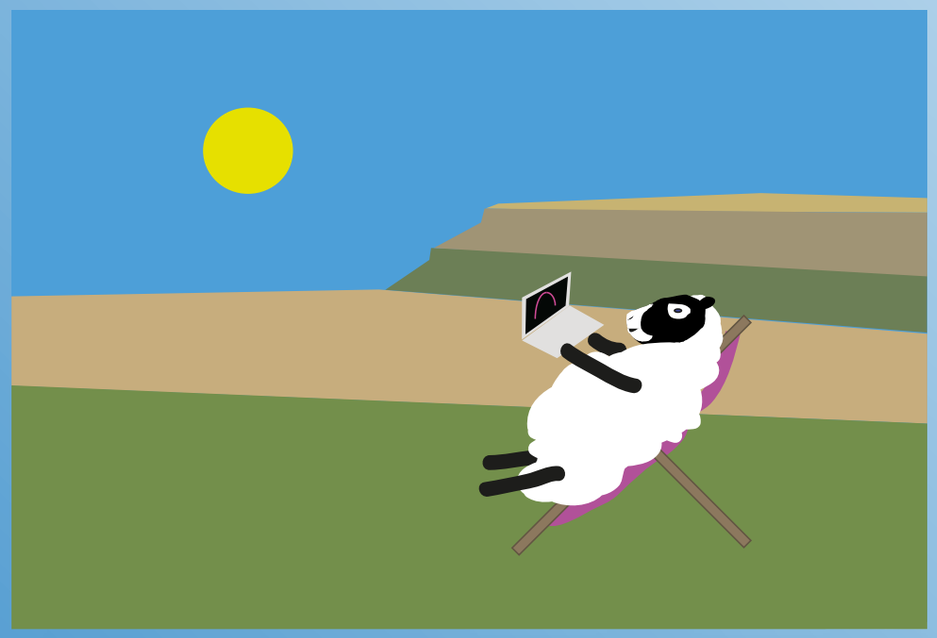
Data Assimilation/Reconstructions

The paleoclimate perspective

James Annan
Julia Hargreaves
(and others)



BlueSkiesResearch.org.uk

Contents

- Intro to Data Assimilation, reconstructions/reanalysis
- Various paleoapplications
- Future prospects/ideas

NWP origins(?) of DA

- Mostly forecast initialisation
- Iterative process
  - Take yesterday's model forecast as prior
  - Update with new observations
  - Run model forward for tomorrow's forecast

Principles

- It's all basically Bayes' Theorem
- $P(Climate|Obs) \propto P(Obs|Climate) \times P(Climate)$
- Differences in how we represent P(Climate) and how to approximate calculation

Kalman equations

- Given forecast x_f, obs O and their related uncertainties (cov matrices) P_f, R we update via:
  - $x_a = R(P_f+R)^{-1}x_f + P_f(P_f+R)^{-1}O$
  - $P_a = RP_f(P_f+R)^{-1}$
- Analysis: x(t) uses data up to time t
- Reanalysis: x(t) uses future data too

Paleo context

- Typically interested in (time-averaged) “climate” rather than (instantaneous) “state”
- → Parameter values more important than state variables
- Augmentation method $M_p(x) \rightarrow M'(x,p)$ can work (with various tricks)
- Limitations...but successes

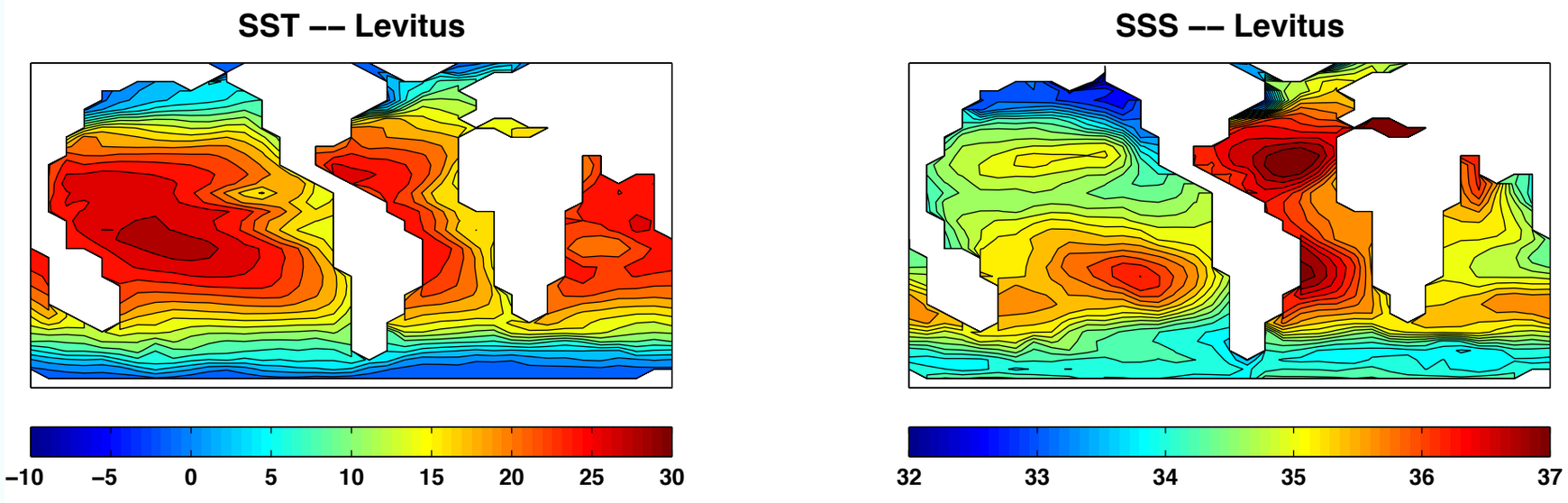
Parameter estimation in EMICs - an example

(Annan et al 2005, Hargreaves et al 2006)

- GENIE - 3D ocean, EMBM atmosphere
- Estimating ~12 parameters using 2D and 3D fields of modern data
- EnKF method with a few tweaks

