Inference and Analysis of Climate Models via Bayesian Approaches

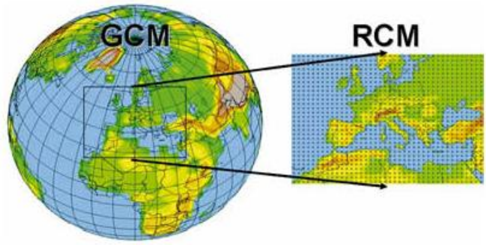
Gabriel Huerta

Department of Mathematics and Statistics
University of New Mexico
http://math.unm.edu

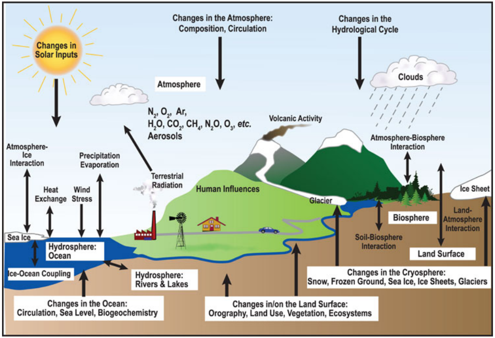
joint work with **Charles Jackson** (UT-Austin) **Deborah Khider** (USC, UT-Austin) **Mohammad Hattab** (UNM & VCU)

Climate Models

- Defined via differential equations (numerical model).
- *General Circulation Models* (GCM's)
  - cover the Earth. Grids boxes on scale of 100's kms.
- *Regional climate models* (RCM's)
  - resolve processes at smaller scale.



Graphical representation of a climate model



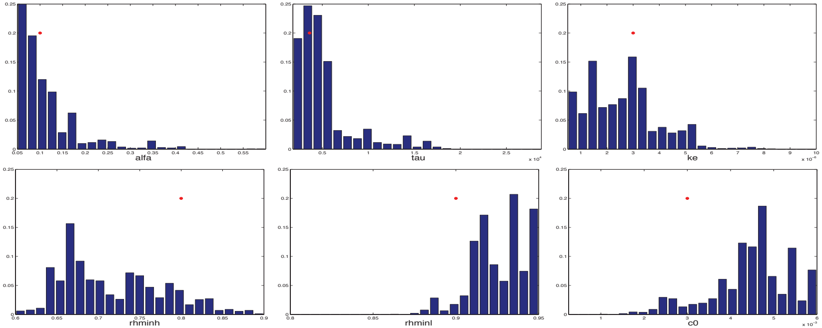
- Complex physical system!
- Requires high performance computing.

Calibration of Parameters for NCAR Community Atmosphere Model v. 3.1. [1]

Parameter	Definition	Value Ranges
RHMINL [%/100]	Low cloud critical relative humidity	0.80 – 0.95
RHMINH [%/100]	High cloud critical relative humidity	0.60 – 0.90
ALFA [fraction]	Initial cloud downdraft mass flux	0.05 – 0.60
TAU [hours]	Consumption rate of CAPE	0.5 – 8.0
ke $[(kg\ m^{-2}\ s^{-1})^{-1/2} s^{-1}]$	Environmental air entrainment rate	3.0e-6 – 10.0e-6
c0 $[m^{-1}]$	Precipitation efficiency	3.0e-3 – 6.0e-3

[1] C.S. Jackson et al Error Reduction and Convergence in Climate Prediction. Journal of Climate. Vol. 21 (2008), 24, pp. 6698-6709.

Samples for 6 climate parameters



- ≈ 1500 experiments (or runs).
- (*) denotes *default values*.

