INSTITUTE FOR GEOPHYSICS
JACKSON SCHOOL OF GEOSCIENCES

THE UNIVERSITY OF TEXAS AT AUSTIN

The Use of Observations in Uncertainty Quantification

Charles Jackson

University of Texas at Austin

ICTP workshop on "Uncertainty Quantification in Climate Modeling and Projection"
Trieste Italy
13-17 July, 2015

Topics

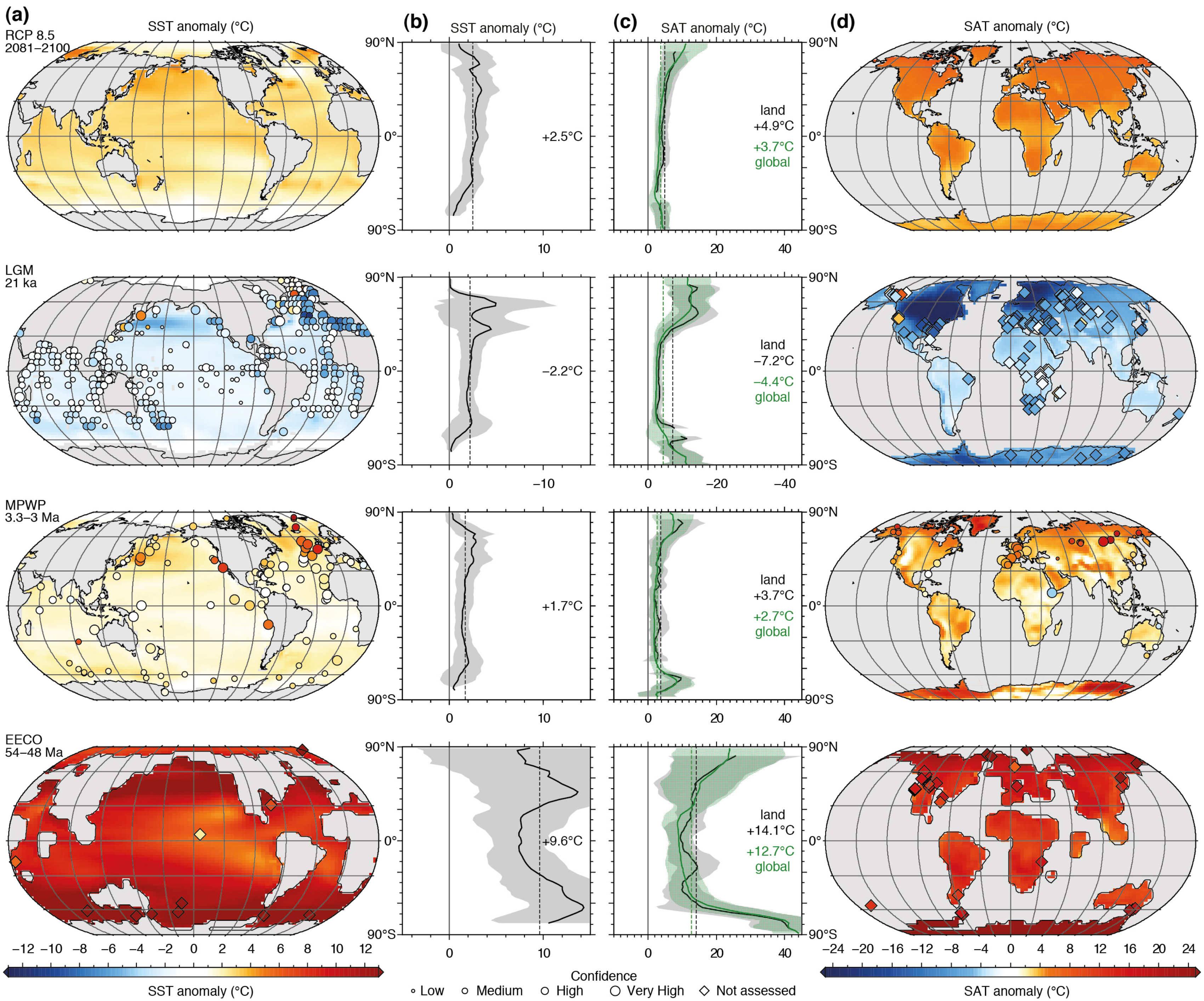
1. Paleoclimate data
2. Probabilistic framework
3. Irreducible errors
4. Emergent Constraints

Summary

- Data is used in the process of model selection.
- Not all data is important to predictions.
- Most important is to identify how data will used to test processes and to reduce the influence of data that is unrelated.

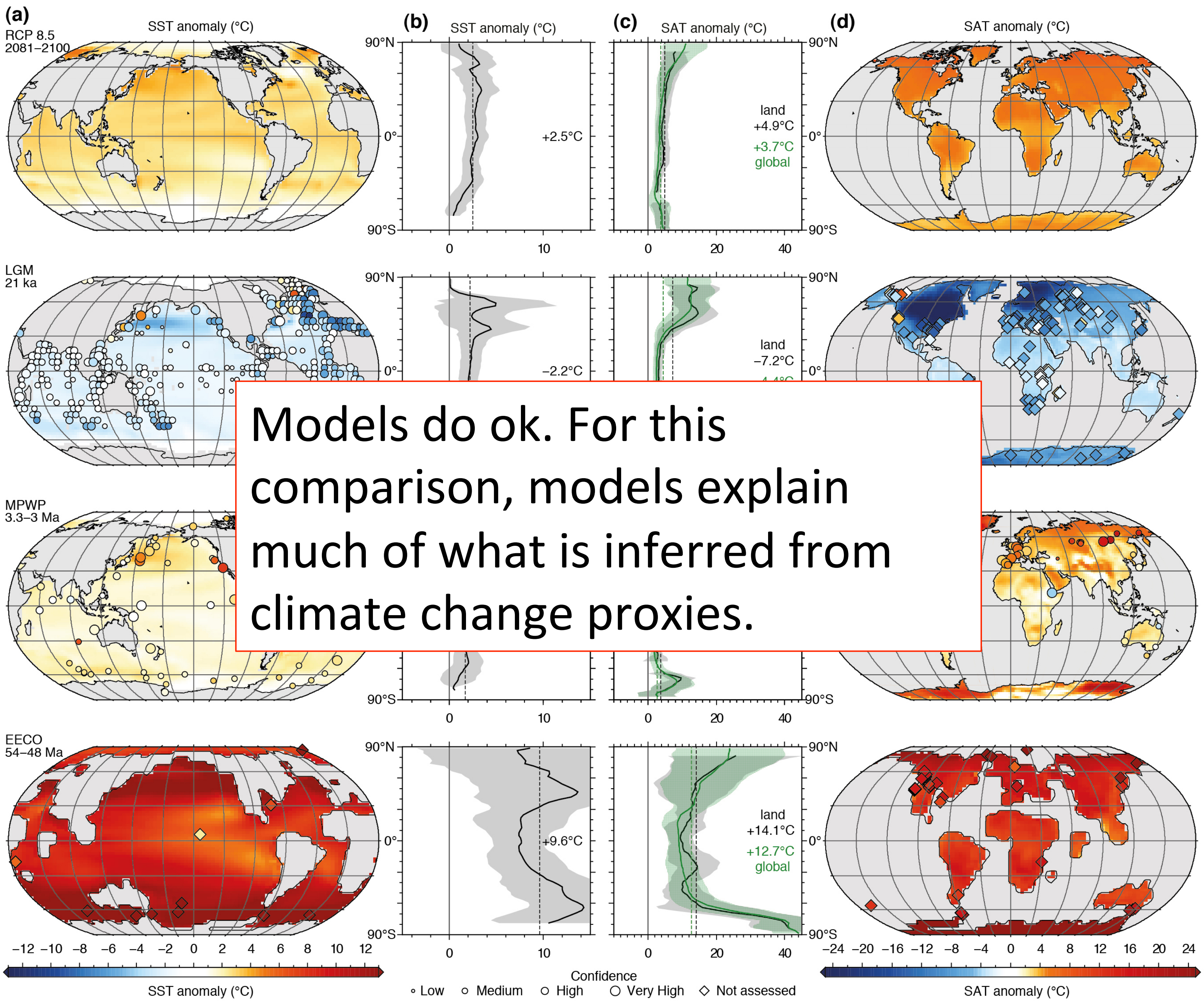
Paleodata: inferences of past climate change as recorded by fossils

- Reflects response to known forcings.
- Provides unique test for climate models … Gets away from problem of using same data to test models as was used to develop them.
- Shows potential risks of a changing climate.



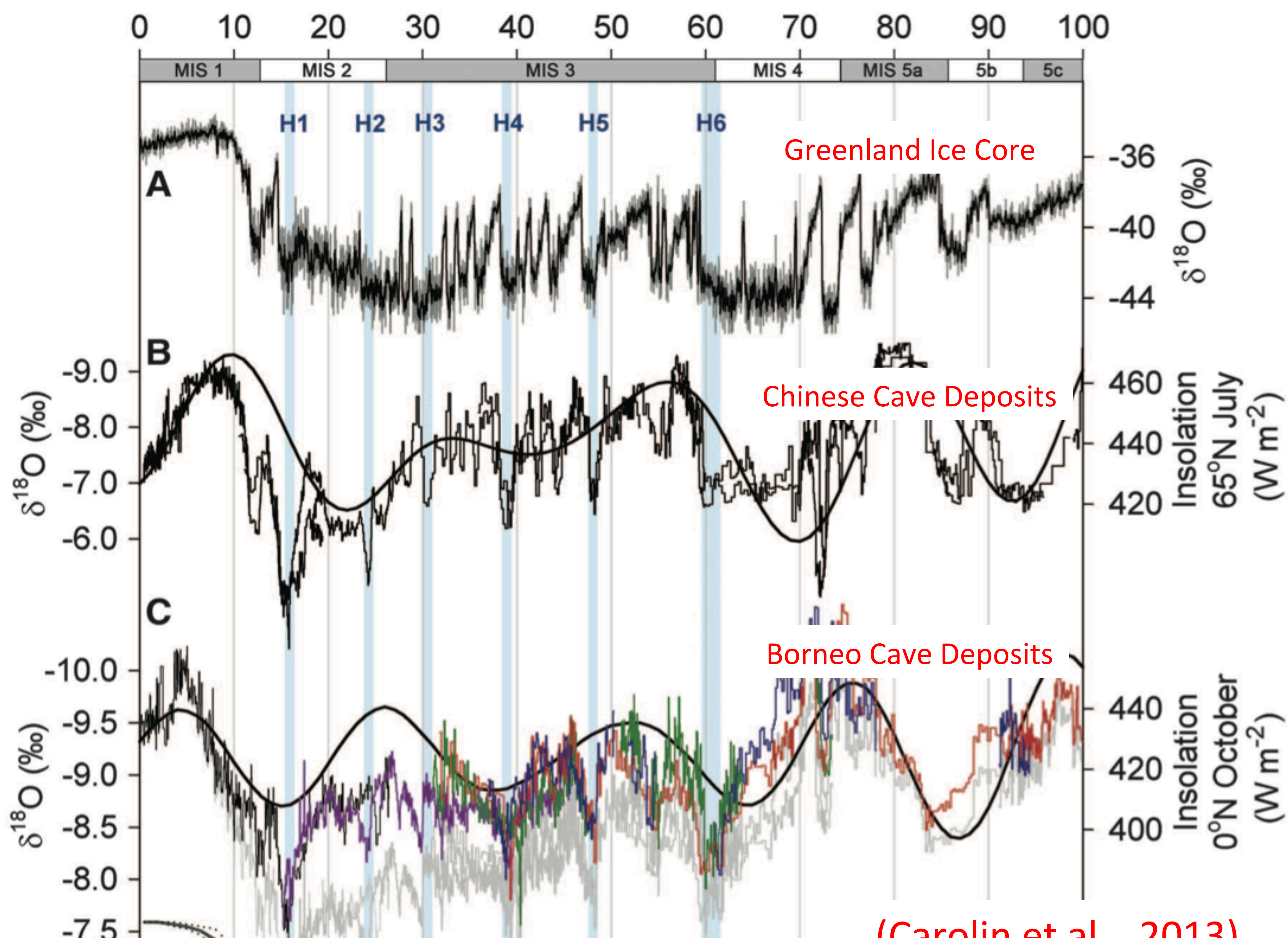
(a)
RCP 8.5
2081–2100
SST anomaly (°C)
(b)
SST anomaly (°C)
+2.5°C
(c)
SAT anomaly (°C)
land
+4.9°C
+3.7°C
global
(d)
SAT anomaly (°C)
LGM
21 ka
−2.2°C
land
−7.2°C
−4.4°C
global
MPWP
3.3–3 Ma
+1.7°C
land
+3.7°C
+2.7°C
global
EECO
54–48 Ma
+9.6°C
land
+14.1°C
+12.7°C
global
90°N
0°
90°S
−12 −10 −8 −6 −4 −2 0 2 4 6 8 10 12
SST anomaly (°C)
Confidence
Low
Medium
High
Very High
Not assessed
−24 −20 −16 −12 −8 −4 0 4 8 12 16 20 24
SAT anomaly (°C)

IPCC 2013



Models do ok. For this comparison, models explain much of what is inferred from climate change proxies.

IPCC 2013



(Carolin et al., 2013)

Paleodata summary

- Climate models do ok reproducing broad features of forced climate change.
- Questions remain concerning the nature of observed abrupt transitions in climate.
- During the past 10 kyrs, there appears to be significant power at millennial time scales which can not be cleanly linked to known forcings (Khider et al, 2014).