Japan Uncertainty Modelling Project

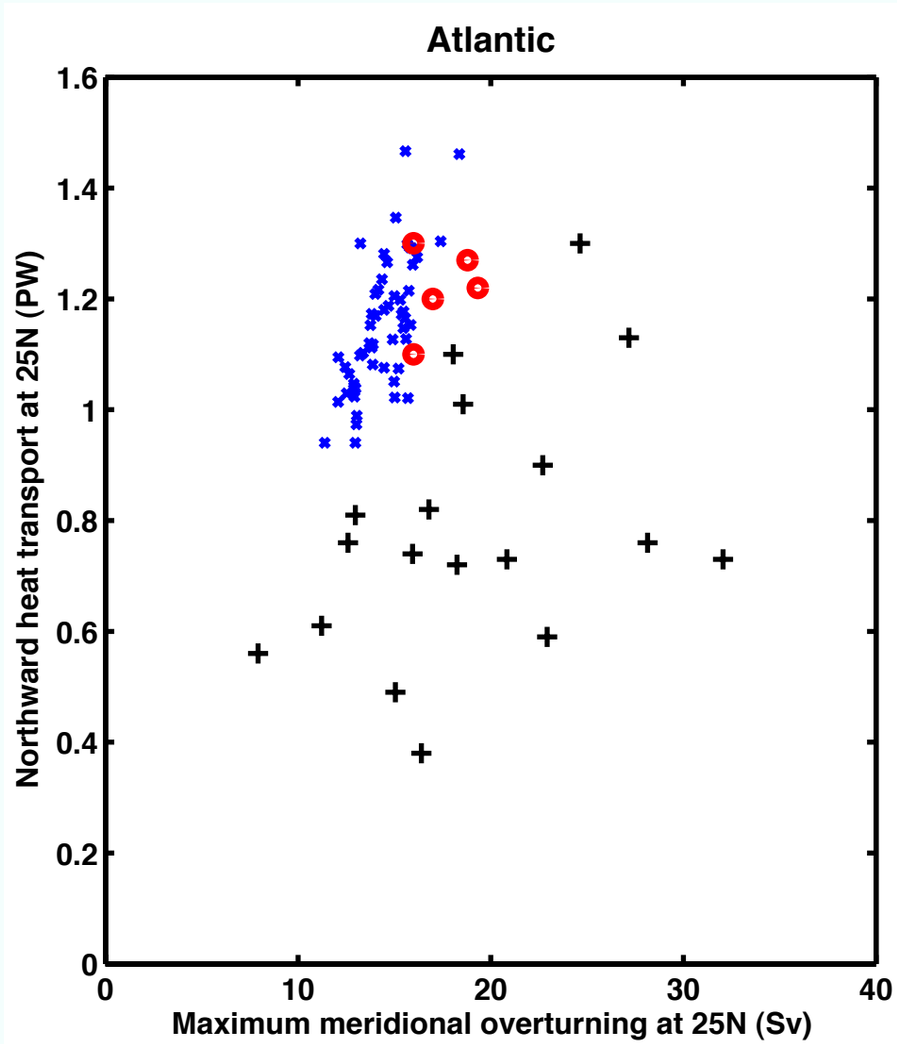
Results of tuning

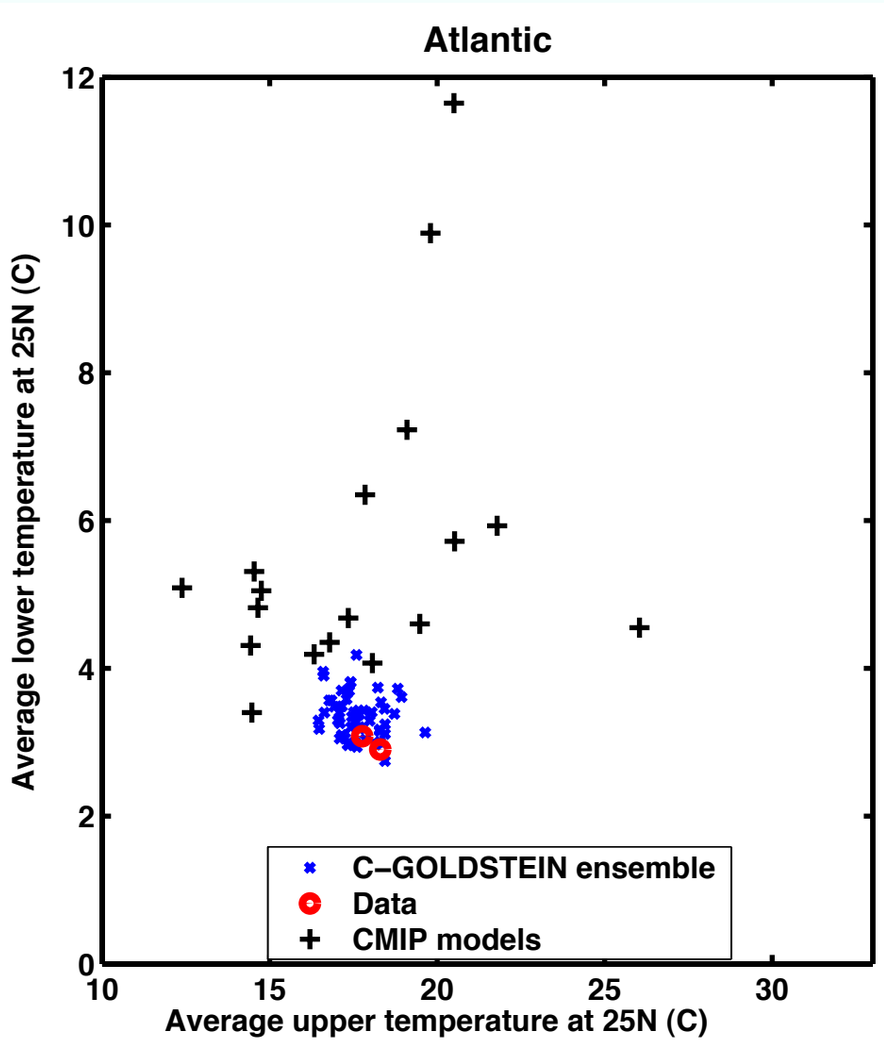


RMS error best to date
(and not subsequently bettered)

Japan Uncertainty Modelling Project

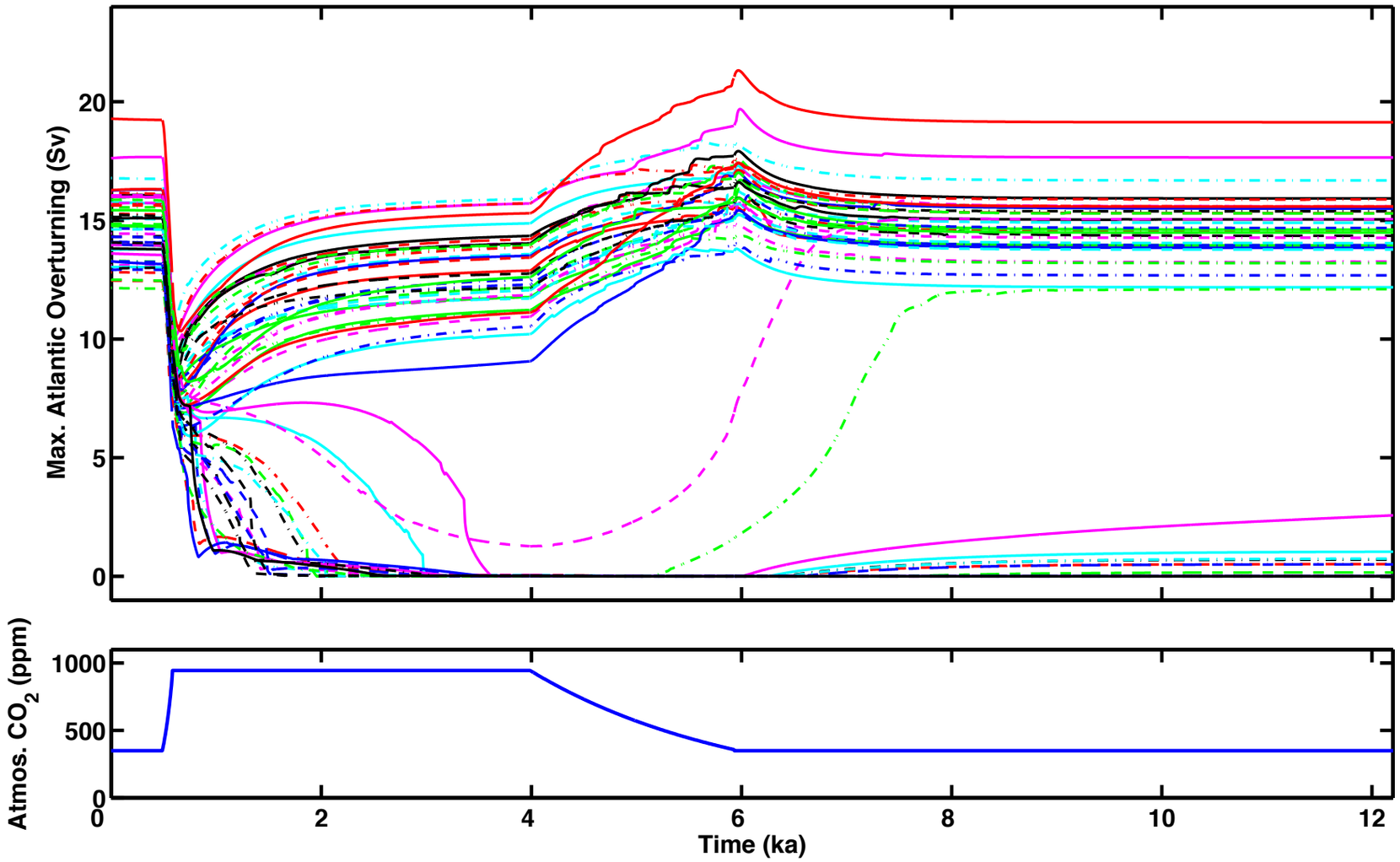
Some validation





Japan Uncertainty Modelling Project

What does GENIE say about future MOC?