The climate modeler's formulation

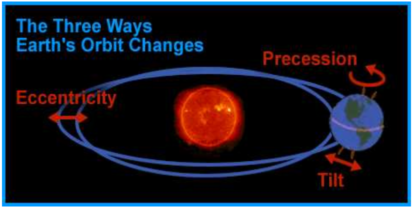
- Approximate

$$\text{posterior}(m) = \frac{\exp[-0.5 * E(m)] \times \text{prior}(m)}{\int exp[-0.5 * E(m)] \times \text{prior}(m)\ dm}$$

- $E(m)$ is a metric of *model skill* (cost function).
- m are *some climate parameters* .
- $E(m)$ considers observations (d_{obs}) and *model runs* $g(m)$.
- Stochastic sampling based on *optimization*.

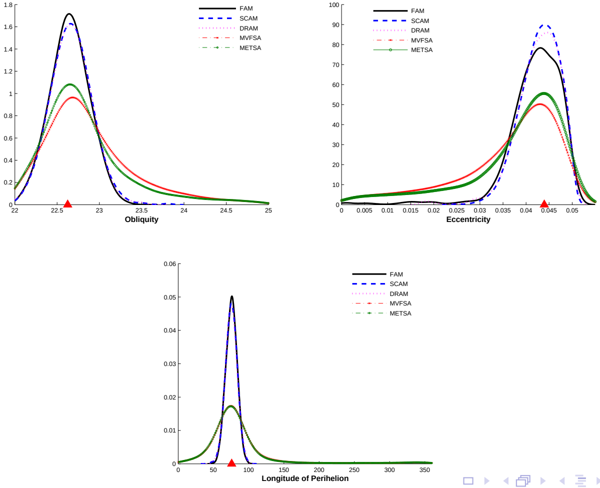
Surrogate climate model [2]

- **Response:** surface air temperature anomalies.
- Obliquity (Φ'): Earth's axial tilt.
- Eccentricity (e): How elliptical is the Earth's orbit around the Sun.
- Longitud of Perihelion: (λ) Point of closest approach to Sun.



[2] Computational methods for parameter estimation in climate models. A. Villagran et al. Bayesian Anal. Vol. 3, No. 4 (2008), pp. 823-850.

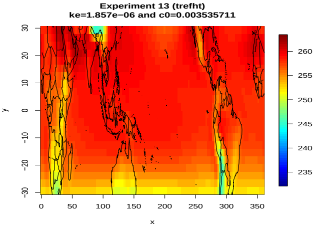
Posterior probability distributions for $m = (\Phi', e, \lambda)$

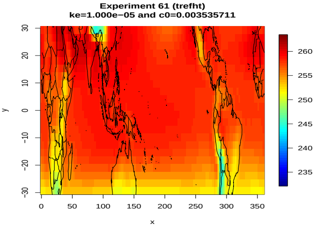


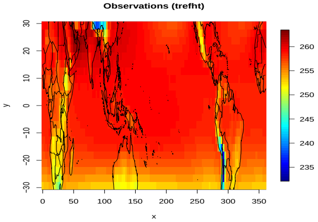
NCAR-CAM 3.1:Output

- Model runs of climate model simulation with physical observations.
- **Fields** (regressors):
  - *2-m air temperature* 'field T' or TREFHT.
  - Other outputs : *shortwave cloud forcing (SWCF)* , *precipitation over ocean (PRECT)* , *longwave cloud forcing (LWCF)* , *longwave flux at TOA*, *shortwave flux at TOA*, *zonal averaged relative humidity*, *latent heat flux over ocean*, etc.
- 7 spatial fields on a grid of 128 $\times$ 64.
- 4 spatial fields on a grid of 64 $\times$ 26.