How observations fit into a Bayesian probabilistic framework:

Bayesian inference

$$ppd(\mathbf{m} \mid \mathbf{d}) \propto p(\mathbf{d} \mid \mathbf{m})\, p(\mathbf{m})$$

$$PPD(\mathbf{m} \mid \mathbf{d}_{obs}, g(\mathbf{m})) \propto \exp\left[-\frac{1}{2}\left(g(\mathbf{m}) - \mathbf{d}_{obs}\right)^T \mathbf{C}_{noise}^{-1}\left(g(\mathbf{m}) - \mathbf{d}_{obs}\right)\right] \cdot prior(\mathbf{m})$$

Likelihood test statistic

Uncertainty of a model estimate of a single observation

$$d_1 = \hat{x}_1 + \varepsilon_x$$

$$\hat{x}_1 = g(\mathbf{m})$$

$$\mathbf{m} = \text{ model parameters}$$

$$\varepsilon_x \sim N(0, \sigma_x)$$

$$p(d_1 \mid \mathbf{m}) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(\hat{x}_1 - d_1)^2}{2\sigma_x^2} \right)$$

Uncertainty of a model estimate of two correlated observations

$$p\left(d_1,d_2 \mid \hat{x}_1,\hat{x}_2,\mathbf{m}\right)=\frac{1}{\sigma_{x_1}\sigma_{x_1}2\pi\left(1-\rho^2\right)}\exp\left(-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(\hat{x}_1-d_1\right)^2}{\sigma_{x_1}^2}+\frac{\left(\hat{x}_2-d_2\right)^2}{\sigma_{x_2}^2}-2\rho\frac{\left(\hat{x}_1-d_1\right)\left(\hat{x}_2-d_2\right)}{\sigma_{x_1}\sigma_{x_2}}\right]\right)$$

Or expressed with matrix notation:

$$p\left(\mathbf{D}\mid\mathbf{X},\mathbf{m}\right)=\frac{1}{2\pi\left|\mathbf{C}\right|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left[\mathbf{X}-\mathbf{D}\right]^T\mathbf{C}^{-1}\left[\mathbf{X}-\mathbf{D}\right]\right)$$

$$X=\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix},\quad \mathbf{C}=\begin{bmatrix}\sigma_{x_1}^2 & \rho\sigma_{x_1}\sigma_{x_2}\\ \rho\sigma_{x_1}\sigma_{x_2} & \sigma_{x_2}^2\end{bmatrix}$$

Now consider the chi-squared test statistic

The sum of k independent normal random variables

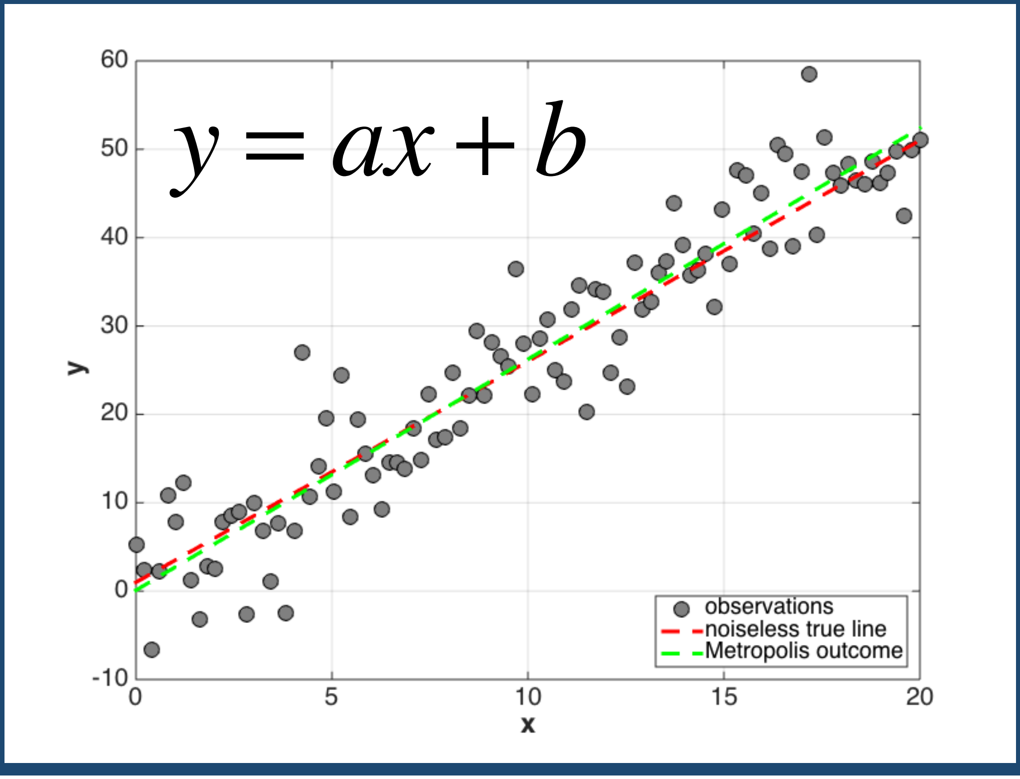
$$\chi^2 \sim \sum_{i=1}^{k} \frac{\left(x - \overline{x}\right)^2}{\sigma^2}$$

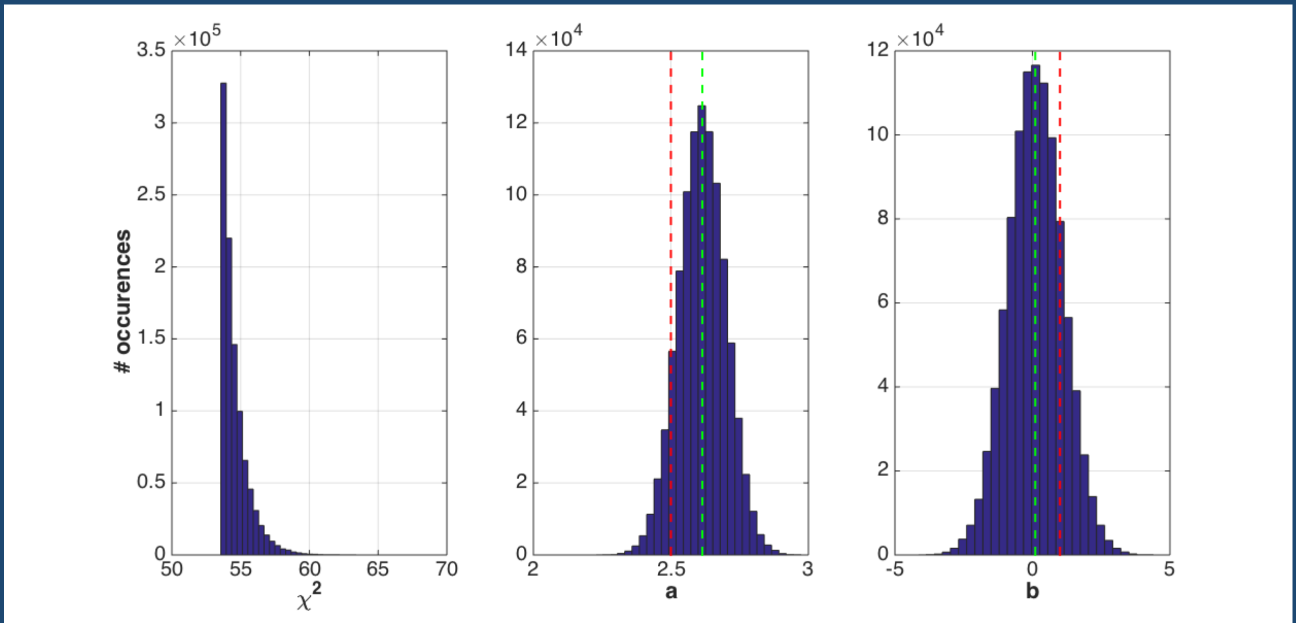
$$\left\langle \chi^2 \right\rangle = k$$

$$\mathrm{var}\left(\chi^2\right) = 2k$$

Sampling from the likelihood distribution provides a measure of those choices of **m** that are consistent with the data given the uncertainties in the data and the effective degrees of freedom in the data in the same way that the chi-squared statistic tests a null hypothesis.

So the uncertainty in the slope and intercept when fitting a line through a set of points depends on the number of points, and whether all the errors in the data are independent or not.



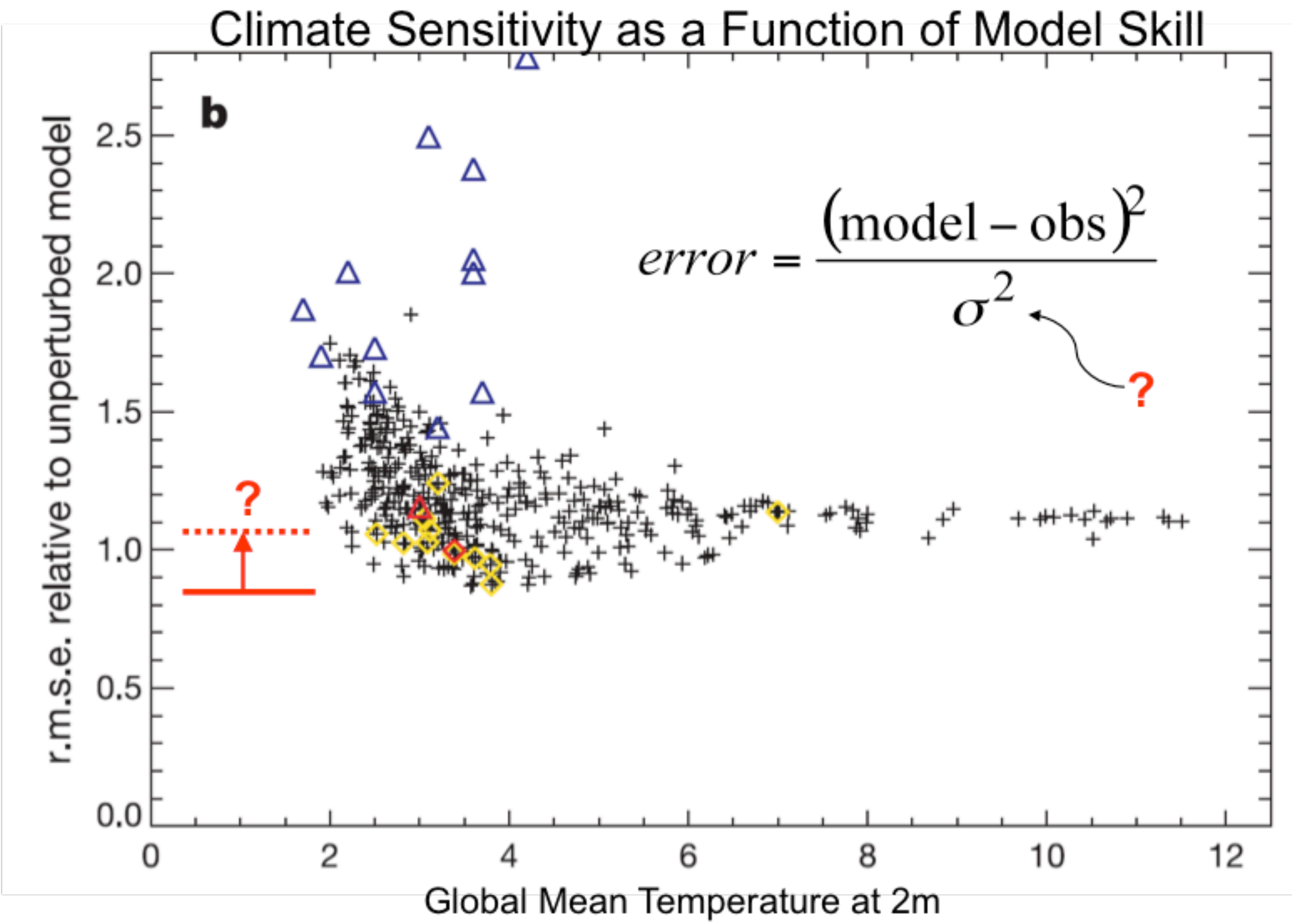


Probabilistic framework summary

- Climate data affects uncertainties through a test statistic that is incorporated within the likelihood function.
- Test statistics involve “degrees of freedom” which is a measure of the independent bits of information.
- Major challenges remain in knowing how to represent dependencies that exist within the data we use to test the representation of climate phenomena.