MOC can switch off in the future

Japan Uncertainty Modelling Project

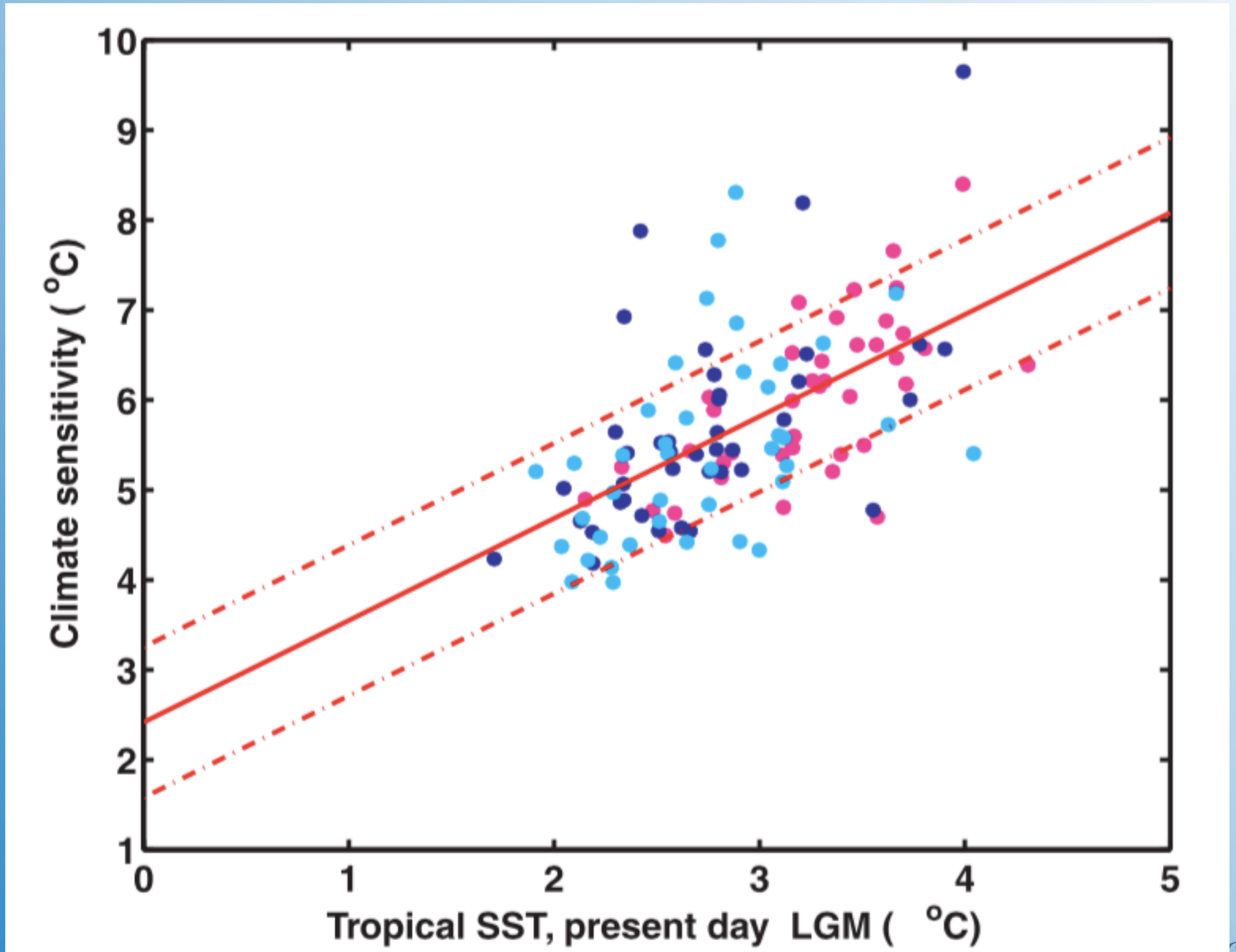
Also GCMs

(Annan et al 2005 SOLA)

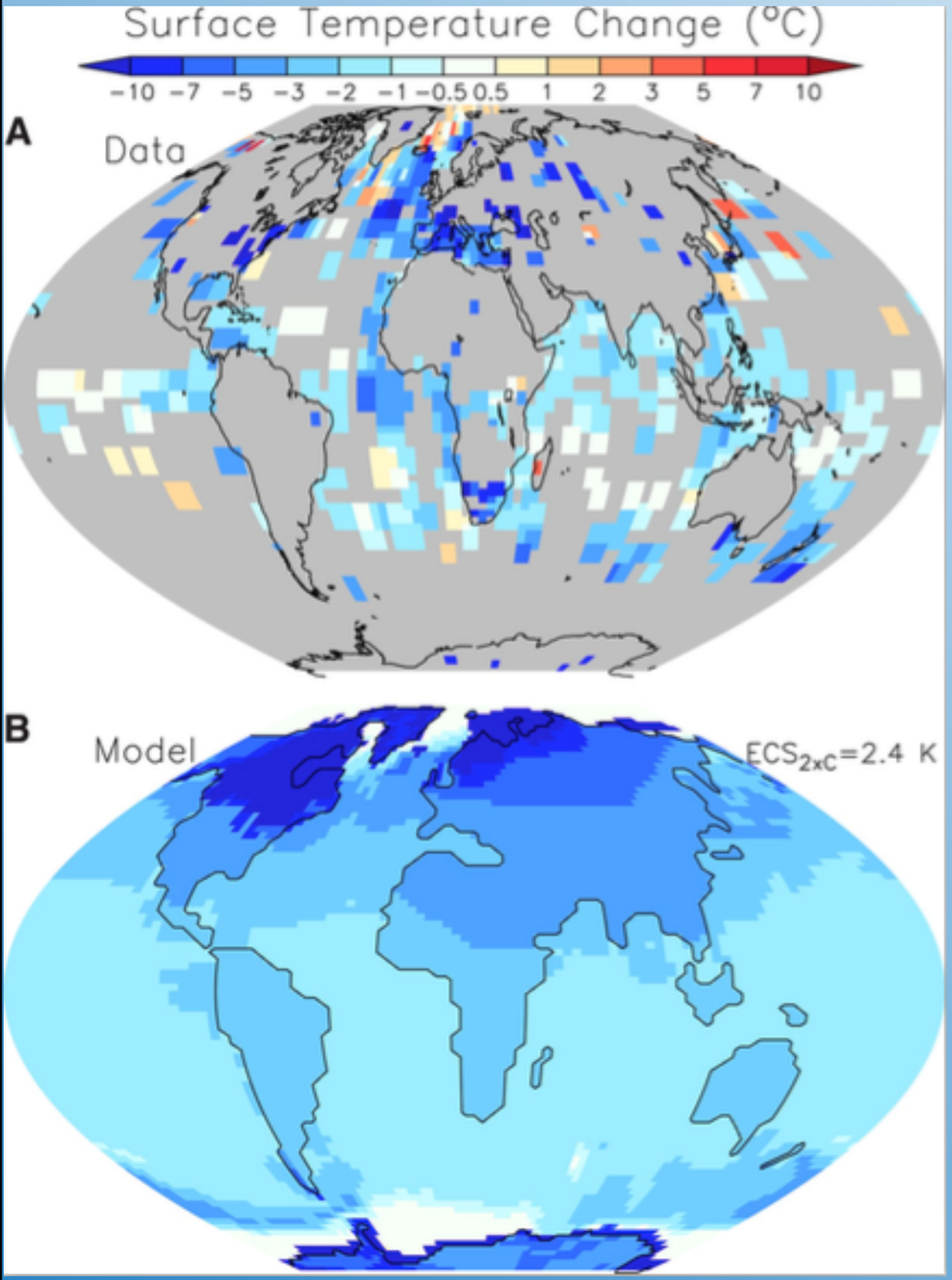
- MIROC3.2 T21L20 ASGCM (CMIP3)
- 25 parameters, 15 2D modern data fields
- 3x40 member ensembles
- LGM sims for validation/further constraint

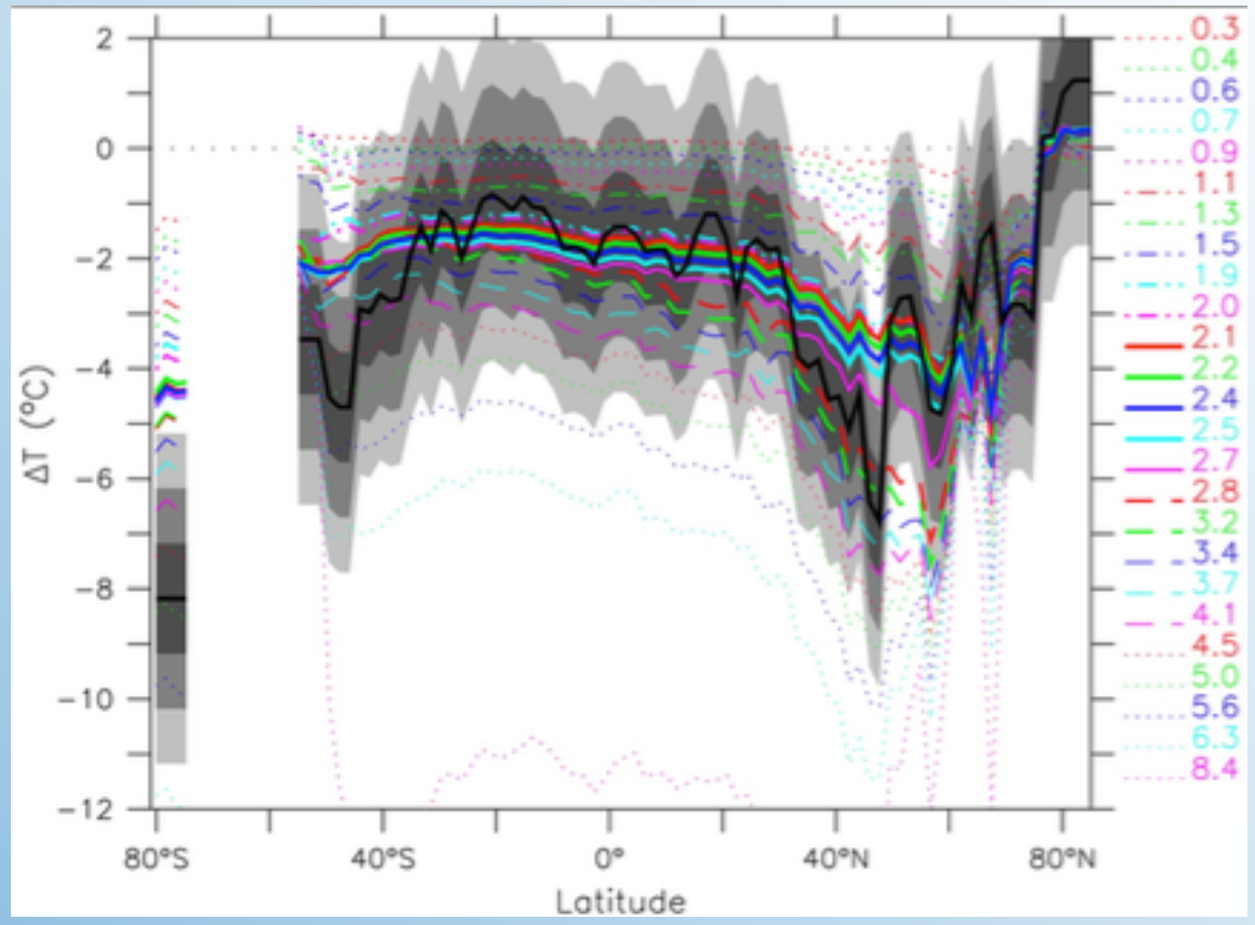
Also GCMs

(Annan et al 2005 SOLA)