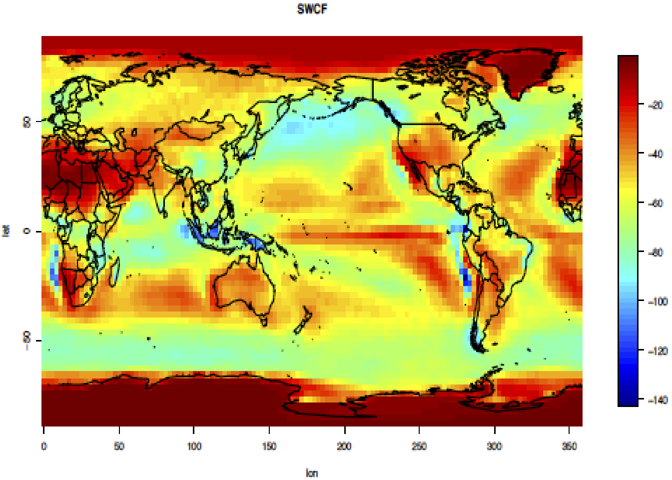
Model runs for field ”TREFHT”





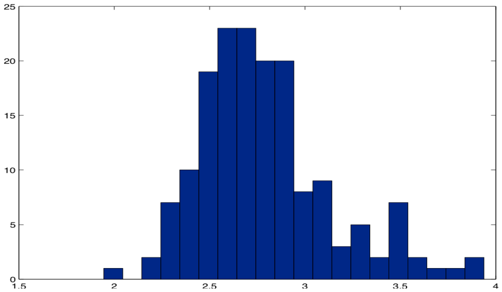


Output run for Shortwave Cloud Forcing



Global warming experiments

- *Coupled* CAM 3.1 to a ”slab” ocean.
- 165 experiments (runs) were performed
- Types: *Control* and $2 \times CO2$.
- *Climate sensitivity*: ”change in global mean temperature after doubling CO2”.



Relating Sensitivity to Climate Fields [3]

- Regression problem with $n = 165$ and $p = 63872$.
- $\boldsymbol{Y} = \boldsymbol{X}\beta + \epsilon_n$.
- $\boldsymbol{X}$ given by fields.
- β represents predictors effects.
- $p >> n$.
- Predictions are also of main interest to predict other climate models (CMIP5).
- CMIP5 = list of climate models for other centers.

[3] M.H. Hattab et al. (2015) A regression between bias and climate sensitivity within a perturbed physics ensemble of CAM3.1

Approach

- Perform a *principal component regression* (PCR)

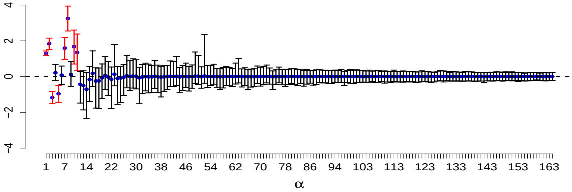
$$\boldsymbol{Y} = \boldsymbol{X}\beta + \epsilon_n = \boldsymbol{W}\alpha + \epsilon_n.$$

- $\boldsymbol{W}$ is a $n \times k$ matrix with the first k PCs as columns.
- From ”super saturated” to ”saturated model”.
- Frequentist and Bayesian framework.
- Estimating α.
- Mapping of α's back to β's.