Irreducible error

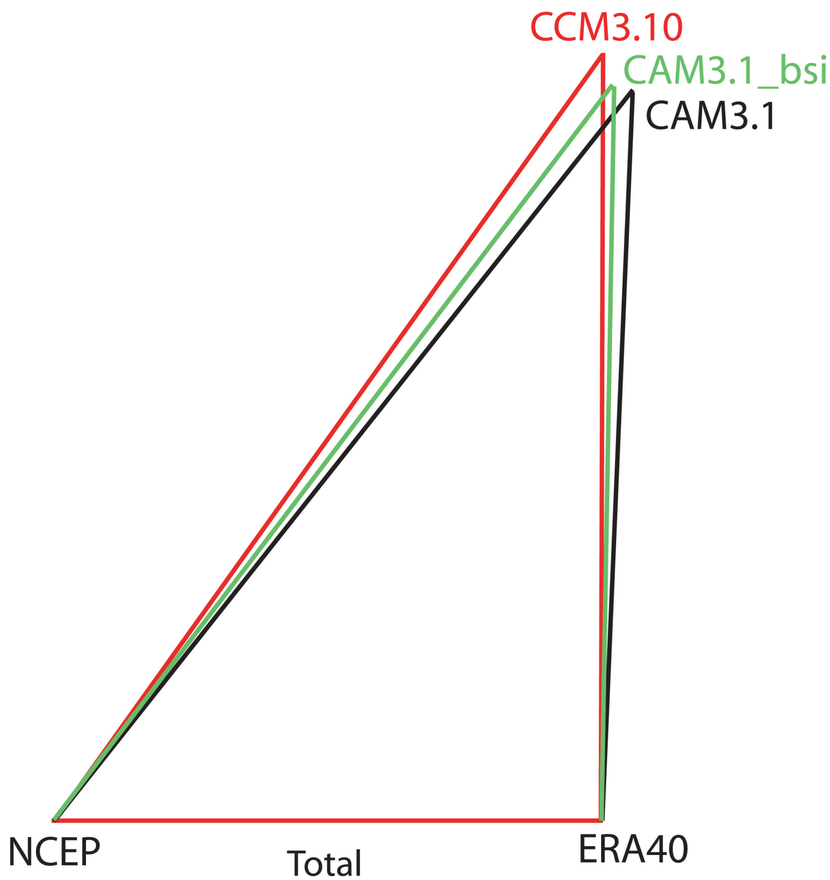
- Refers to the gap that exists between a model and data that no amount of tuning will eliminate.
- Sometimes referred to as “structural error” or “model discrepancy”.
- Has the potential to throw off model calibration, producing good results for the wrong reasons.

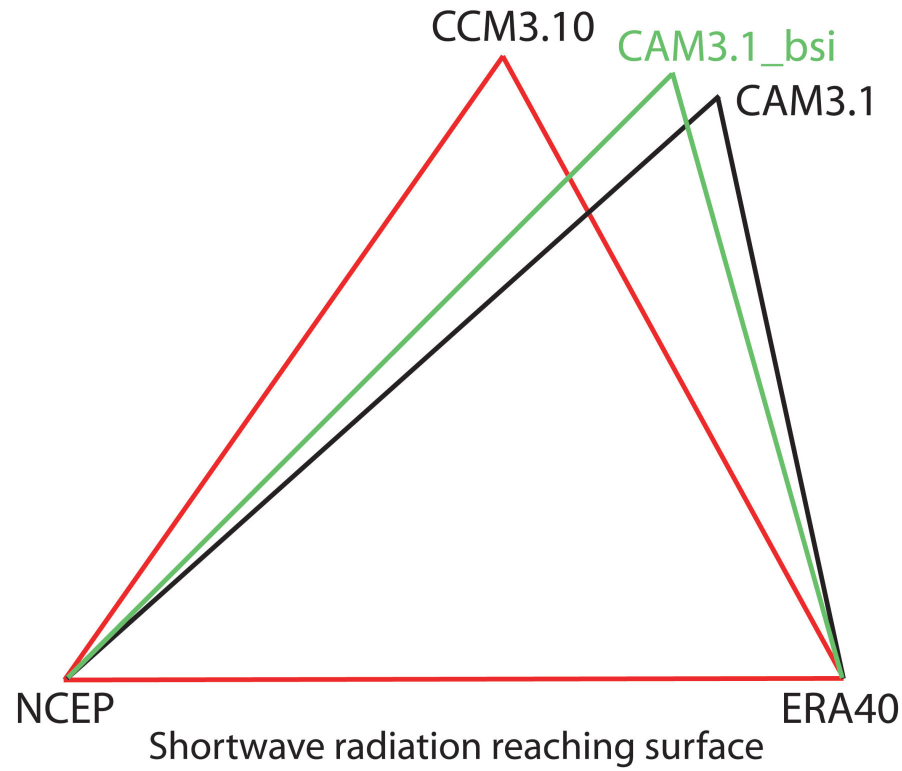


Grid-search, no rejections,
no weighting

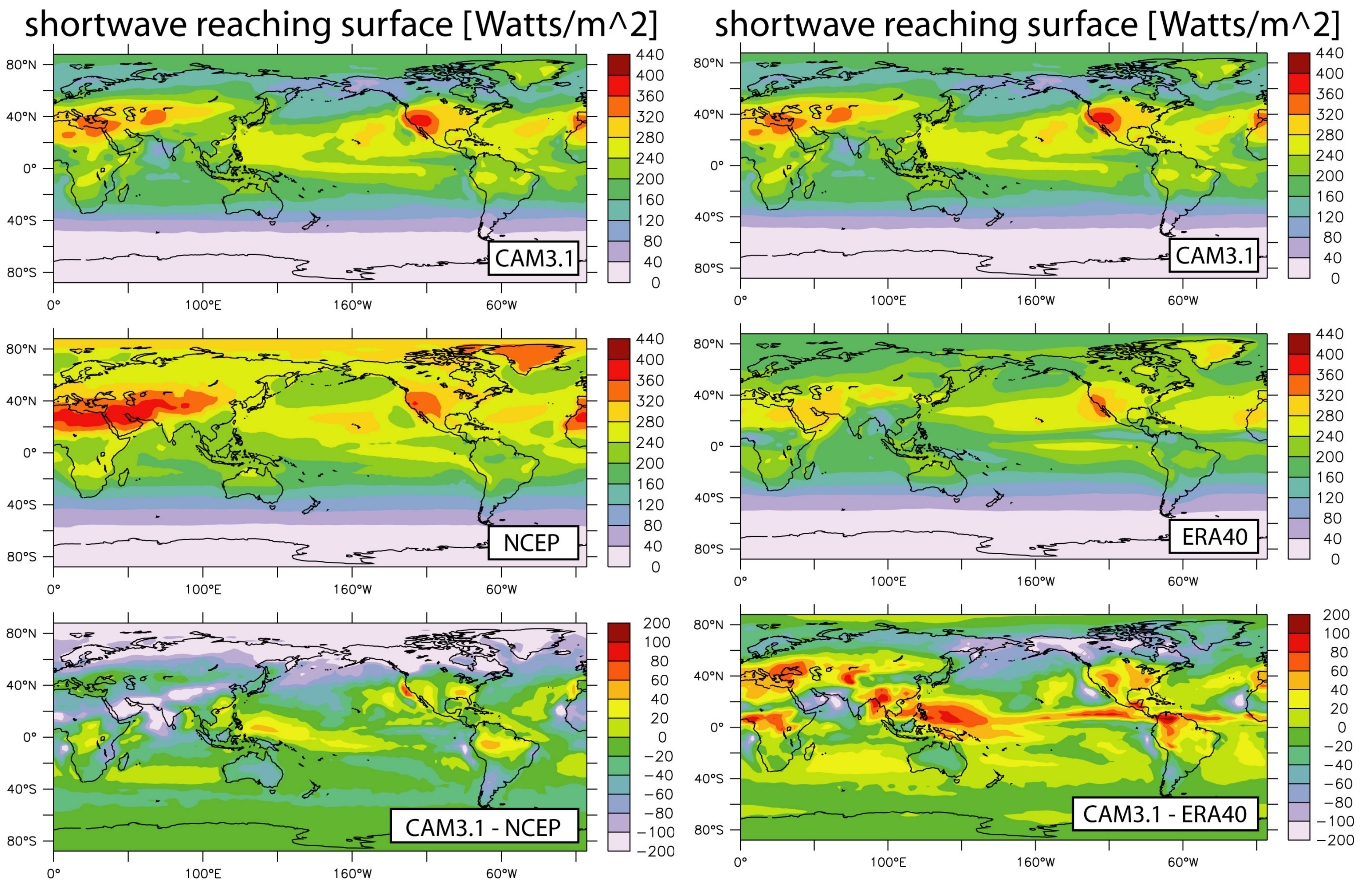
(Stainforth et al., Nature 2005)

Distance between different reanalysis products is similar in size to any model and those products.



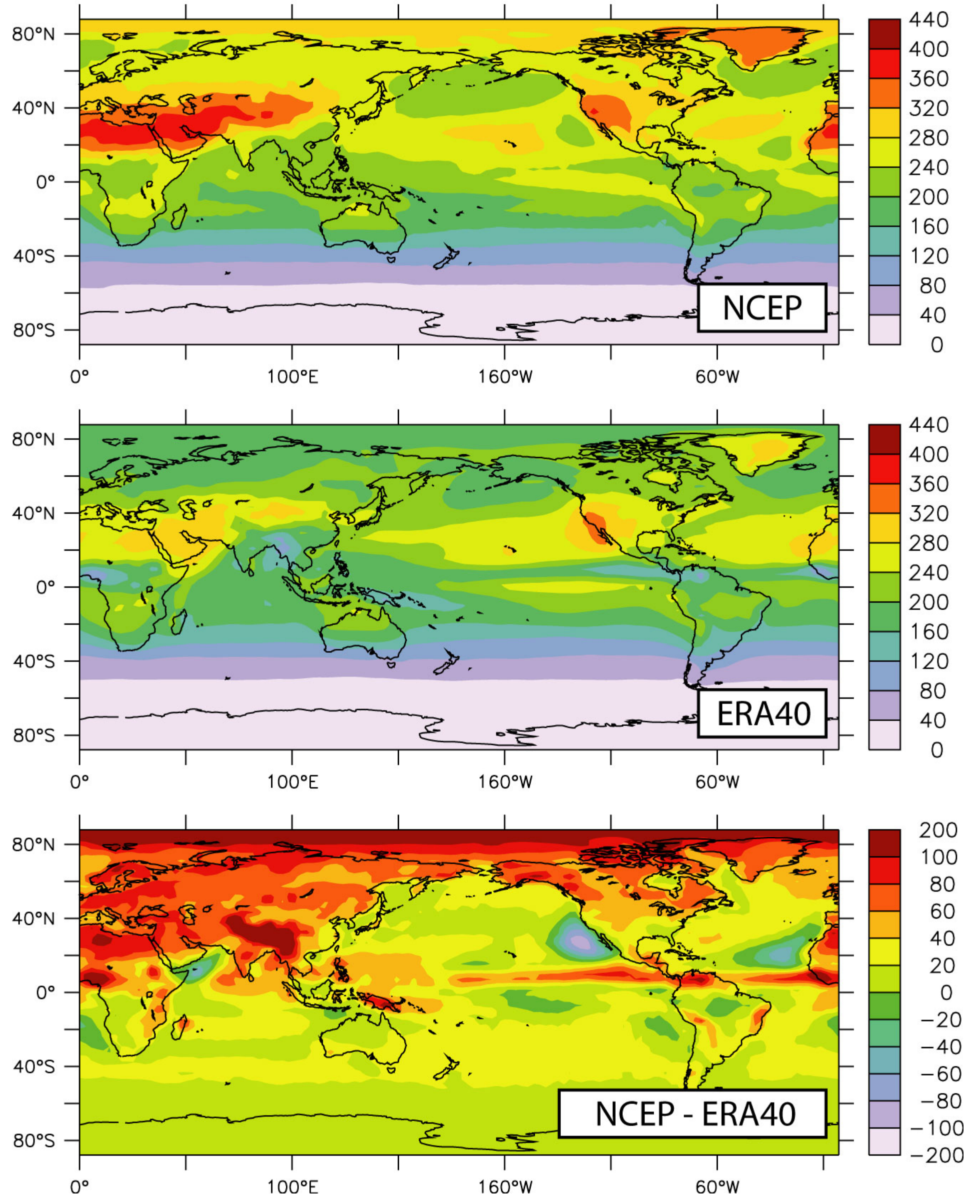


June – July model discrepancies with reanalysis data

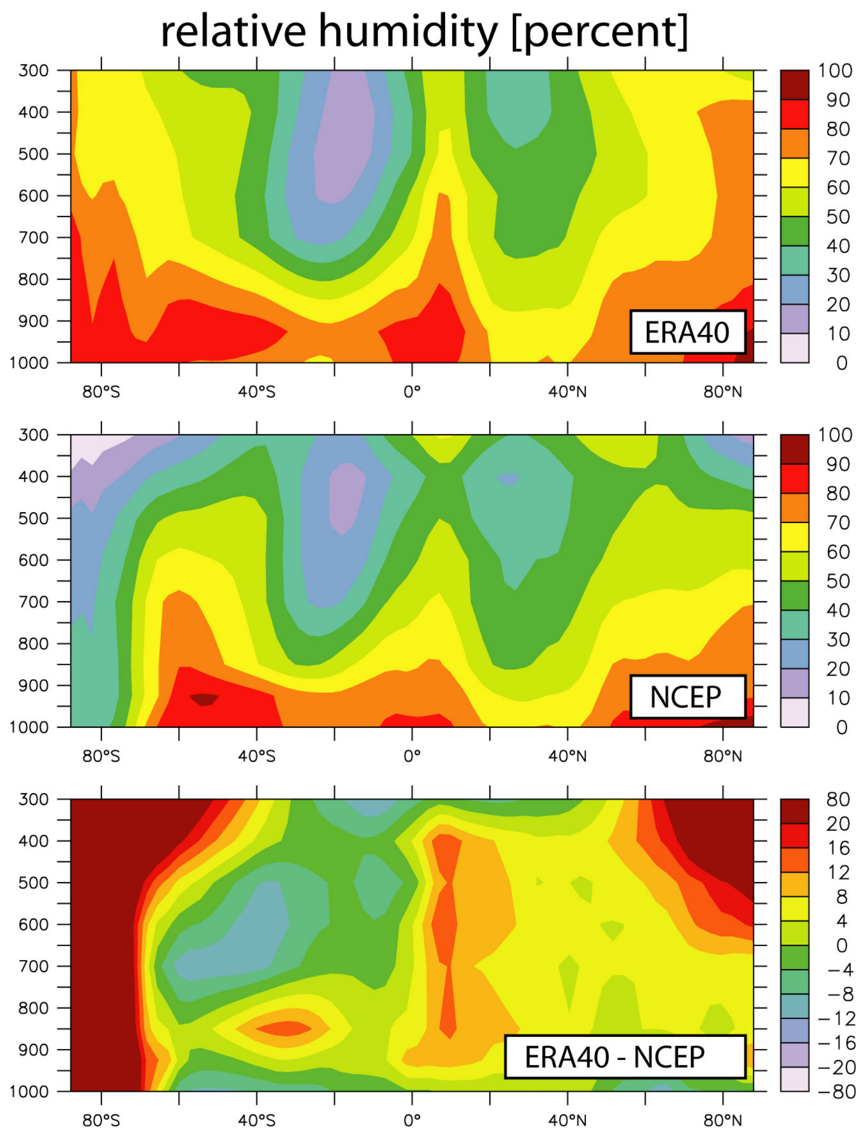


June – July NCEP discrepancies with ERA40

shortwave reaching surface [Watts/m^2]



June – July NCEP discrepancies with ERA40



What to do about it?