Schmittner et al 2011 LGM





- ΔT_{LGM}=3.0C!
- (Too simple) EMBM model
- (Too) sophisticated stats

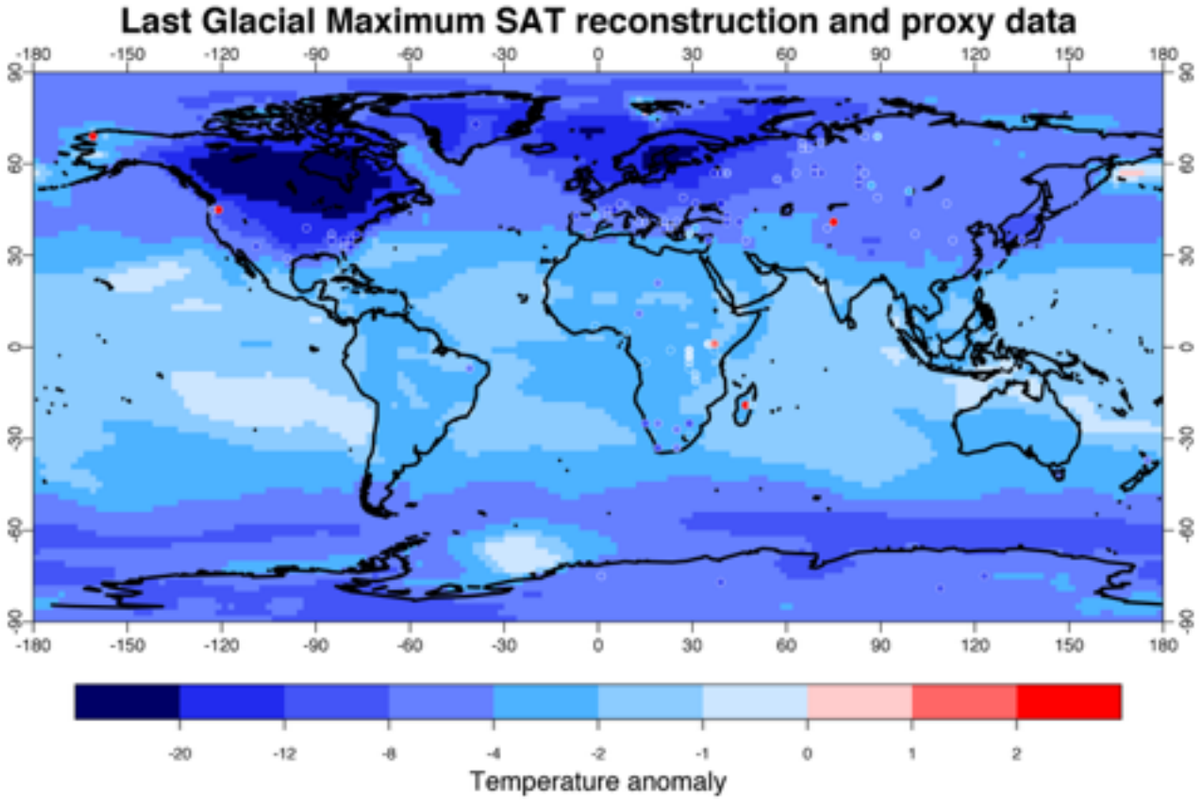
How to do better....quickly

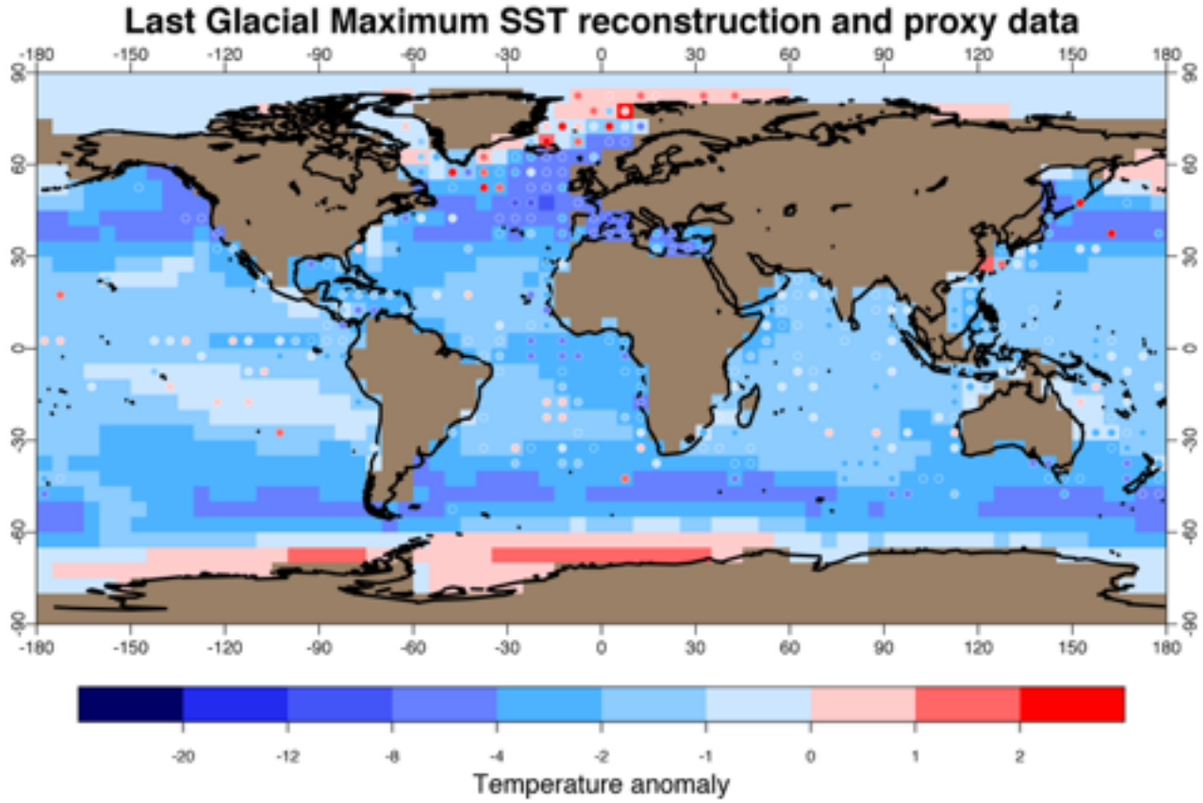
- Want a better model but cannot run even one
- Pattern scaling on PMIP sims?
- Generalise to multiple linear regression
- "Superensemble" approach of Krishnamurti
  - Sum of scaled model anomaly fields $\sum \alpha_i M_i$
  - Not BMA or weighting, scaling factors can be negative and don't sum to 1