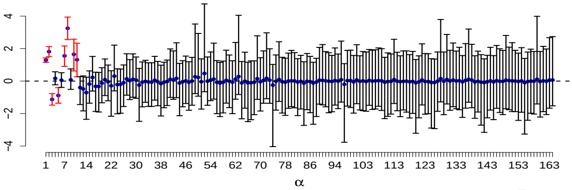
Bayesian Analysis

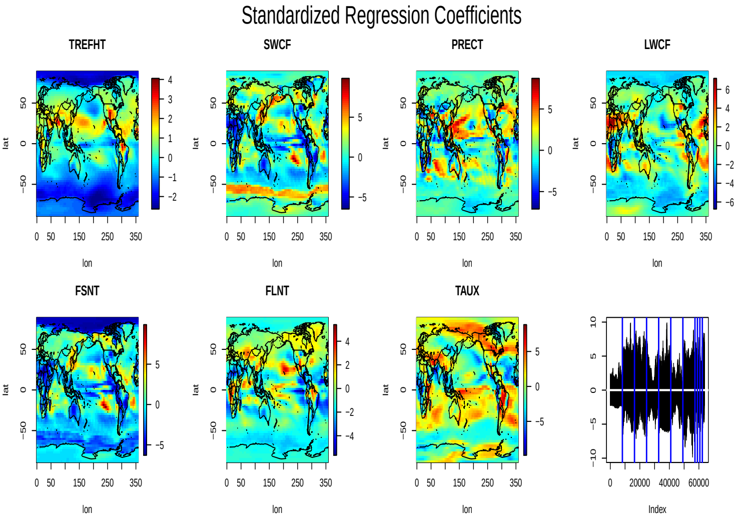
- Exploring *Bayesian* solution under two priors for $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_k)^t$.
- *Prior 1:* $\alpha_i \sim N\left(0, \frac{g(i^{-2})}{\phi_i}\right)$.
- *Prior 2:* $\alpha_i \sim N\left(0, \frac{g}{\phi_i}\right); i = 1, 2, \ldots, k$.
- In both scenarios,
  - $\phi_i \sim Gamma(\delta, \delta), i = 1, 2, \ldots, k$.
  - $g \sim Unif(0, a)$.
- Main results not sensitive to prior specification.

Posterior intervals for α_i under two priors.

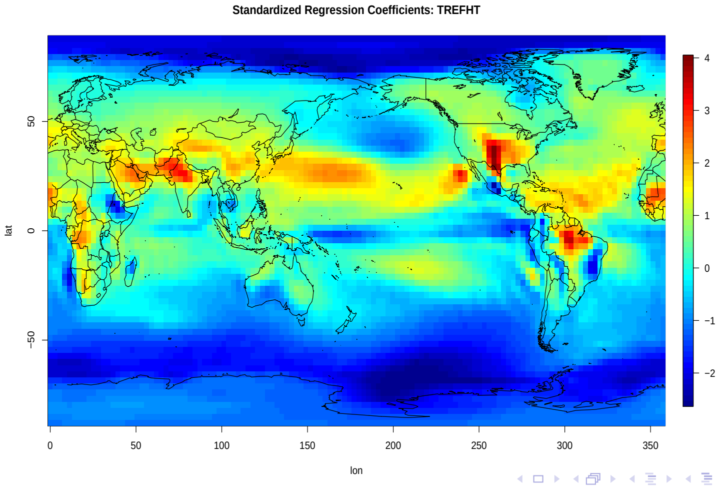


(a) M_1

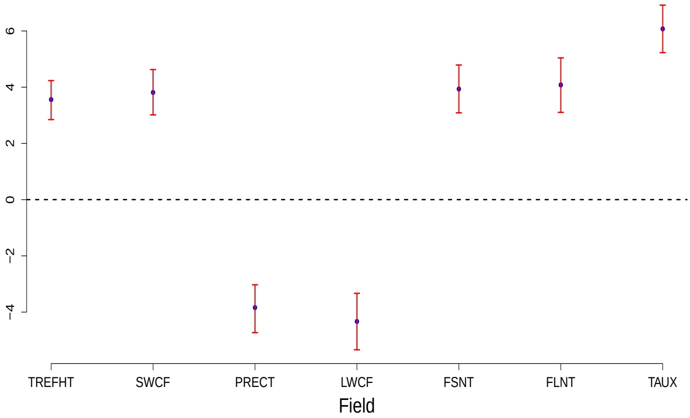



Standardized Regression Coefficients
TREFHT
SWCF
PRECT
LWCF
FSNT
FLNT
TAUX
lat
lon
Index

