How ‘best model’ MSE increases as you consider more variables

Best model performance for each variable often used as a way to include model error but considering multiple variables shows models not quite as good as that assumption indicates.
Summary

• Getting the physics/dynamics right! Accounting for model imperfections, and using multiple observations are important as it means:
  • development of models and observational data sets matter!
  • need to consider how model imperfections affect results e.g. Bjorn’s discussion - how about discuss two timescale model?

• Tries to avoid over-confident results when we have model imperfections, and results that jump around as methods/models develop

• Need to make sure that constrained ranges/PDFs consider the set of observational uncertainty, simulator/emulator uncertainty and discrepancy. So historical constraints ok as long as this is done correctly.
Any questions?
Can use several metrics to constrain climate projections

First of six metrics used in Sexton et al (2012) and UKCP09

More metrics, less chance for rewarding a poor model
Model not perfect so there are processes in real system but not in our model that could alter model response by an uncertain amount.

Discrepancy captures the uncertainty about what would happen to model outputs if all model errors magically corrected.

\[ y = \beta f(x^*) + d \]

\( \mathcal{B} = 1 \) for perfect model

\( d = 0 \) for perfect model
Multiple constraints

Climate Prediction Index
Probabilistic prediction of equilibrium response to double CO2

Does not account for systematic errors common to all climate models.
MSE metric v Climate sensitivity for small sample

MSE v climate sensitivity for nine different sub-samples of size 50 from 1 million Monte Carlo sample
Moving from uncertainty to probability

UKCIP02
Single projection

Very unlikely to be less than (10%)

UKCP09
Central estimate (50%)

Very unlikely to be more than (90%)

Summer Rainfall 2080’s
Sensitivity studies - sampling parameters in a controlled way

Global climate sensitivity (Rougier et al., 2009)
From QUMP and cpdn ensembles of HadSM3