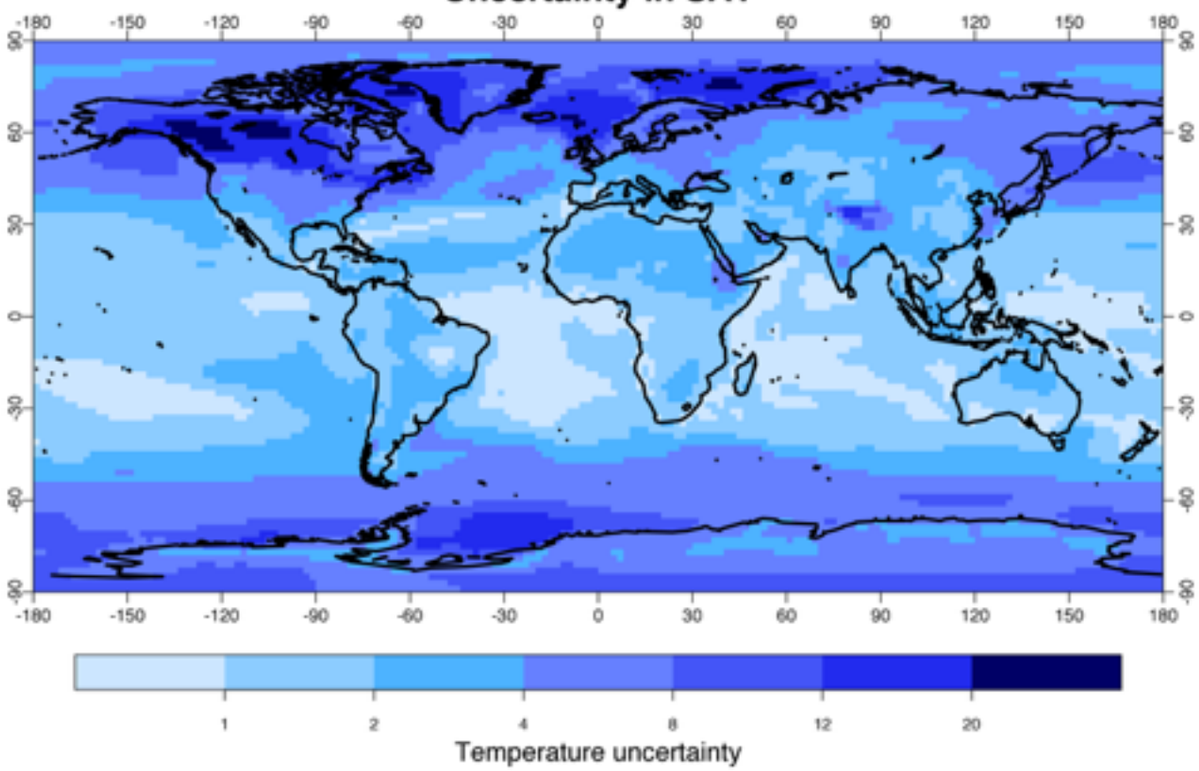
LGM estimate

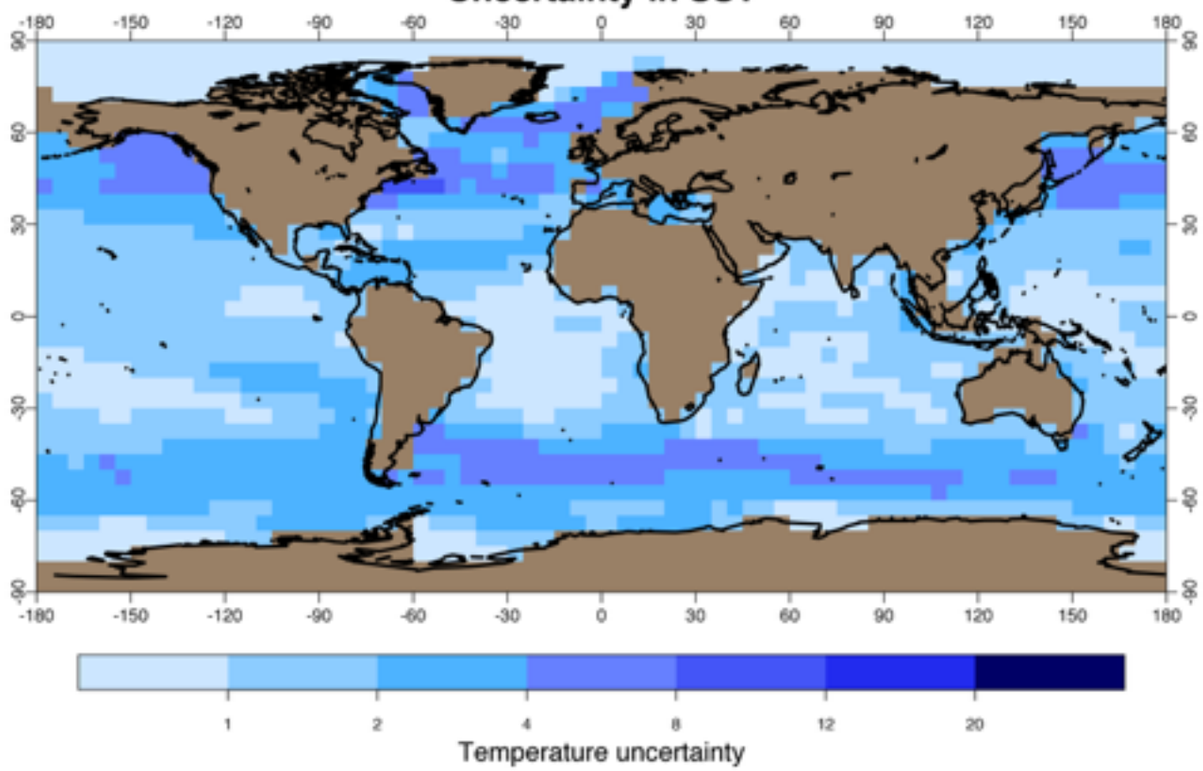
(Annan and Hargreaves 2013)

$\Delta T_{LGM} = 4.0 \pm 0.8C$
(and better fit to data)



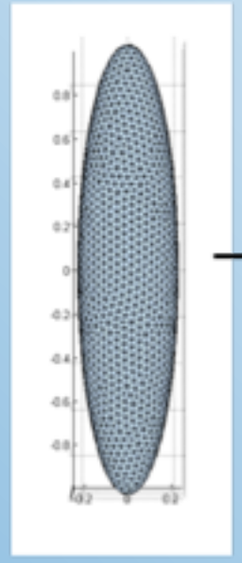






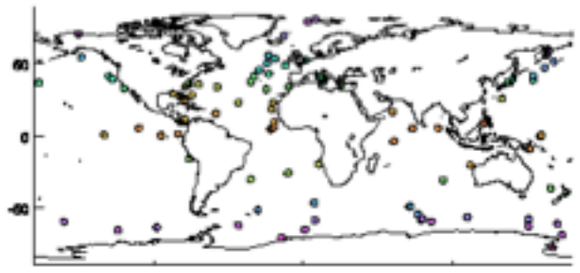
Mid-Pliocene Warm Period

(Bragg, Zammit Mangion)

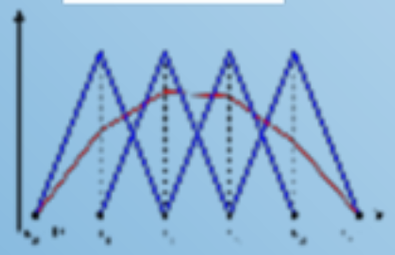


Q – precision matrix
Depends on FE mesh
Describes relationship between nodes.
Markov property means only relate to neighbouring elements and matrix is sparse
smooth goes in here.

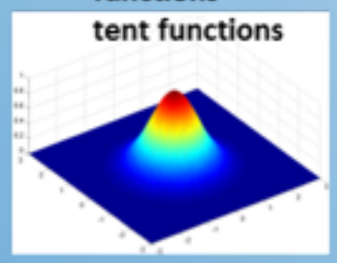
Data points



Weights for basis functions

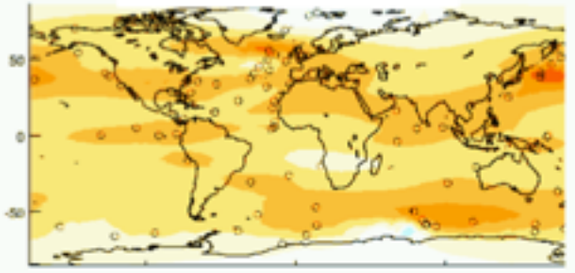


Basis functions – tent functions

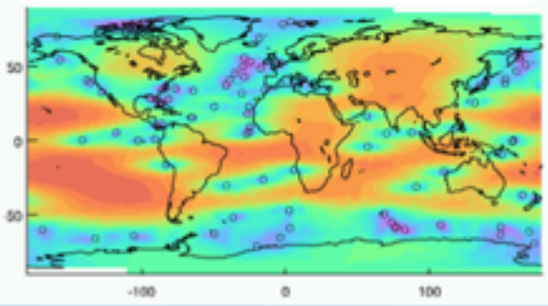


Gaussian updating
Prior GMRF=0
Prior uncertainty

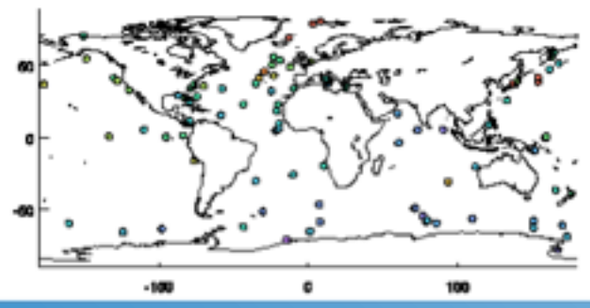
Output field



Output uncertainty



Uncertainties



Gaussian uncertainties

SPDE

- Continuous, infinite dimensional field

Gaussian field

- n elements
- n basis functions
- n dimensional field

Last Millennium etc
(H. Goosse et al)

- Naive particle filter:
- Run 100 climate models for a year
- Pick the best one
  (annually resolved tree ring data)
- Add perturbations
- Goto step 1

Particle filter collapse (Snyder et al)