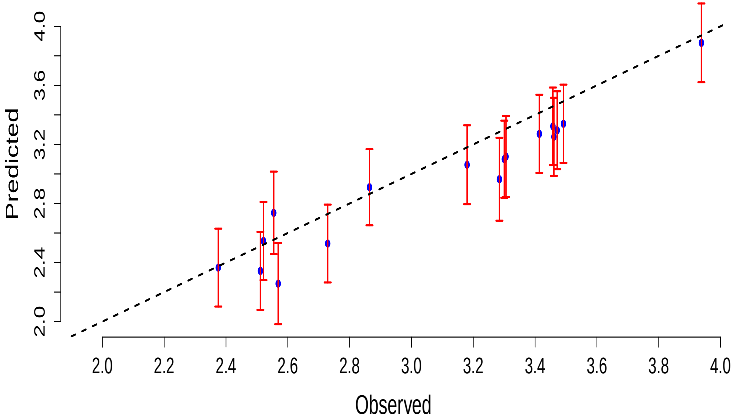
TREFHT β


Standardized Regression Coefficients: TREFHT
lat
50
0
−50
lon
0
50
100
150
200
250
300
350
4
3
2
1
0
−1
−2

Location (261.5625, 29.30136)



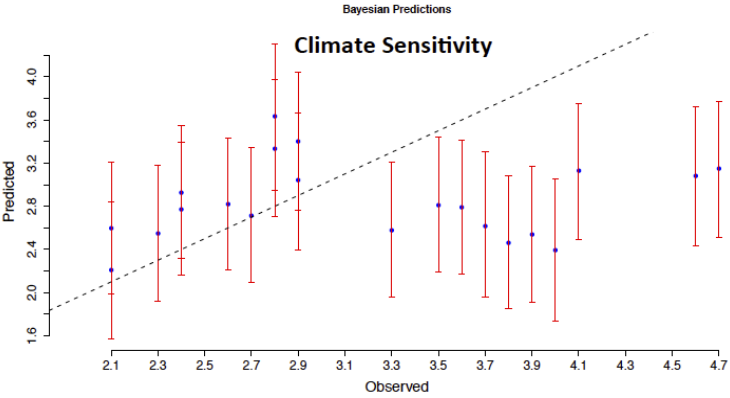
Predictions for CAM3.1



Sub-list of CMIP5 models

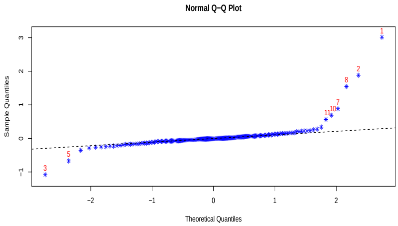
Modeling Center	Model	Institution	terms of use
BCC	BCC-CSM1.1 BCC-CSM1.1(m)	Beijing Climate Center, China Meteorological Administration	unrestricted
CCCma	CanAM4 CanCM4 CanESM2	Canadian Centre for Climate Modelling and Analysis	unrestricted
CMCC	CMCC-CESM CMCC-CM CMCC-CMS	Centro Euro-Mediterraneo per I Cambiamenti Climatici	unrestricted
CNRM-CERFACS	CNRM-CM5	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	unrestricted
CNRM-CERFACS	CNRM-CM5-2	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	unrestricted
COLA and NCEP	CFSv2-2011	Center for Ocean-Land-Atmosphere Studies and National Centers for Environmental Prediction	unrestricted
CSIRO-BOM	ACCESS1.0 ACCESS1.3	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	unrestricted
CSIRO-QCCCE	CSIRO-Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence	unrestricted
EC-EARTH	EC-EARTH	EC-EARTH consortium	unrestricted
FIO	FIO-ESM	The First Institute of Oceanography, SOA, China	unrestricted
GCESS	BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University	unrestricted
INM	INM-CM4	Institute for Numerical Mathematics	unrestricted

Regression model prediction of CMIP5 archive



Comments about work with GCMs

- PCR capable of estimating the ”field” coefficients for CAM3.1 and predictions.
- 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.