Option 1: Add a discrepancy term to likelihood which is an additional unknown. (Brynjarsdottir and O'Hagan 2014) … very readable summary paper.

$$d_1 = \hat{x}_1 + \varepsilon_x + \delta_x$$

$$\hat{x}_1 = g(\mathbf{m})$$

$$\delta_x \sim N(\mu_x, \sigma_\delta) \text{ i.e. use Gaussian Process Model}$$

$$\varepsilon_x \sim N(0, \sigma_x)$$

$$p(d_1 \mid \mathbf{m}) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(\hat{x}_1 - d_1 - \delta_x)^2}{2\sigma_x^2} \right)$$

What to do about it (2)?

Option 2: Scale the variance. (Jackson et al., 2007)

$$d_1 = \hat{x}_1 + \varepsilon_x$$

$$\hat{x}_1 = g(\mathbf{m})$$

$$\varepsilon_x \sim N\left(0, \frac{1}{S}\sigma_x^2\right)$$

$$S \sim gamma\left(\alpha, \beta + E(\mathbf{m})\right)$$

$$E(\mathbf{m}) = \text{cost function (i.e. log-likelihood)}$$

$$p(d_1 \mid \mathbf{m}) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{S(\hat{x}_1 - d_1)^2}{2\sigma_x^2}\right)$$

Hierarchical Bayes strategy to add in "hyper" parameter S

Using Bayes' theorem we now have

$$p(\mathbf{m}, S|\mathbf{d}_{obs}) \propto l(\mathbf{d}_{obs}|\mathbf{m}, S)p(m)p(S)$$

A very important point from this last expression is that,

$$\begin{aligned} p(\mathbf{m}|S, \mathbf{d}_{obs}) &\propto l(\mathbf{d}_{obs}|\mathbf{m}, S)p(m) \\ &\propto exp(-SE(\mathbf{m}))p(m) \end{aligned}$$

and

$$p(S|\mathbf{m}, \mathbf{d}_{obs}) \propto l(\mathbf{d}_{obs}|\mathbf{m}, S)p(S)$$

Jackson et al. 2007

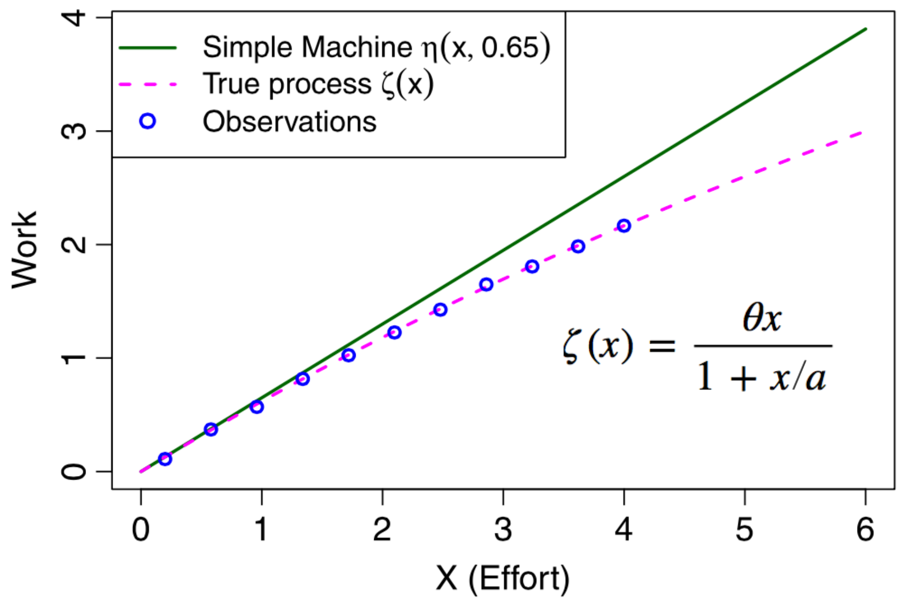
This implies that for each step we can iteratively generate a value of $\mathbf{m}$ conditional on S and a value of S conditional on $\mathbf{m}$ in the following way:

1. To simulate $\mathbf{m}$ conditional on S, apply sampling algorithm for $\mathbf{m}$ but just one iteration.
2. To simulate S conditional on $\mathbf{m}$:
   - For the informative gamma distribution, we have

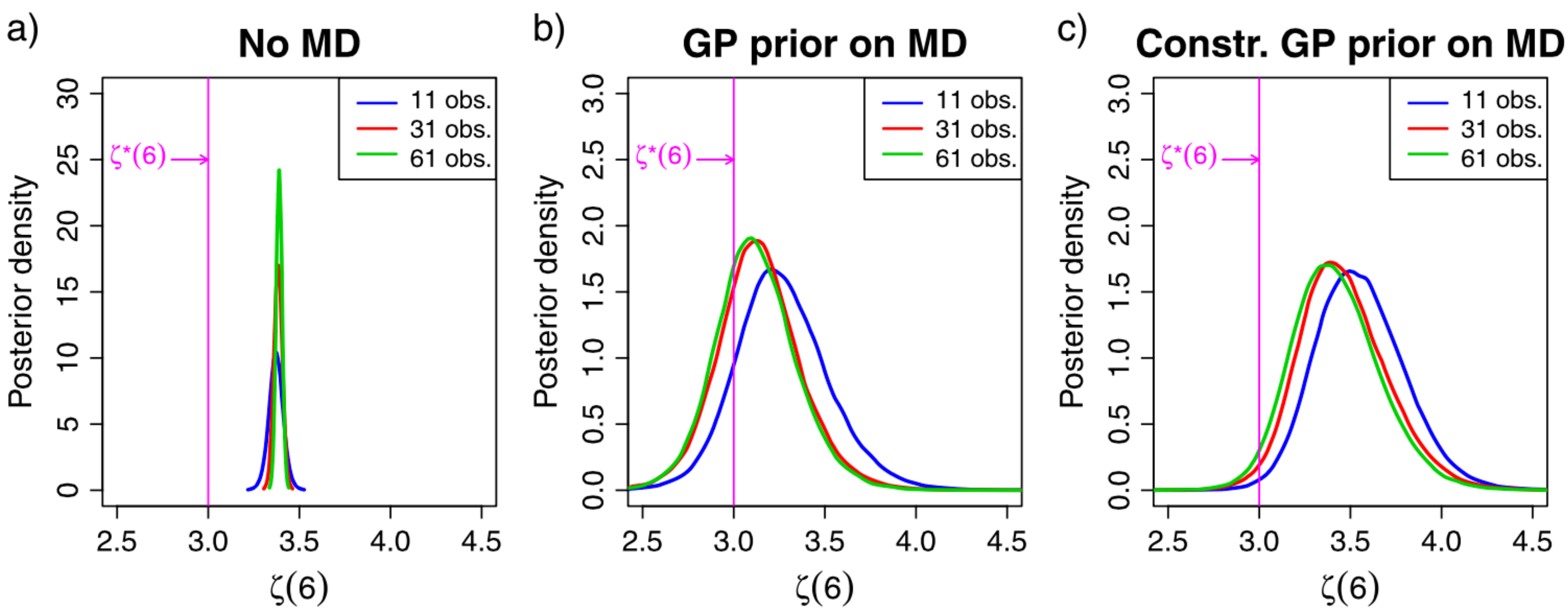
     $$p(S|\mathbf{m}, \mathbf{d}_{obs}) \propto S^{\frac{k_e}{2}+\alpha-1} exp(-S[E(\mathbf{m}) + \beta])$$

     which results in a gamma distribution of parameters $\frac{k_e}{2} + \alpha$ and $E(\mathbf{m}) + \beta$ where k_e is the effective degrees of freedom determined earlier.
   - Repeat steps 1 and 2 several times until convergence is achieved.

Example about discrepancy term using Gaussian Process Modeling

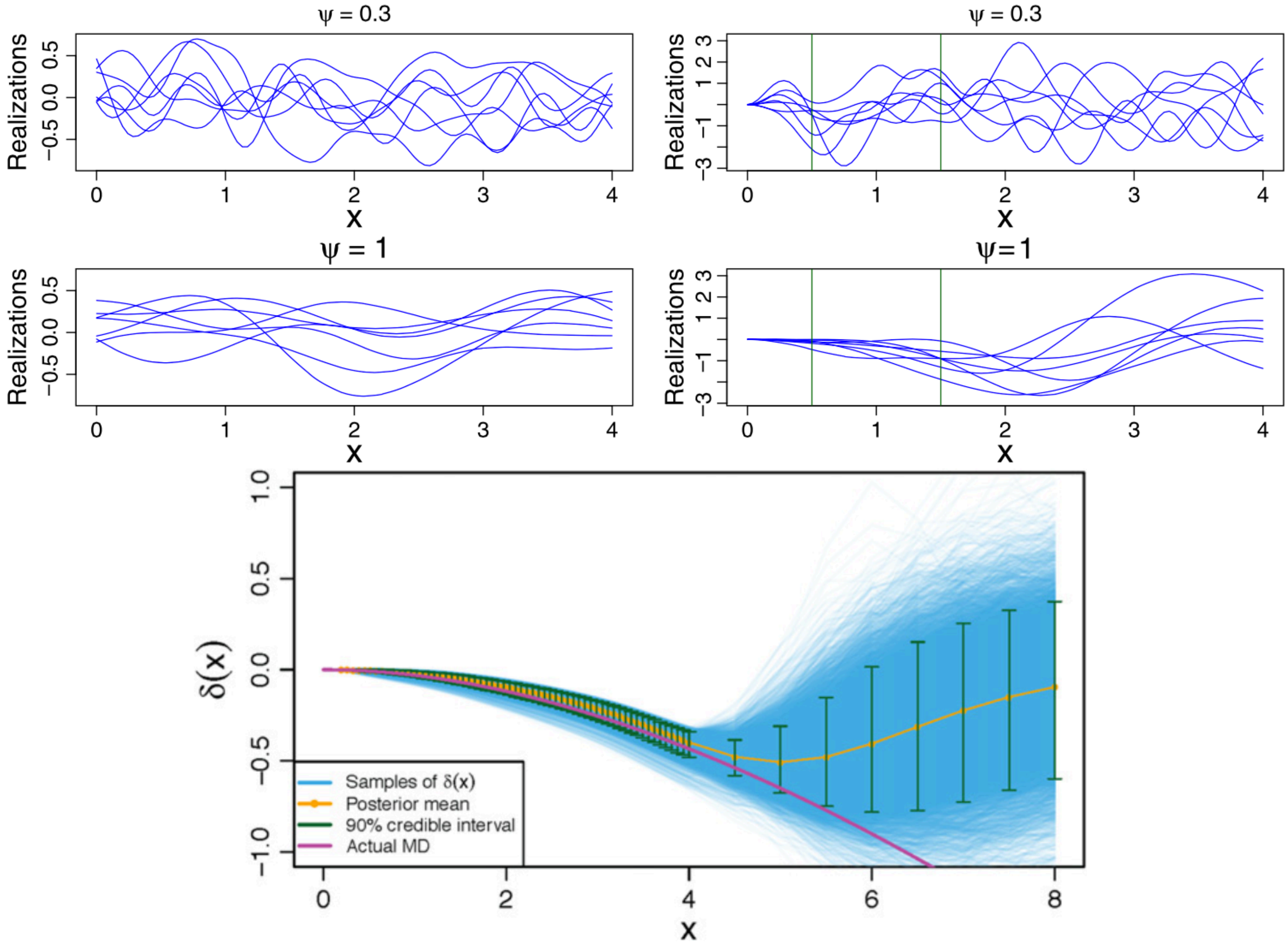


Brynjarsdottir and O'Hagan (2014) "Learning about physical parameters: The importance of model discrepancy"



Predictions of True process at x=6 (extrapolation)