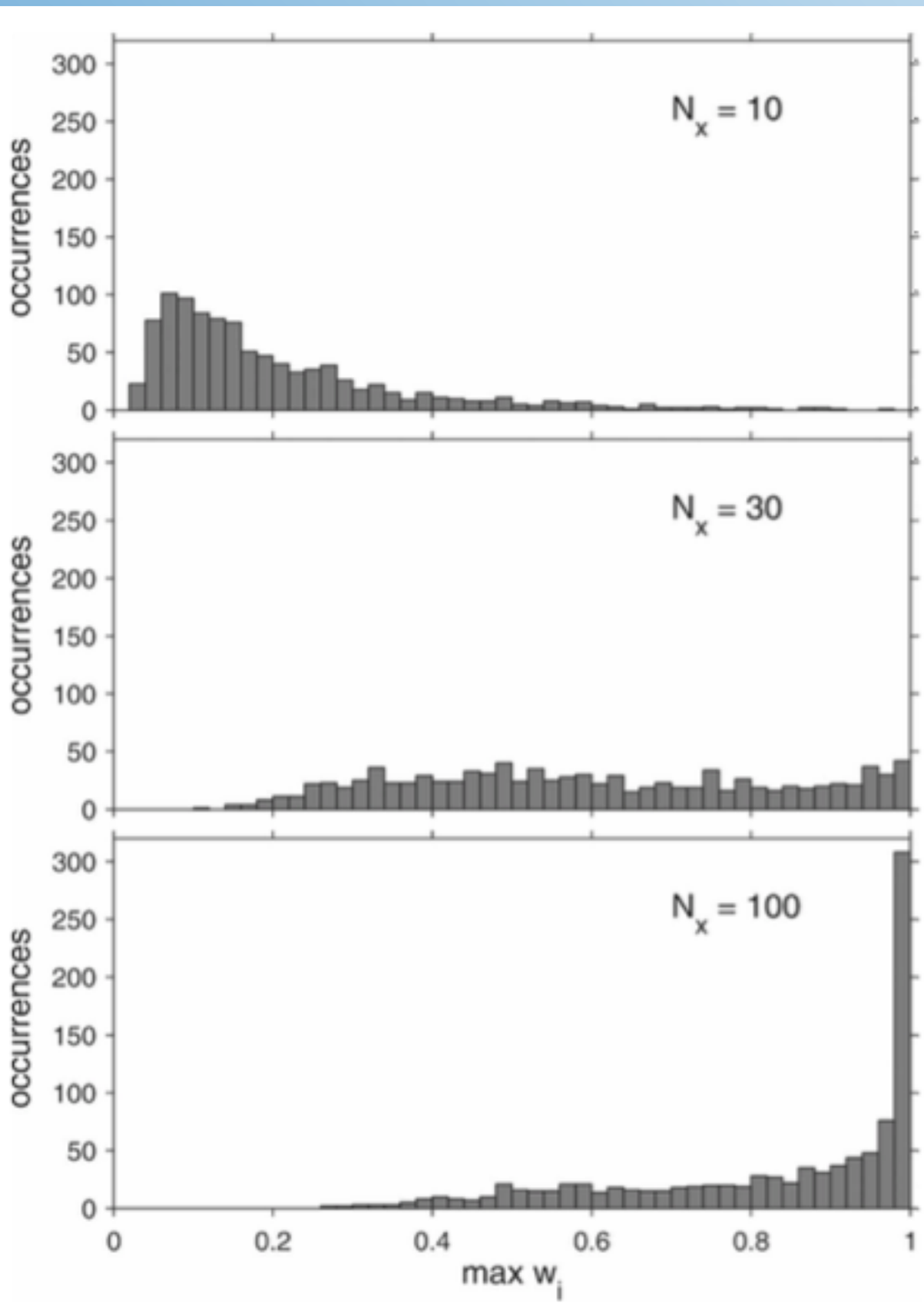


FIG. 1. Histograms of maxw_i for N_x = 10, 30, and 100 and N_e = 10^3 from the particle-filter simulations described in text: $N_e = 10^3$, $\mathbf{x} \sim N(0, \mathbf{I})$, $N_y = N_x$, $\mathbf{H} = \mathbf{I}$, and $\boldsymbol{\epsilon} \sim N(0, \mathbf{I})$.

- Prob(collapse) high for high dimensions
- Required ensemble size increases exponentially with dimension

Problems/successes

- Shouldn't really work (“Best one” is still rubbish)
- $O(10^6)$ or more simulations needed to work properly
- However it does something...

Last Millennium

(Goosse et al)

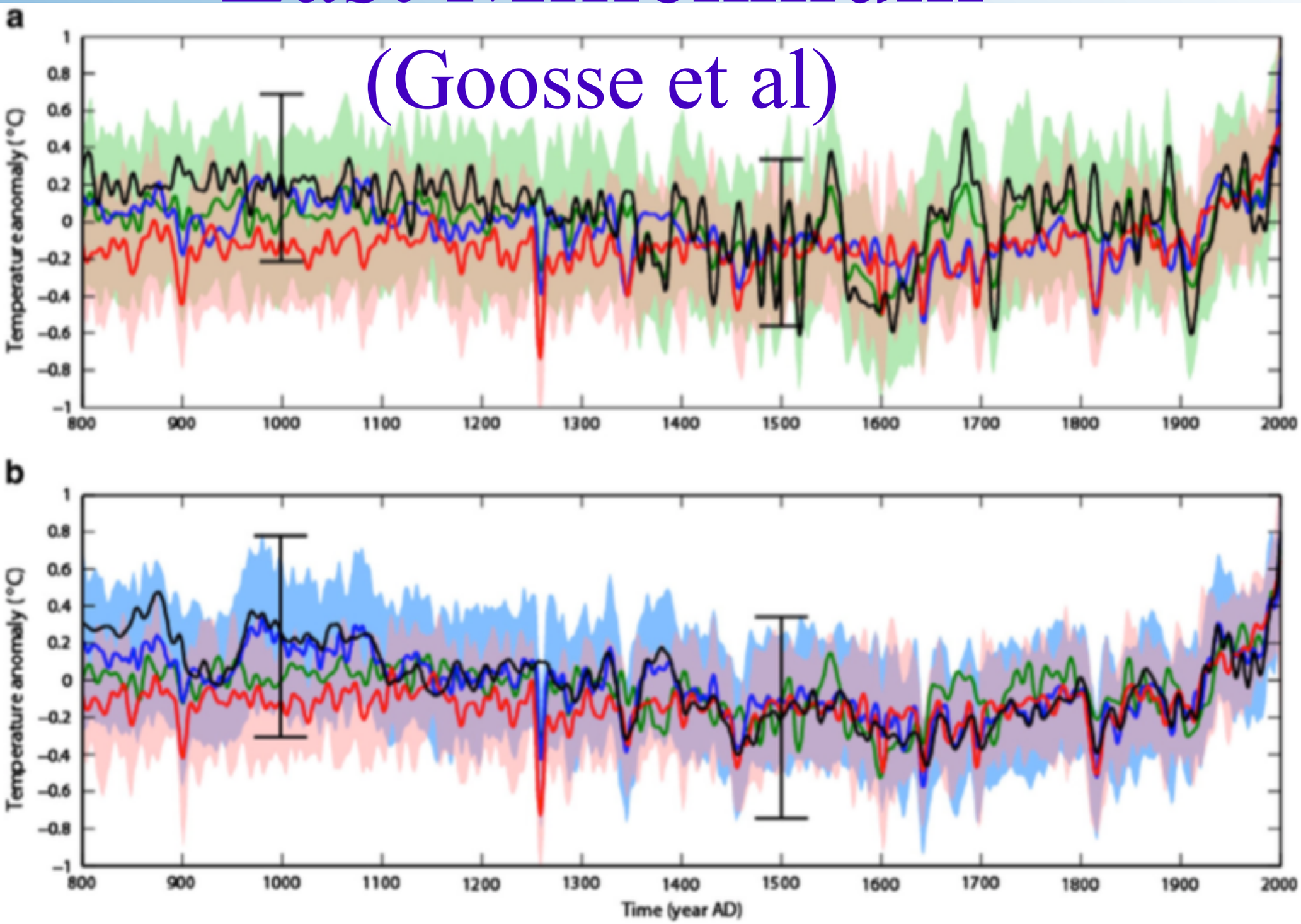
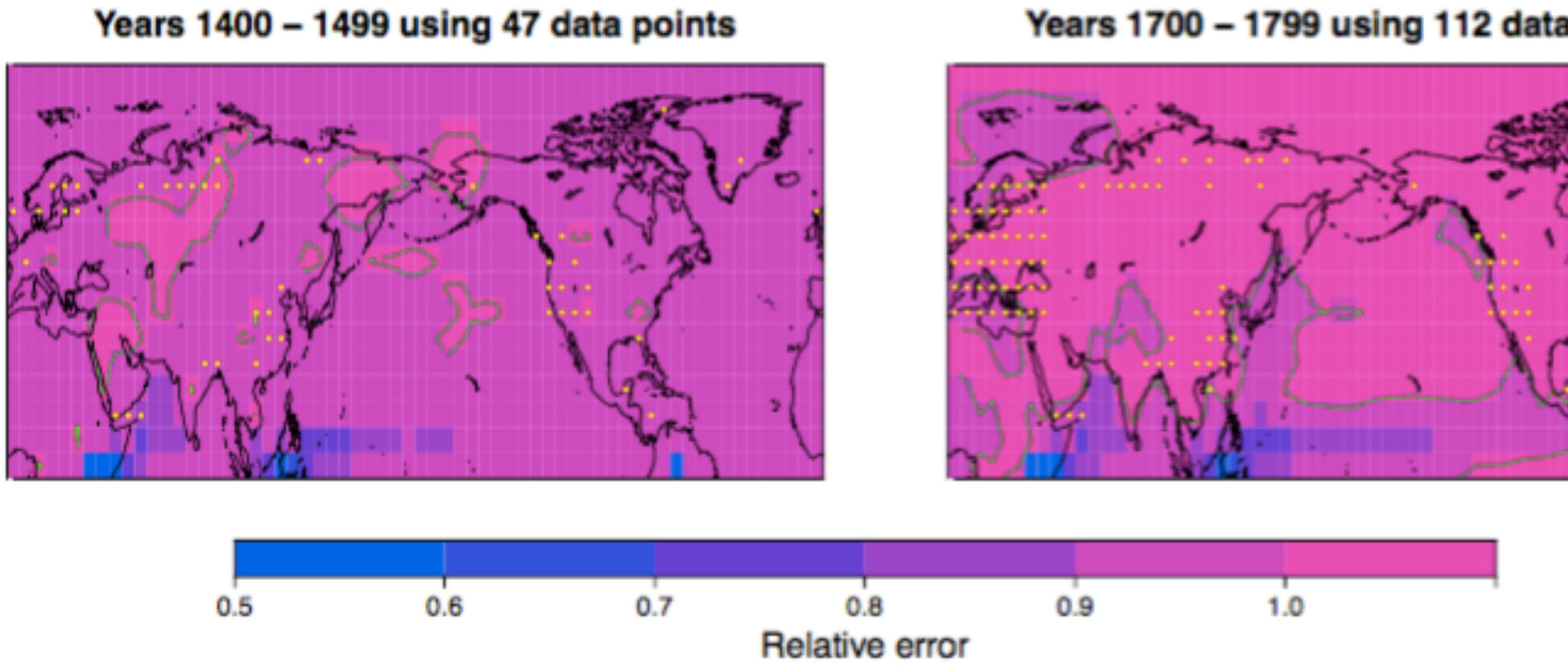
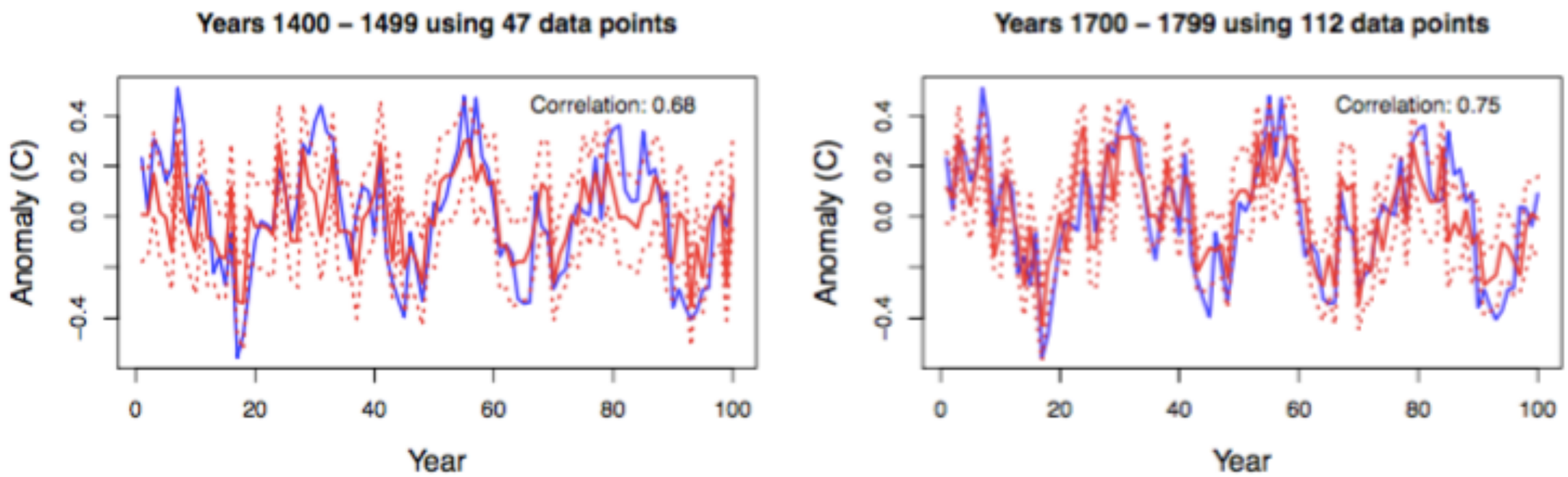