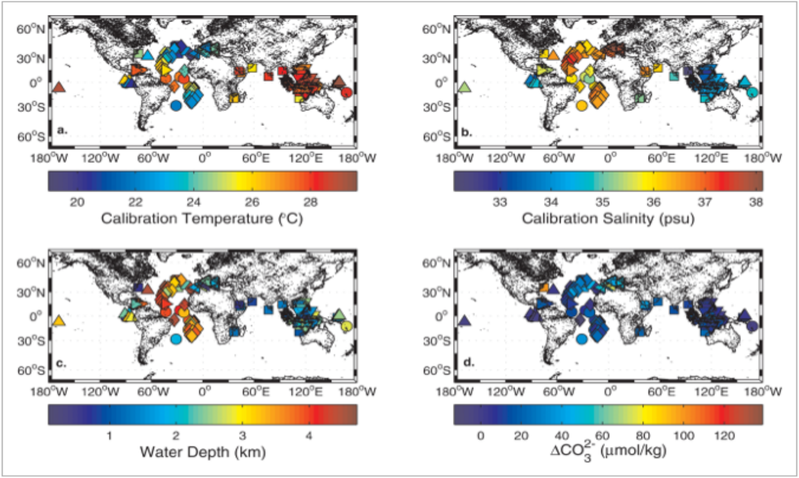
Bayesian Calibration of *Globigerinoides ruber* Mg/Ca [4]

- Estimate probability distributions of Mg/Ca sensitivity to SST, sea level salinity (SSS) and deep water ΔCO_3^{2-} (dissolution).
- Data set of 186 core top-samples with "global coverage".

$$\pi(unknows \mid data) \propto f(data \mid unknows) \cdot \pi(unknowns)$$

- $\pi(unknowns|data)$ "probability of unknowns given data" (posterior).
- $f(data|unknowns)$ "likelihood of the data given the unknowns".
- $\pi(unknowns)$ "prior probability of unknowns".

[4] D. Khider et al. accepted for $G-cubed$ (2015).



60°N
30°N
0°
30°S
60°S
a.
180°W 120°W 60°W 0° 60°E 120°E 180°W
20 22 24 26 28
Calibration Temperature (°C)
b.
33 34 35 36 37 38
Calibration Salinity (psu)
c.
1 2 3 4
Water Depth (km)
d.
0 20 40 60 80 100 120
ΔCO$_3^{2-}$ (µmol/kg)

Calibration equation for Mg/Ca

- A *piecewise regression:*

$$\text{if } \Delta CO_{3\ (i)}^{2-} \geq 21\,\mu mol/kg,\ Mg/Ca_{(i)} = \left(\exp\left(\alpha_1 T_{(i)} + \alpha_2 S_{(i)} + \alpha_0\right) + \alpha_3 21\right) \Big/ \left(1 + \alpha_4 C_{(i)}\right) + \varepsilon_{(i)}$$

$$\text{if } \Delta CO_{3\ (i)}^{2-} < 21\,\mu mol/kg,\ Mg/Ca_{(i)} = \left(\exp\left(\alpha_1 T_{(i)} + \alpha_2 S_{(i)} + \alpha_0\right) + \alpha_3 \Delta CO_{3\ (i)}^{2-}\right) \Big/ \left(1 + \alpha_4 C_{(i)}\right) + \varepsilon_{(i)}$$

$$\varepsilon_{(i)} \sim N\left(0, \tau^2\right)$$

- $\mathbf{\Phi} = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \tau^2)$. By Bayes theorem,

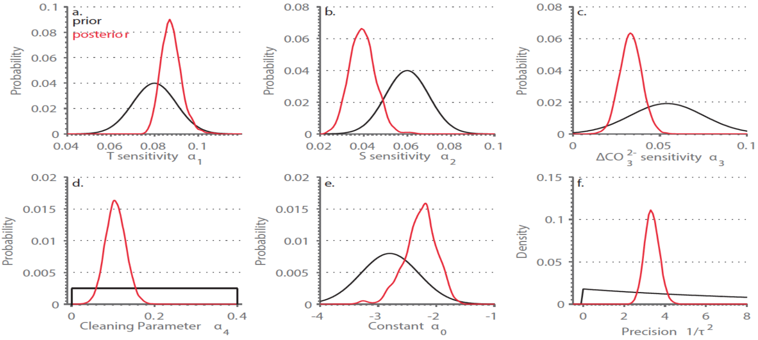
$$\pi\left(\Phi \middle| Mg/Ca, T, S, \Delta CO_3^{2-}, C\right) \propto f\left(Mg/Ca \middle| T, S, \Delta CO_3^{2-}, C, \Phi\right) \cdot \pi(\Phi)$$