Prior samples of discrepancy using Gaussian Process



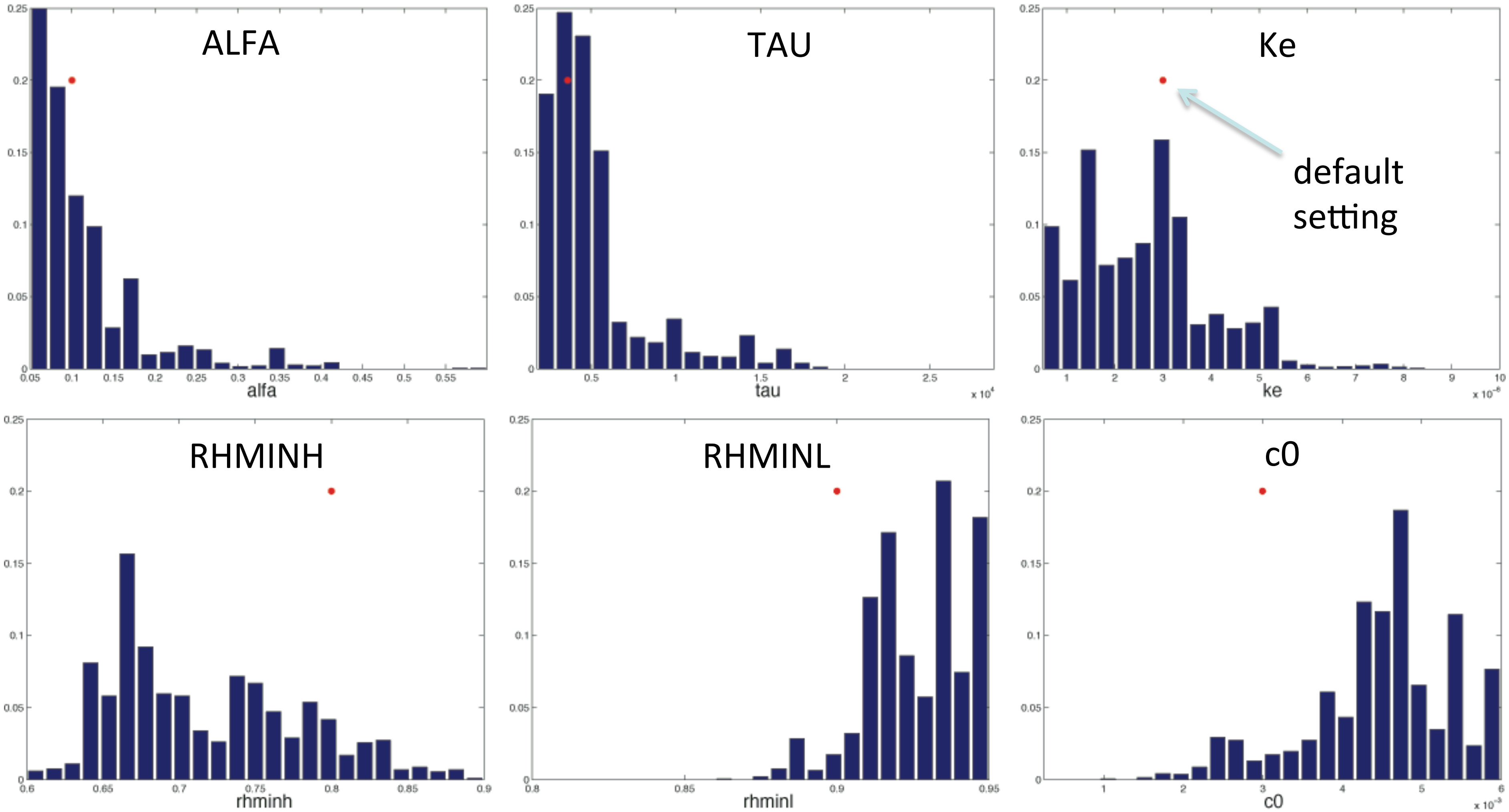
"In order to obtain a realistic extrapolation we need realistic prior information about $\delta\,(\,\cdot\,)$, both in the range of the data and out to the control variable values that we wish to predict."

Brynjarsdottir and O'Hagan (2014)

Calibration of CAM3.1

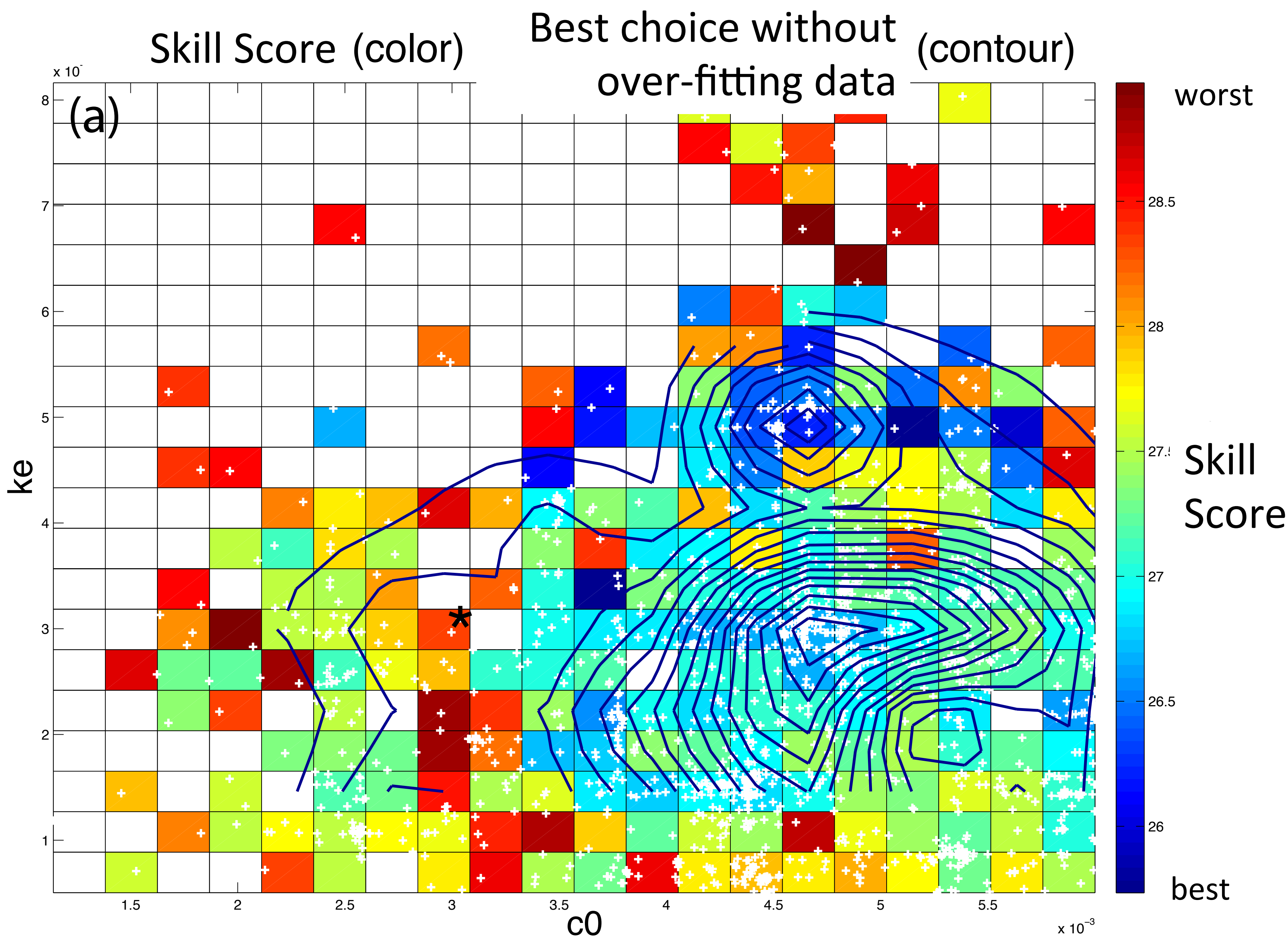
- Jackson et al. 2007 and Jackson 2009
- Despite significant “irreducible errors” get reasonable calibration results.
  - Model can not get all observations well at the same time.
  - Region calibration selected was “in the middle” of competing constraints and similar to where experts had wanted parameters values to be.

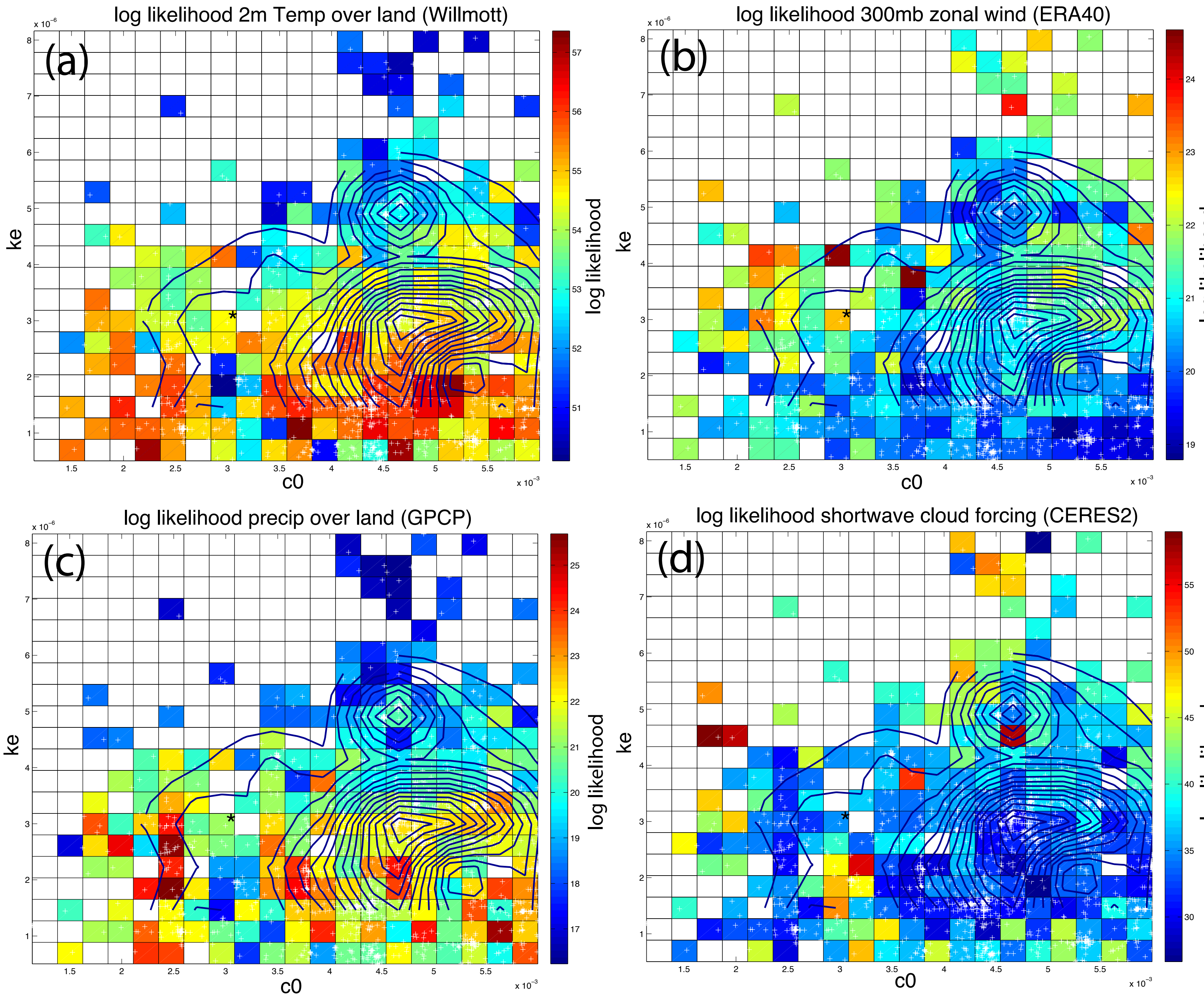
Calibration of 15 parameters of CAM3.1 using 11 observational products.