Fig. 1. a) Anomaly of growing season temperature (April to September, °C) averaged over Europe (25°–65°N, 0–60°E) in the reconstruction of Guiot et al. (2010, black), in LOVECLIM simulations without data assimilation (red), in LOVECLIM simulations with data assimilation constrained by Mann et al. (2008) reconstruction (ASSIM-MANN, blue) and in LOVECLIM simulations with data assimilation constrained by Guiot et al. (2010) reconstruction (ASSIM-GUIOT, green). The reference period is 1850–1995. The time series has been filtered using an 11-year running mean. b) Same as a) but for annual mean and the black curve representing the reconstruction of Mann et al. (2008). The pink shading represents the range of the simulations without data assimilation (mean plus and minus two standard deviations), the green shading the one of ASSIM-GREEN (panel a only) and the blue shading the one of ASSIM-MANN (panel b only). The overlap of the uncertainties of the simulations is represented by a darker shading. For the clarity of the plot, the uncertainties of the reconstructions are shown as error bars only for the years 1000 and 1500.

Performance of particle methods
(Annan and Hargreaves 2013)





Why does it work?

- Not a valid estimate, locally rubbish (as per theory)
- But large-scale average has substantial skill
- Consider projection onto dominant (forced response and int. var.) modes?
- cf statistical methods (Tingley, Christiansen)

Bayesian Model Averaging

- Weighted average $\sum w_i M_i$, $\sum w_i = 1$
- (also model-specific bias terms and variances)
- Applicable to multi-model ensembles

Last Millennium

(Fang and Li)

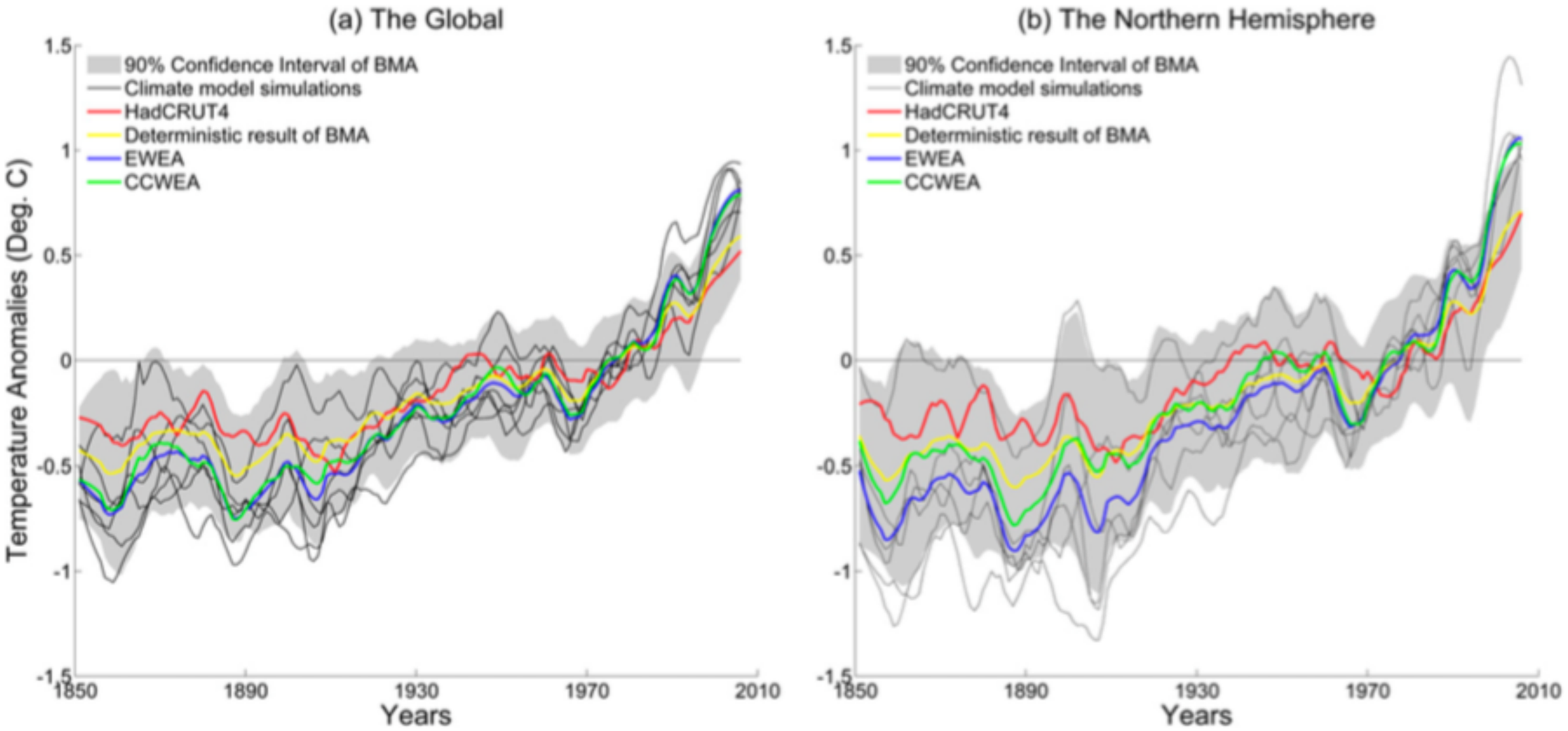


FIG. 5. Annual air temperature anomalies (°C) derived from the multimodel ensemble simulations and the results of BMA, EWEA, and CCWEA over the period of 1850–2005. All of the anomalies are with respect to the mean climatology from 1961 to 1990 and are smoothed using an 11-yr filter. The 90% confidence interval of the BMA estimation is indicated with gray shading.