Prior distribution on coefficients

- Based on "culturing experiments" and expert knowledge (Khider).
- *Apriori independent:*
  $\pi(\mathbf{\Phi}) = p(\alpha_0)p(\alpha_1)p(\alpha_2)p(\alpha_3)p(\alpha_4)p(\tau^2)$

$$\alpha_0 \sim N(-2.8, 0.5)$$
$$\alpha_1 \sim N(0.08, 0.01)$$
$$\alpha_2 \sim N(0.06, 0.01)$$
$$\alpha_3 \sim N(0.054, 0.019)$$
$$\alpha_4 \sim U(0, 0.4)$$
$$1/\tau^2 \sim Ga(1.0, 0.1)$$

Posterior and prior distribution for $\mathbf{\Phi}$.



- Required a large number of MCMC iterations.
- Implementation in *rjags* and *matjags*.

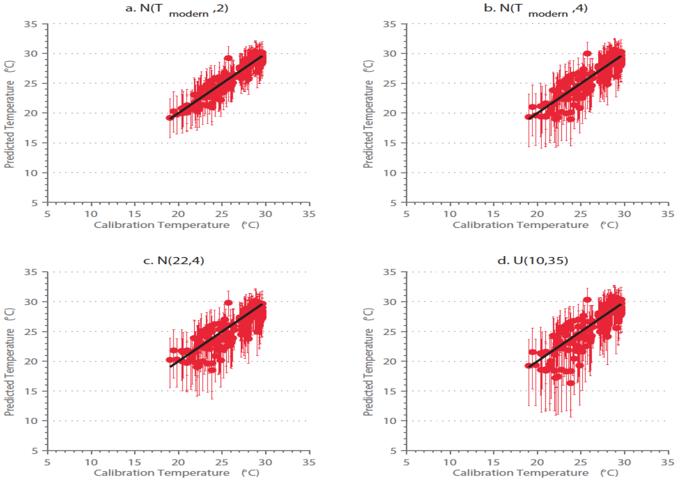
Predicitive framework

- Goal: provide predictions of (say) T.
- Just another application of Bayes theorem and MCMC!

$$\pi\left(T \middle| Mg/Ca, S, \Delta CO_3^{2-}, C, \Phi\right) \propto f\left(Mg/Ca \middle| T, S, \Delta CO_3^{2-}, C, \Phi\right) \cdot \pi(T)$$

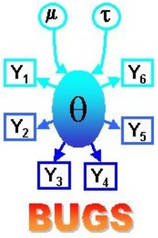
- Requires prior on T.
- Notice conditioning on $\mathbf{\Phi}$.
- Post-processing of posterior samples of $\mathbf{\Phi}$.
- Done also for ΔCO_3^{-2} an S.

Prediction results for T.



Openbugs

- Generic software to implement Bayesian models.
- Available from www.openbugs.net
- BUGS stands for *Bayesian Inference under Gibbs Sampling*.
- Based on *acyclical graphs*:



- *JAGS* interfaces with R.

Rjags programs for Mg/Ca calibration

- Calibration of Mg/Ca model. Files: *calibration-script.R*. Dataset: *calibration.txt*. JAGS/Openbugs model: *model-khider3.txt*, *predictive1.txt*. Version available at https://github.com/khider/
- I*ICTP workshop of uncertainty quantification in Climate Modeling*, July 2015. http://indico.ictp.it/event/a14268/other-view?view=ictptimetable

Portions of work presented funded by the **U.S. Department of Energy, Grant: DE-SC0010843**.

Muchas gracias!