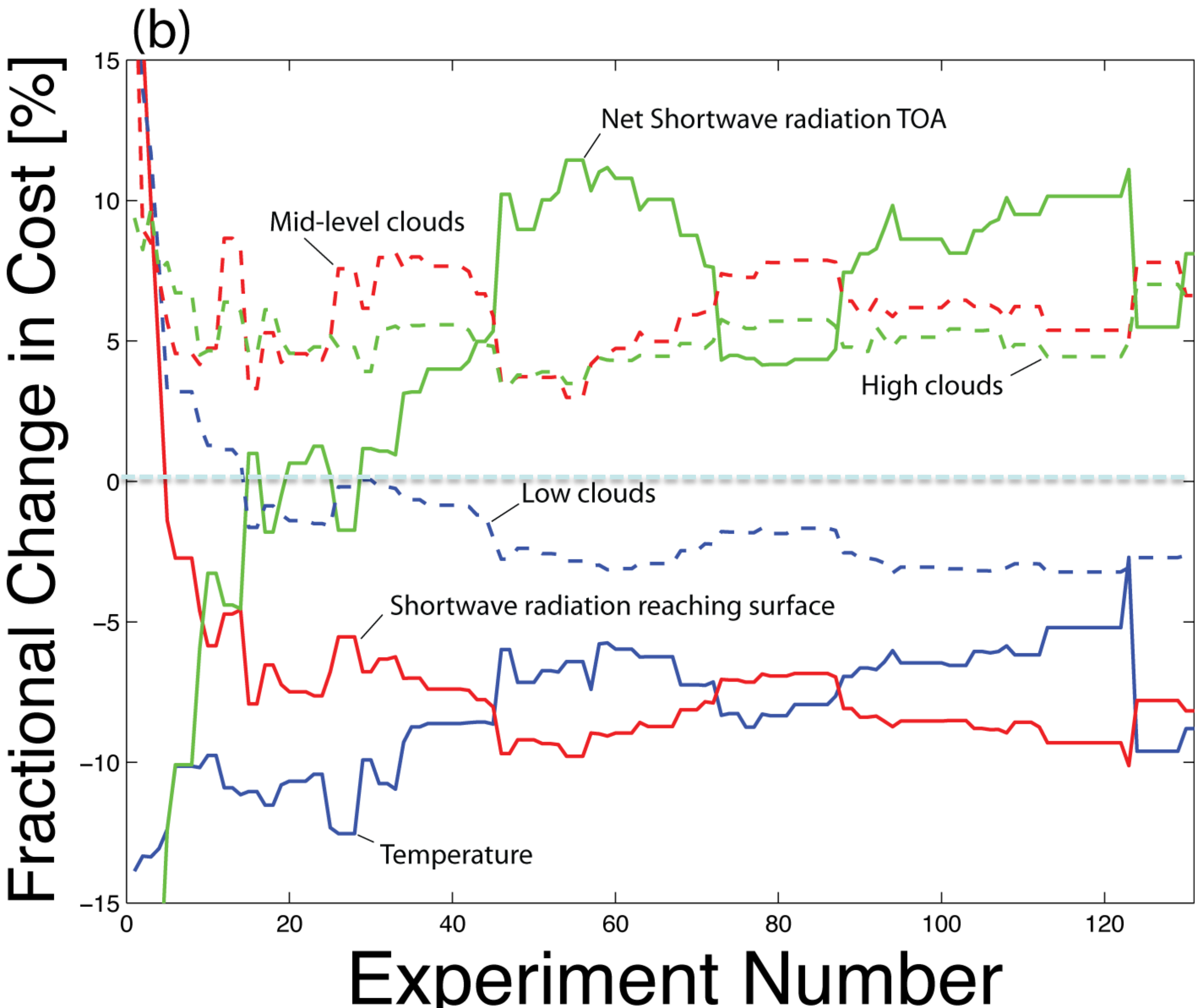
Calibration selects reasonable parameter values

Result summary of how 2 model parameters affect model skill

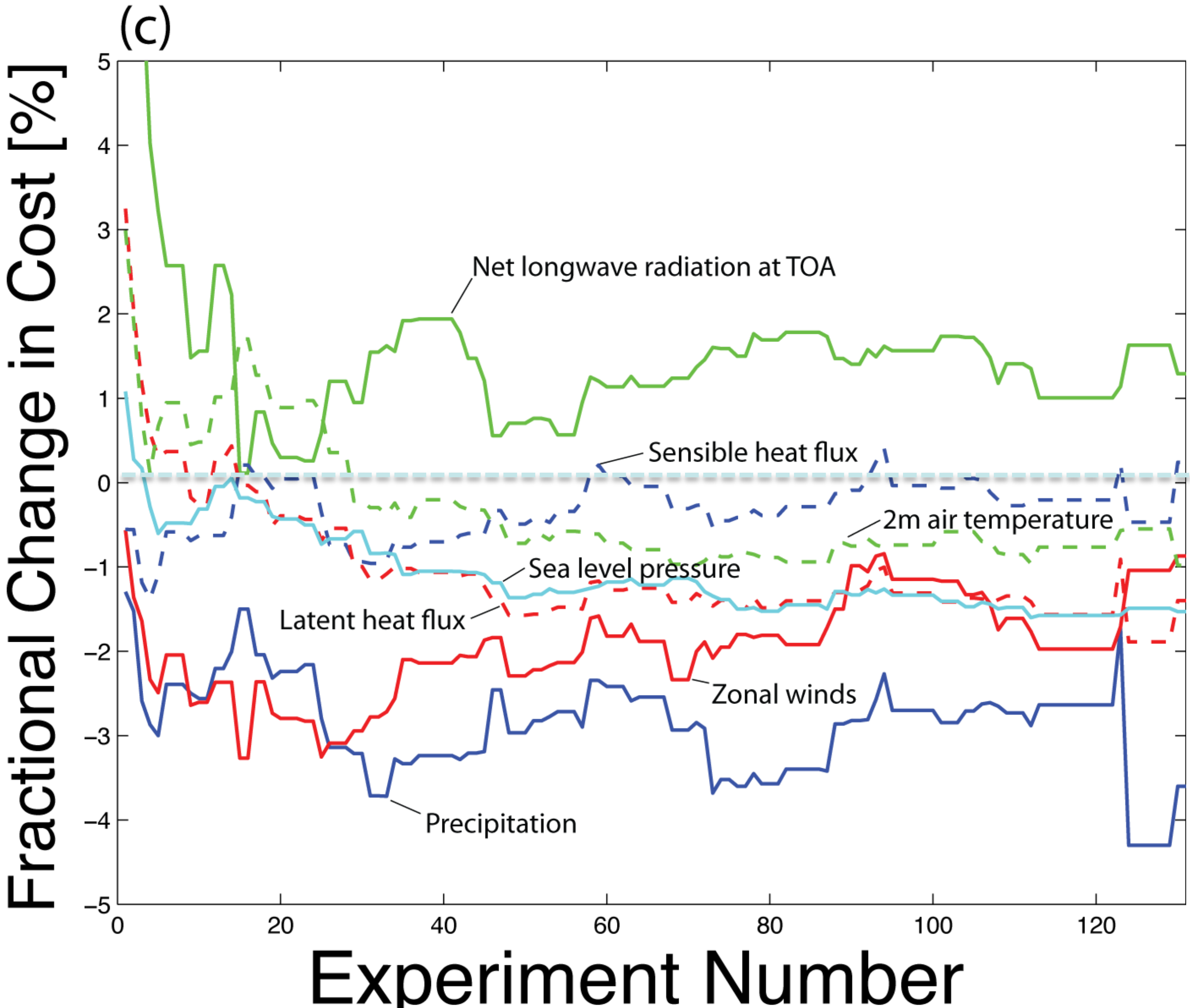




Jackson 2009



Jackson 2009



Jackson 2009

Irreducible error summary

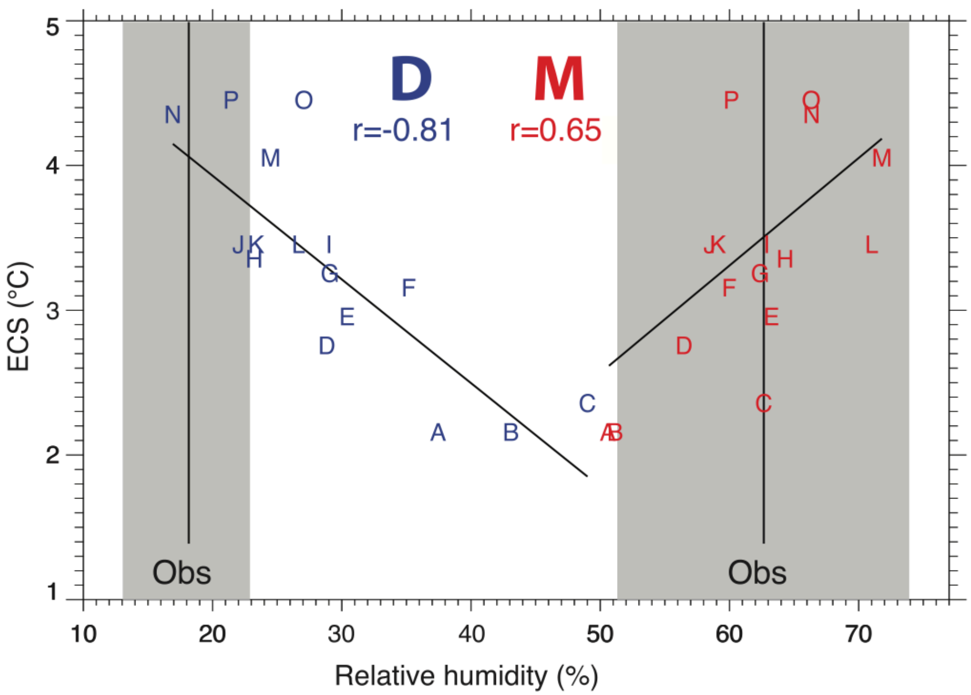
- Major problem for us.
- Seems hopeless to formulate problem with a discrepancy term that predicts how errors evolve with future climate.
- However calibration examples seem to produce reasonable results relative to expert opinion.
- Clearly an area of study that needs more attention.

Emergent constraints

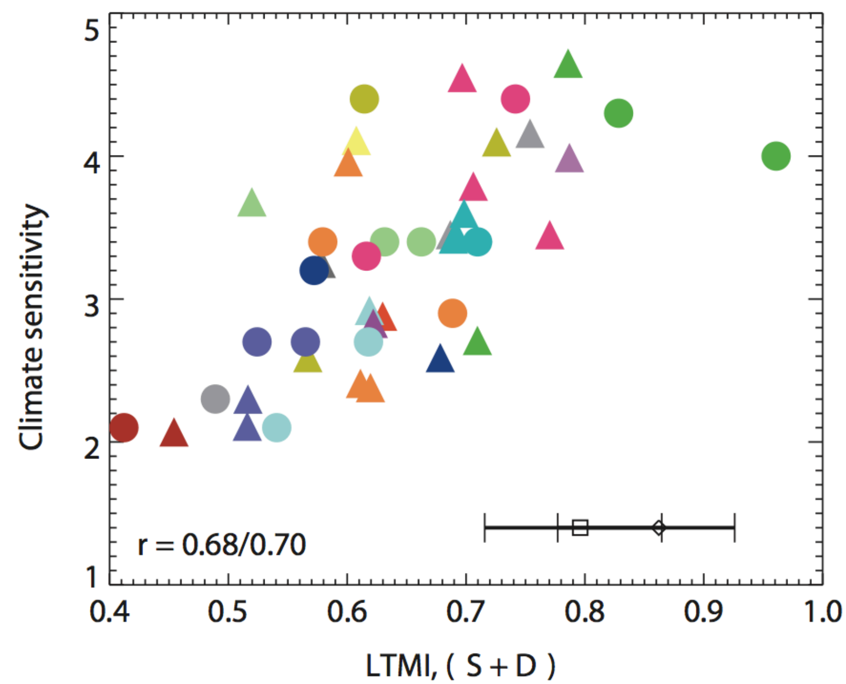
- An emergent constraint can be used to estimate errors in climate predictions using only information about the errors in simulating modern climate.
- Important as a way of establishing credibility of climate model projection information.
- Not clear that we have found any (see next slide)

Emergent constraints of CMIP Archive

a. Fasullo and Trenberth 2012



b. Sherwood et al., 2014



It has been extremely difficult to identify what observables matter to climate sensitivity.

Single model predictors

Use Bayesian inference and stochastic sampling to identify plausible alternatives to the standard CAM3.1 configuration. Determine what observables predict its sensitivity. Try predictors on CMIP5 archive.

Bayesian expression for observational constraints on parameter value selection

$$PPD(\mathbf{m} \mid \mathbf{d}_{obs}, g(\mathbf{m})) \propto \exp\left[-\frac{1}{2}\left(g(\mathbf{m}) - \mathbf{d}_{obs}\right)^T \mathbf{C}_{noise}^{-1}\left(g(\mathbf{m}) - \mathbf{d}_{obs}\right)\right] \cdot prior(\mathbf{m})$$

Likelihood test statistic

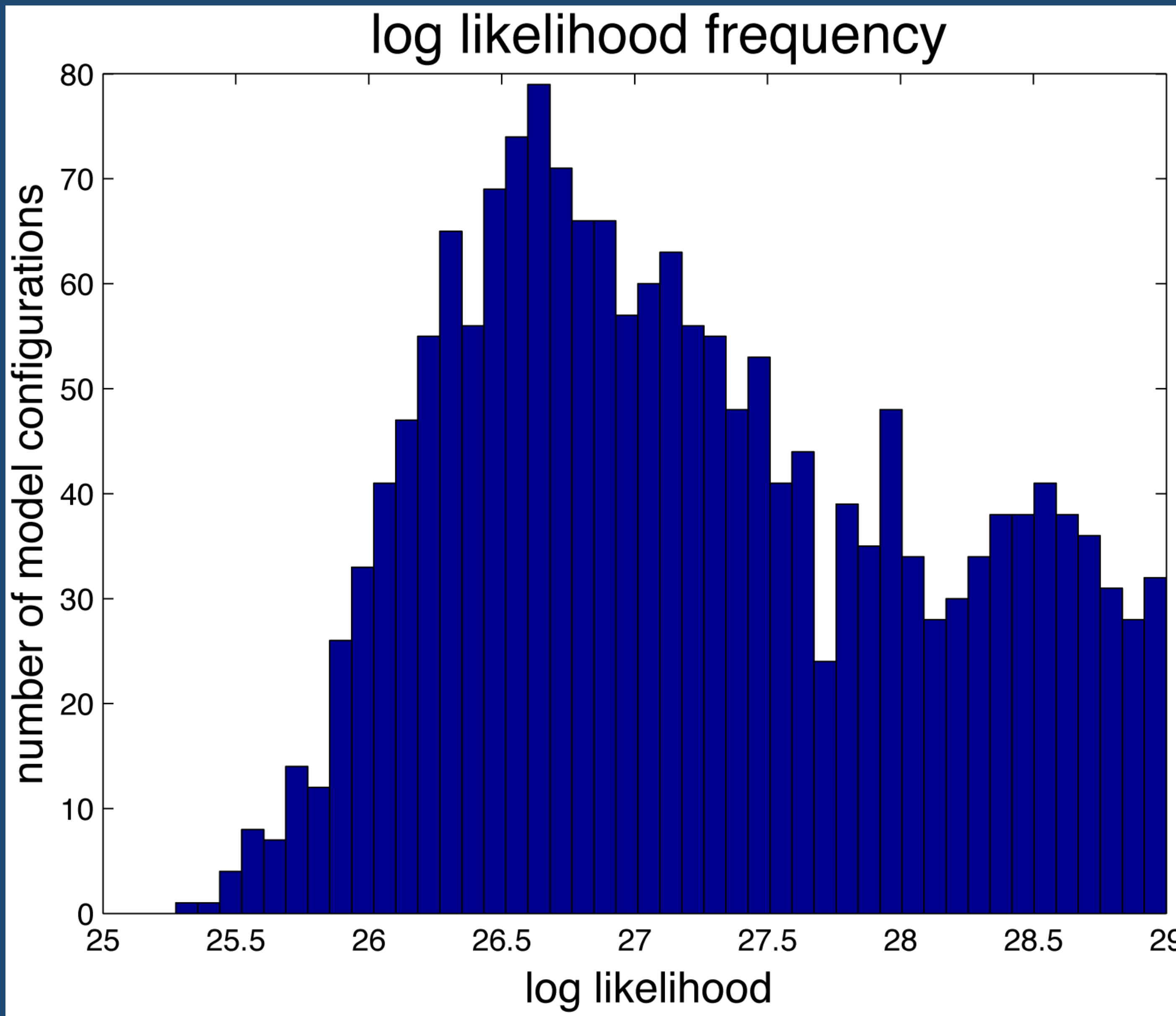
Generate an ensemble of plausible CAM3.1 parameter sets

