An Idiot's Guide to Emergent Constraints

Julia Hargreaves (an idiot)
Martin Renoult, James Annan, Thorsten Mauritsen
14 Feb
Motivation

- Way back in the mists of time....(2012)

mid-Pliocene Warm Period

- But do we know what we are doing or why?
- What does the estimate mean — Bayesian/Frequentist/???
- What are our assumptions?
Problem 1: What are our prior assumptions?

- We tried a Bayesian estimate in 2012, but did not like using the ensemble as the prior.

- With least squares (usually used in emergent constraint studies) we can extrapolate outside ensemble. This feels good, but is it valid?
Problem 2: How to do the regression?

- Assume uncertainties on X or Y axes? Or both? Which is the Predictor and which the Predictand?
Research Goals
(and outline of talk)

- Establish a (Bayesian) framework that makes sense
- Test with some old paleo models
- Apply to the newest PMIP4/CMIP6 models
Bayes

Probability of climate sensitivity (S) given some observations

\[ p(S|T_0) \propto p(S) \, p(T_0|S) \]

Prior beliefs about S

Probability of getting those observations given a value of S

Likelihood

Posterior
What is our likelihood model?

What are we doing when we “do emergent constraints”?

- We believe that a *relationship* in the ensemble between an observable $T_M$ and parameter $S_M$ can be used to constrain “real” $S$ using observations $T_O$.
- So we need to model that relationship in our likelihood.
Problem 2: Which way round to do the calculation?

- Bayes to the rescue!
- The observations enter this process via the likelihood $p(T_0|S)$ which is a model that takes $S$ as an input parameter and predicts the observations $T$ that we expect (probabilistically)
- If we want to use an empirical quasi-linear relationship between $S$ and $T$ it needs to use: $S$ as the predictor and $T$ as predictand.
- This is the reverse of what all emergent constraint work has done!
Calculating the likelihood

1. Bayesian linear regression

- Model: $T_M = \alpha S_M + \beta + \epsilon, \ \epsilon \sim N(0,\sigma)$

- We have priors over the regression parameters $\alpha, \beta, \sigma$

  - This is a good thing as it forces us to make and state explicit judgments about the relationship that we expect to find!

- We test the sensitivity of results to our prior
Calculating the likelihood

2. Doing it with it on a computer

\[ T_M = \alpha S_M + \beta + \varepsilon, \quad \varepsilon \sim N(0,\sigma) \]

- Use MCMC to wander through parameter space \((\alpha, \beta, \sigma)\)
- Likelihood of any triple \((\alpha, \beta, \sigma)\) is probability of the climate model ensemble generating its set of \(T_M\) according to that instance of the likelihood model
- MCMC generates posterior ensemble of triples \((\alpha, \beta, \sigma)\) conditioned on model ensemble \((S_M,T_M)\)
- Likelihood \(p(T_O|S)\) is calculated by integrating over the \(S\) posterior distribution of \((\alpha, \beta, \sigma)\)
  - ie: calculate the probability of \(T_O\) for each value of \(S\) and add it all together
Doing the Bayesian updating

- The likelihood doesn't create the posterior, it only updates a prior:
  \[ p(S|T_0) \propto p(S) \cdot p(T_0|S) \]

- Need a prior on S.

- This can be whatever you like! Specified entirely separately from the rest of the analysis (Another bonus compared to other approaches)
Example: using the Last Glacial Maximum

- LGM: 23-19ka, cold climate with low CO₂ and large ice sheets
- Expect to find a relationship between $S$ and the tropical temp anomaly $T$
- Prior: $\alpha, \beta, \sim N(0,1)$ $\sigma \sim$ half-Cauchy (+ve)
- Prior on $S$: Cauchy (after Annan and Hargreaves 2011)
Results

- Broadly similar results to the ad-hoc OLS approach
- But it makes the assumptions explicit and adjustable
mid-Pliocene Warm Period

• 3.2Ma, 400ppm CO$_2$, warm interglacials

Similar result to previous work

Martin Renoult et al CPD
Different PMIP ensembles

PMIP2

PMIP2+3

PMIP3
Sequential updating!

- Can use the posterior from LGM as the prior for mPWP
- (assumption of independence)
Summary

- Bayesian principles suggest using climate parameter (e.g., S) as predictor with observation as predictand
- Bayesian Linear Regression with priors over parameters and a separate prior over S
- Multiple constraints follow immediately as a consequence
- Still waiting for more PMIP4 simulations
- Ensemble inadequacy can be included in the Bayesian analysis, but we don’t know how big or important it is, so still a fundamental problem here! Also there will always be a very small ensemble size in statistical terms!