- Select 15 parameters important to clouds, convection, and radiation in CAM3.1
- Use Markov Chain Monte Carlo sampling to estimate uncertainty distributions
  - Sampling strategy: Multiple Very Fast Simulated Annealing (MVFSA) (Jackson et al., 2004)
  - 3336 4-year long integrations to estimate 15 dimensional joint probability distribution
- Select 180 ensemble members that represent uncertainty
- Couple CAM3.1 to a slab ocean model and estimate response to 2x CO2

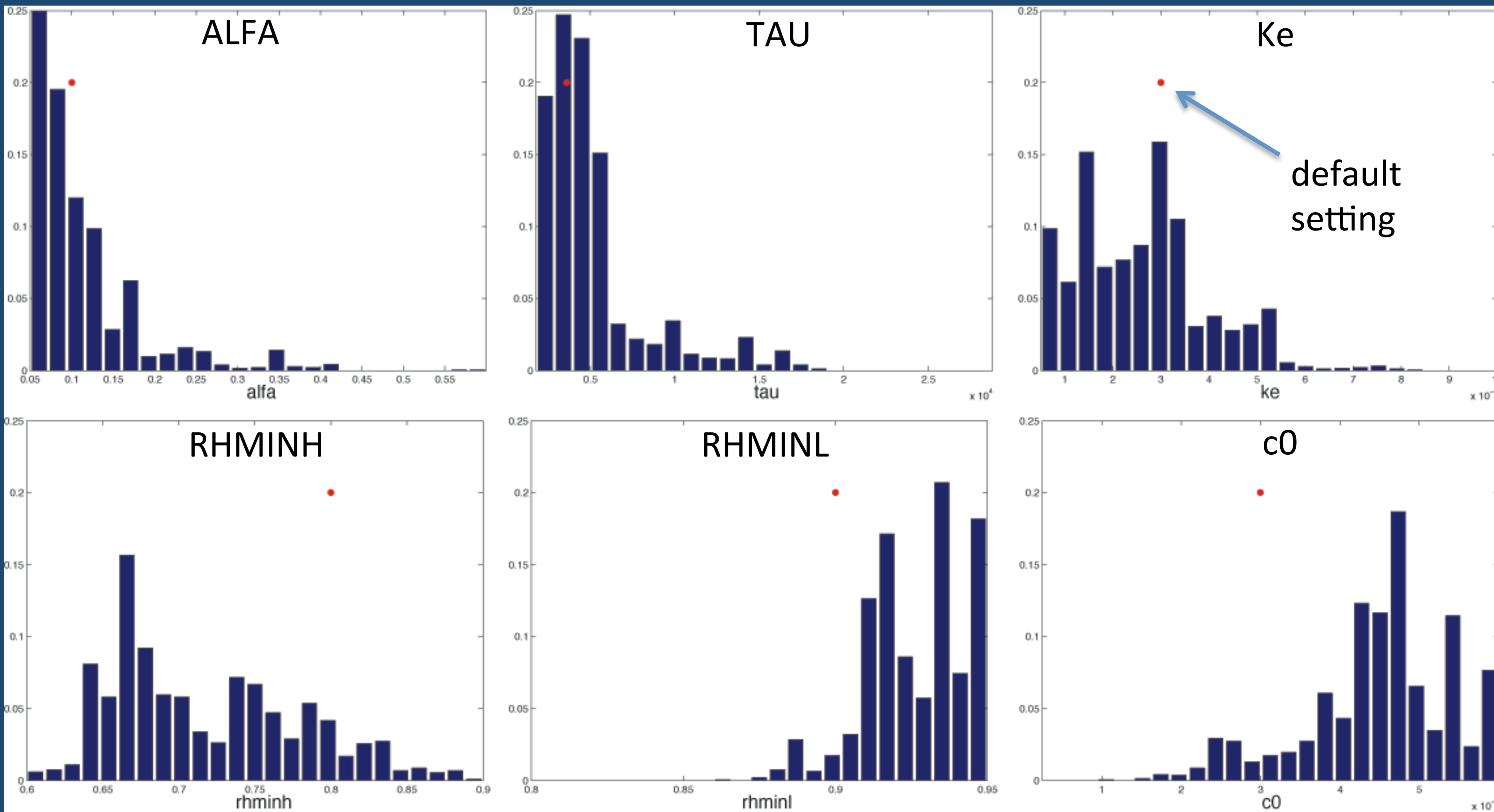
NCAR Top "10 +" observational constraints

(Seasonal means, 30S to 30N, unless noted otherwise)

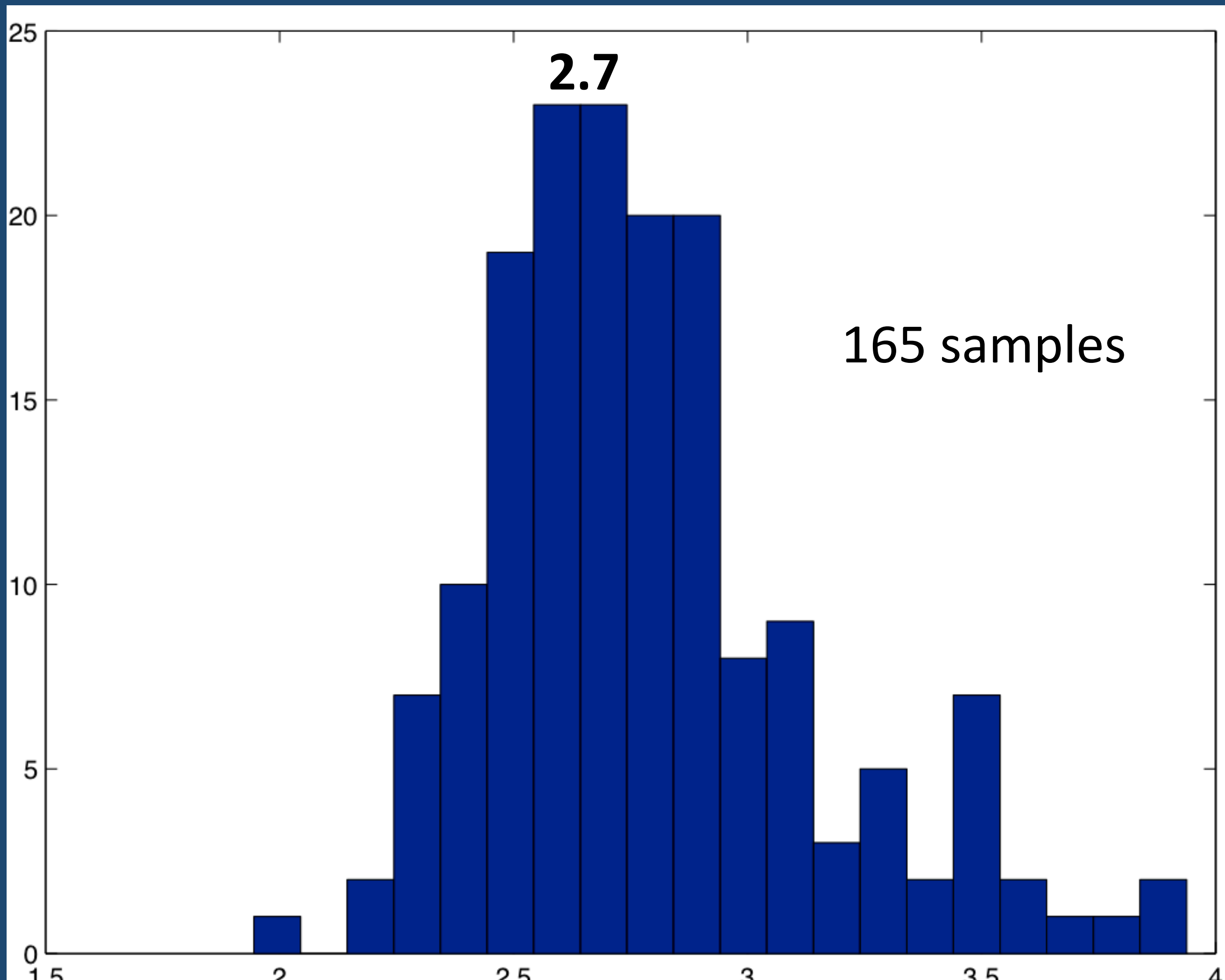
1. Land 2-m air temperature (Willmott)
2. Vertically averaged air temperature (ERA40)
3. Latent heat fluxes over ocean (WHOI)
4. Zonal winds at 300 mb (ERA40)
5. Longwave cloud forcing (CERES2)
6. Shortwave cloud forcing (CERES2)
7. Precipitation over land (GPCP)
8. Precipitation over ocean (GPCP)
9. Sea level pressure (ERA40)
10. Vertically averaged relative humidity (ERA40)
11. Global mean annual mean radiative balance (= 0.5 W/m^2)
12. Pacific ocean wind stress along equator


log likelihood frequency
number of model configurations
log likelihood
0
10
20
30
40
50
60
70
80
25
25.5
26
26.5
27
27.5
28
28.5
29

parameter distributions



Equilibrium Climate Sensitivity

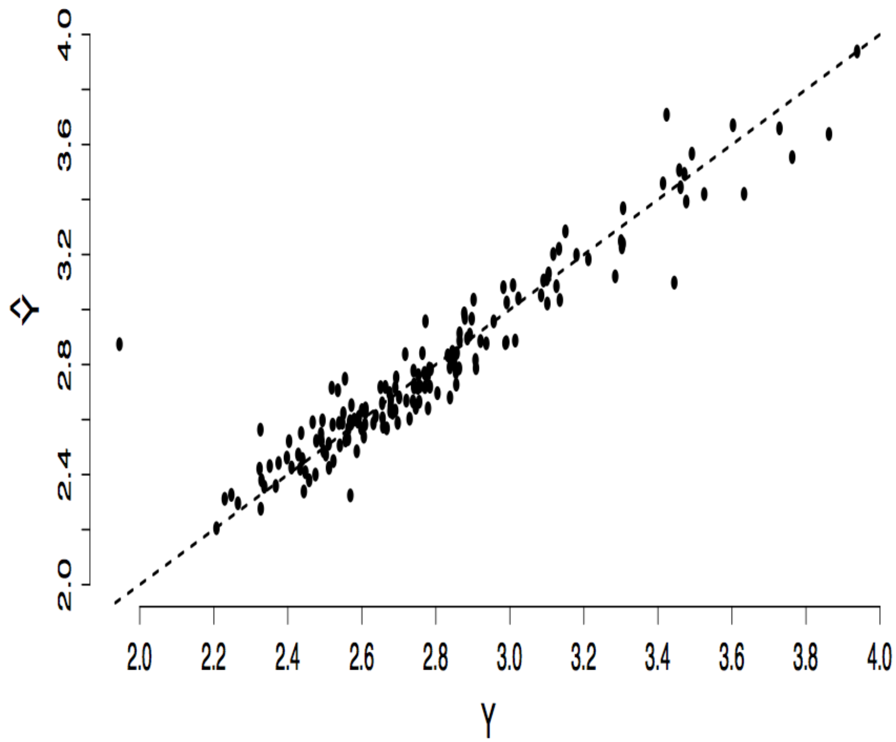


Number of experiments

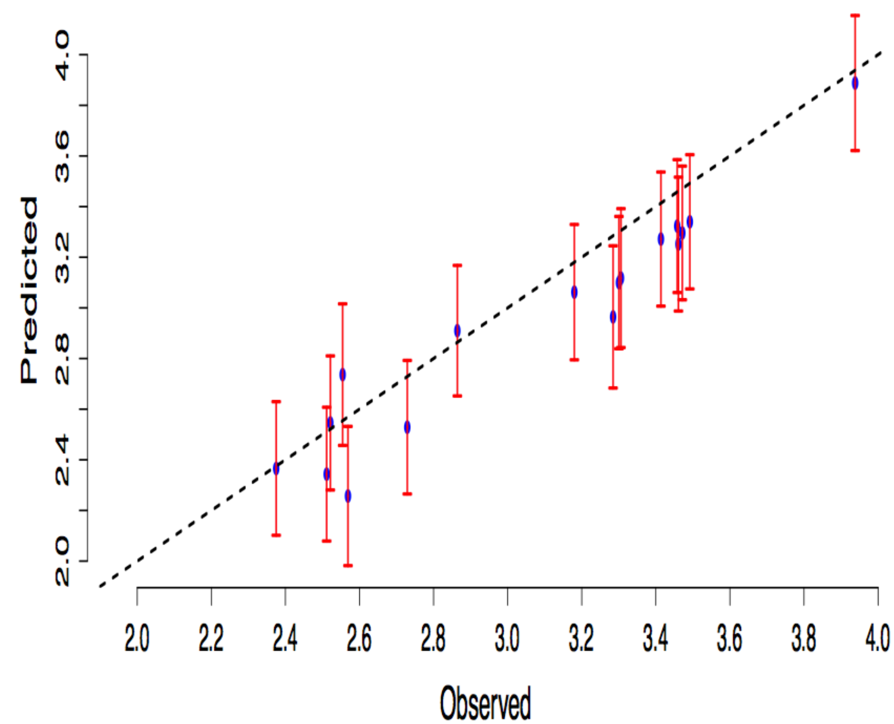
Global Mean Temperature after doubling CO2

Hattab et al. 2015 (in prep) Created a regression model to identify what errors matter to CAM3.1 predictions.

Climate Sensitivity

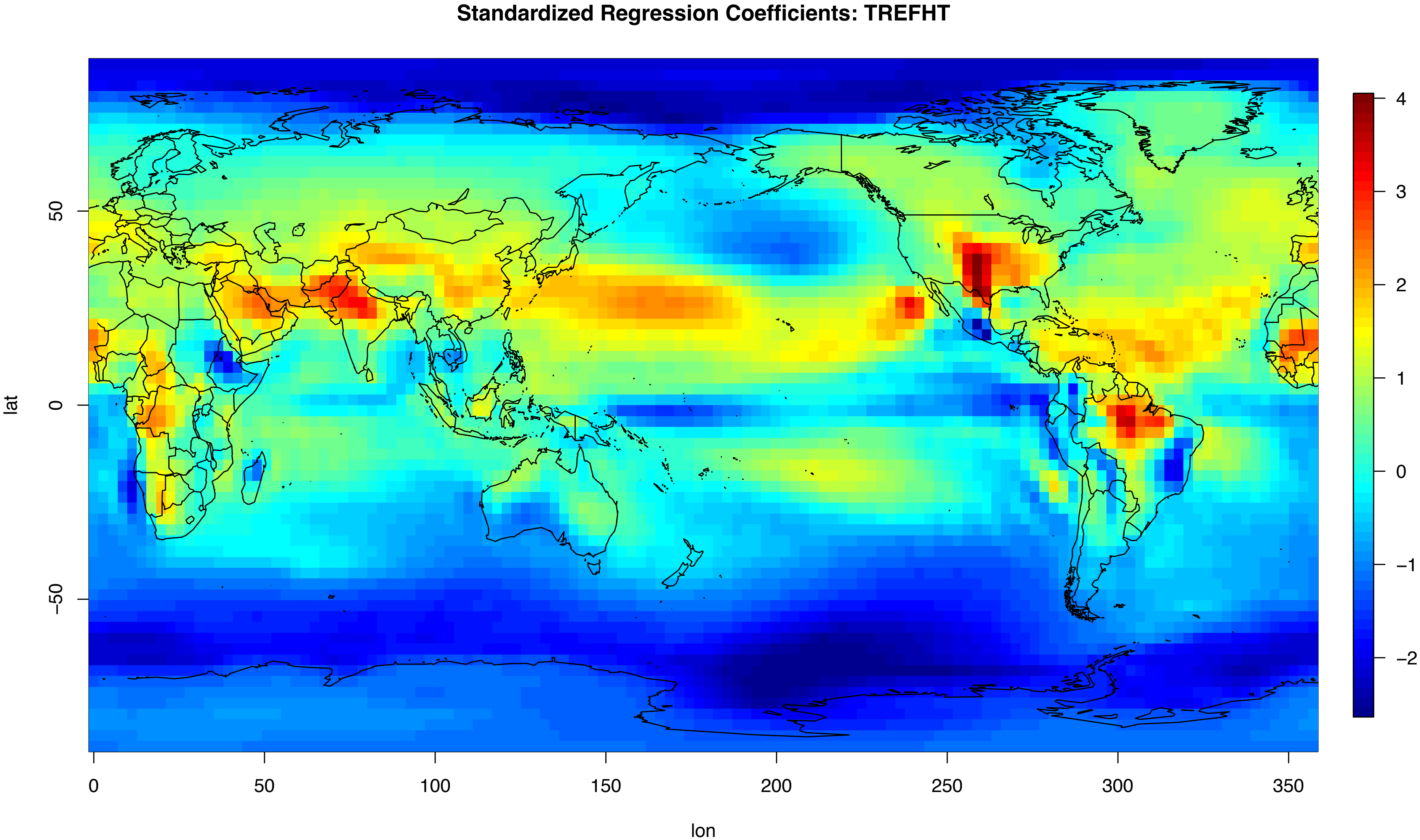


(b) Actual Vs. Fitted

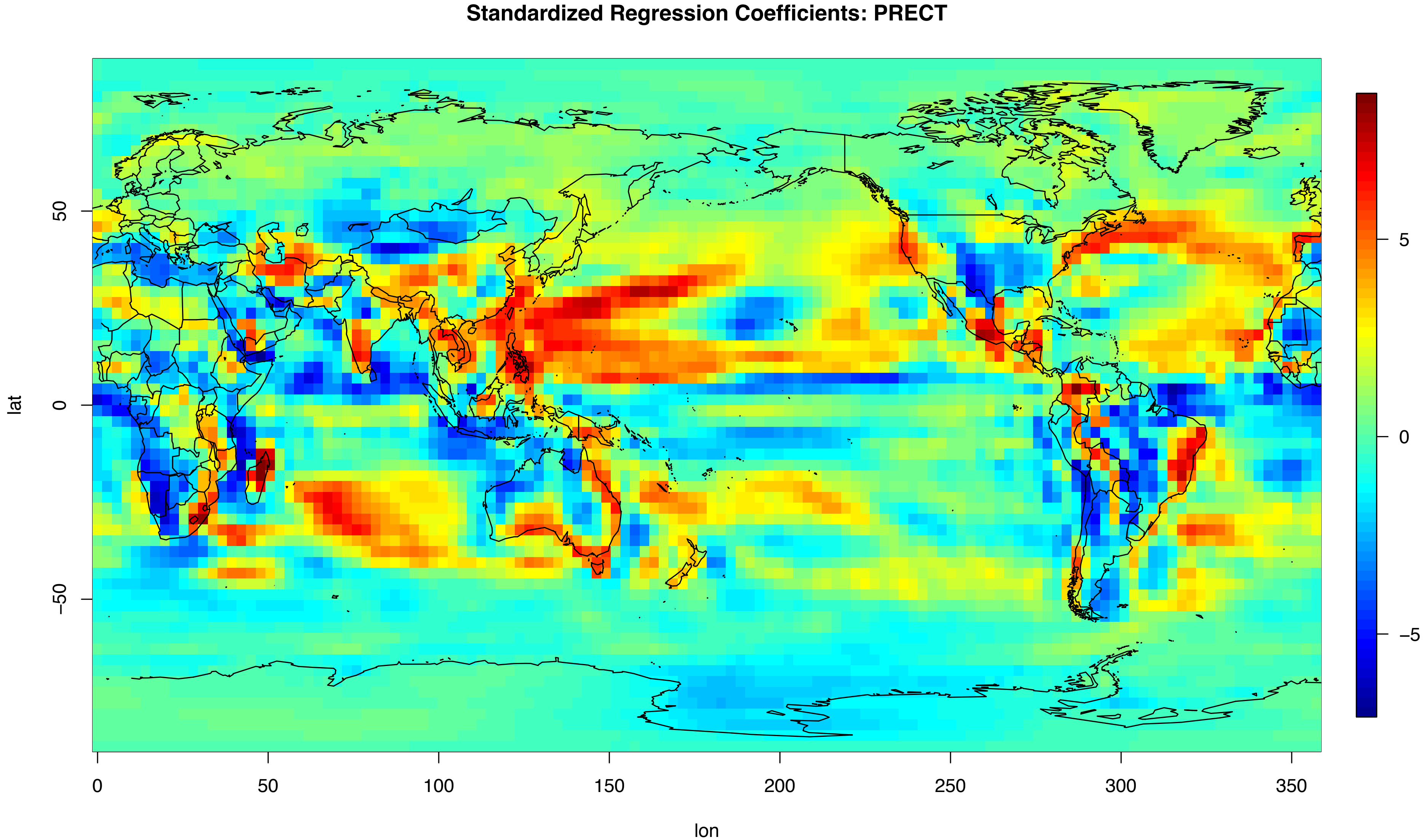


(c) Validation Predictions

β predictor maps; $\mathbf{Y} = \mathbf{X}\beta + e_n$



2 meter air temperature

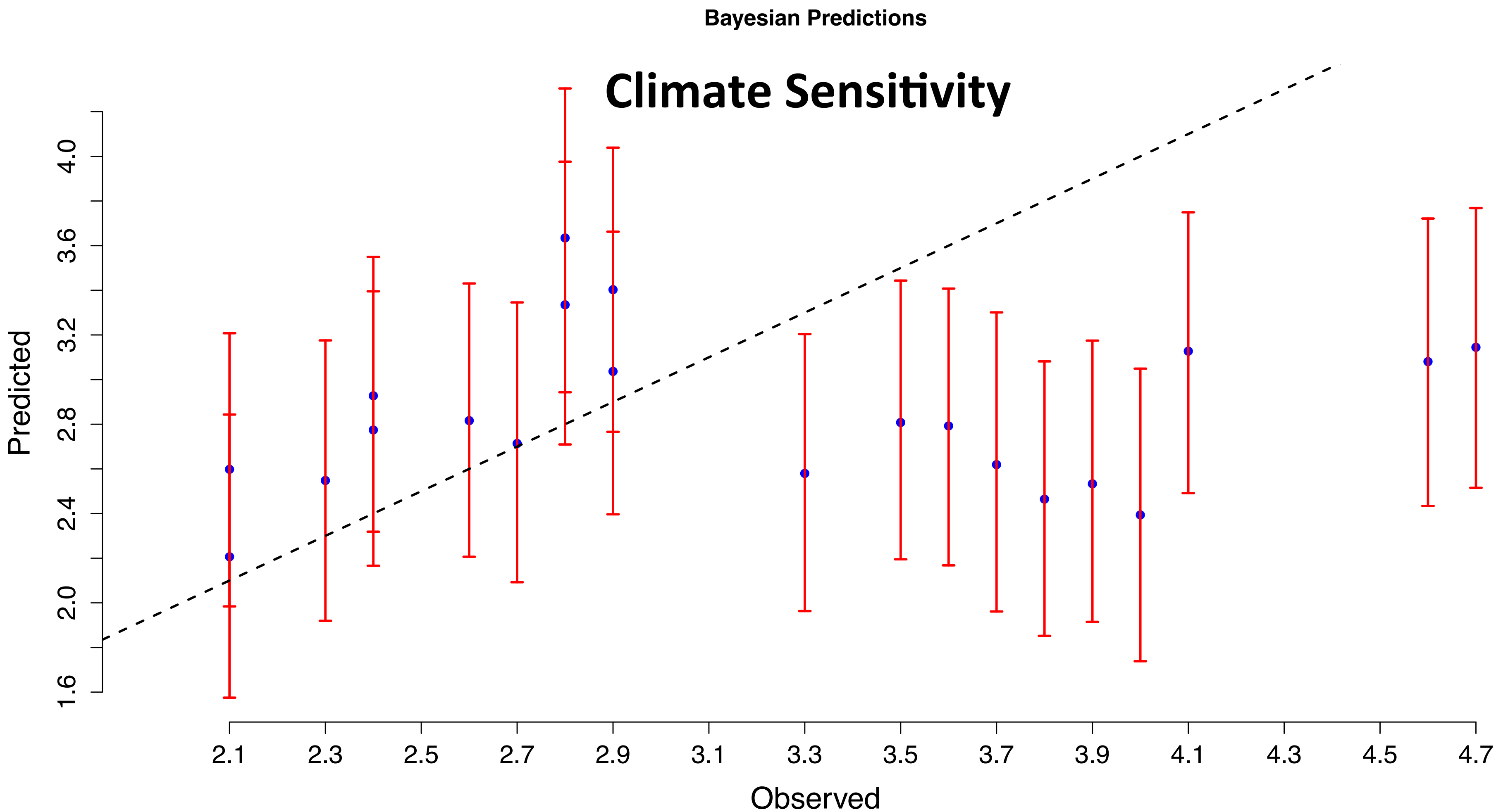


Precipitation

Spectacular failure

Use maps to predict climate sensitivity of CMIP5 archive.

Regression model prediction of CMIP5 archive



Interpretation of failure

- CAM3.1/slab ocean model is not like other models within CMIP5 archive.
- Parameter perturbations within CAM3.1 do not create structures that can be useful to predict other models.

Questions?