Last Millennium

(Fang and Li)

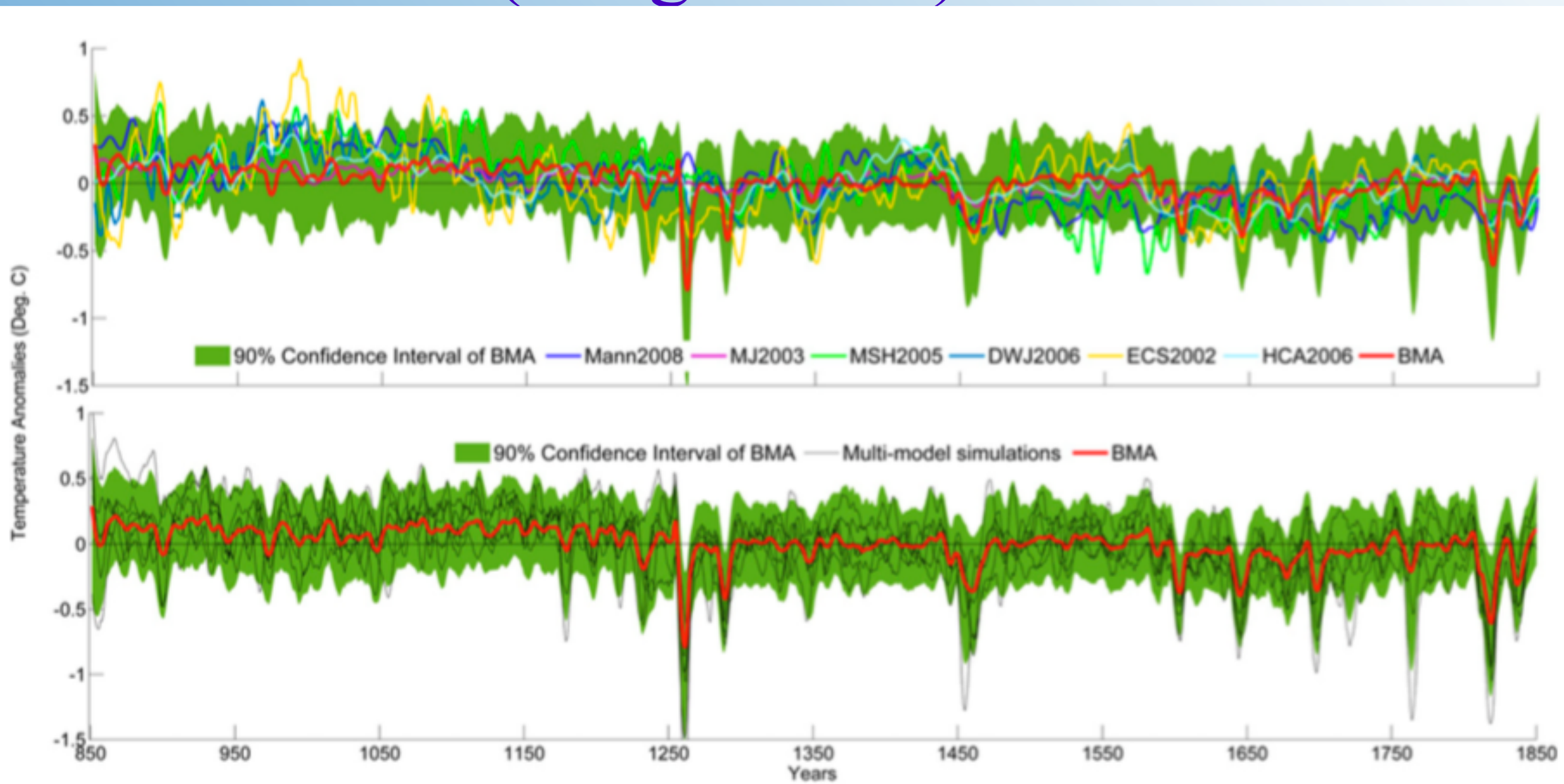


FIG. 7. Annual NH temperature anomalies (°C) over the past millennium (from 850 to 1849) from (top) six proxy-based reconstructions and (bottom) the results of BMA using the multimodel ensemble simulations and the individual model simulations. All of the time series are anomalous relative to their individual millennial means and are smoothed using an 11-yr filter. The 90% confidence interval of the BMA estimation is indicated by green shading.

Data Assimilation reanalyses

- ERA40, JMA25
- 20th Century Reanalysis Project
- Last Millennium
- And now...

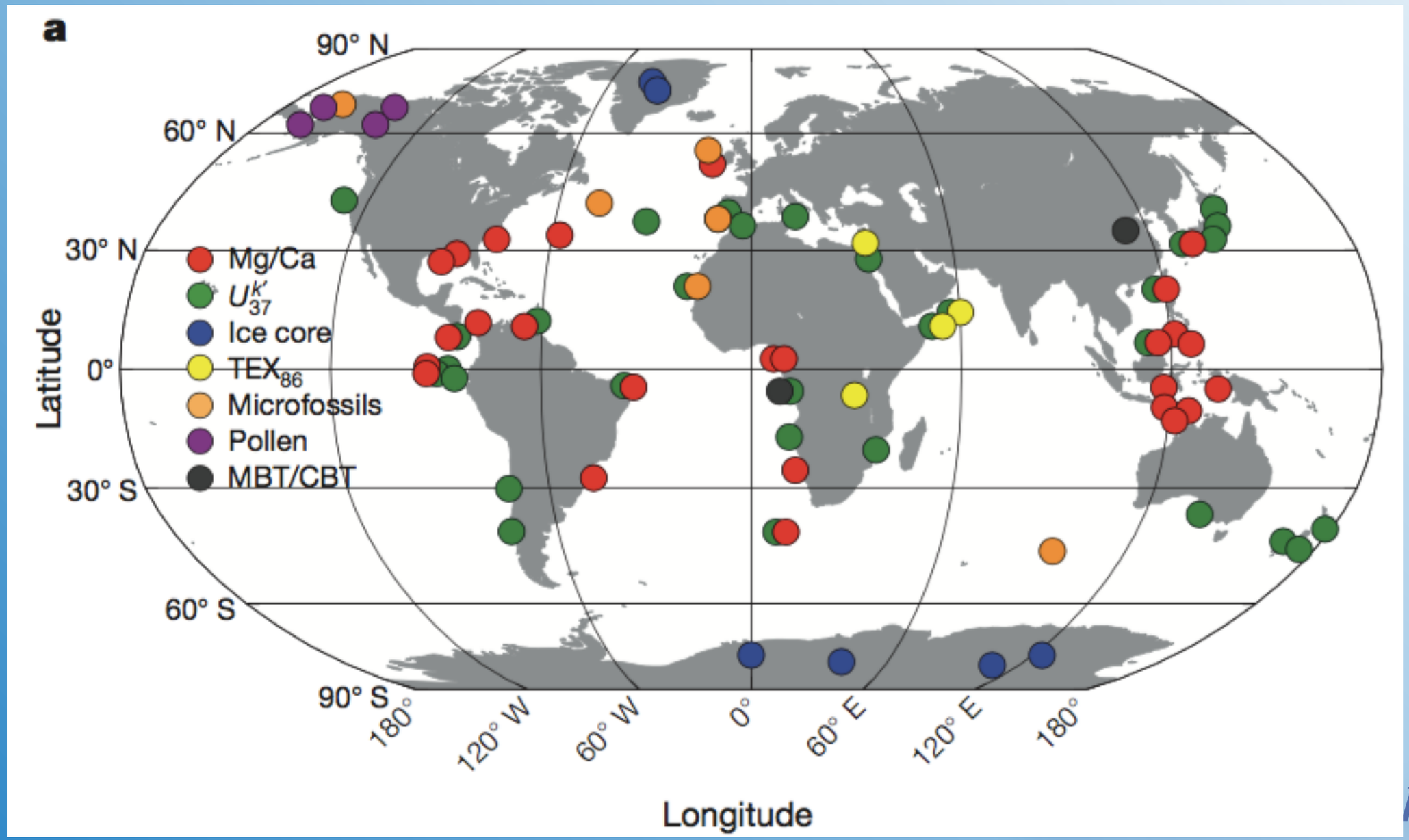
A Grand Challenge - 21kaRP?

- 21ka Reanalysis Project
- Reconstruction of climate state and climate changes over the last deglaciation
- Including - ice sheet collapse, green sahara (monsoon enhancement), CO_2 changes, climate sensitivity, ocean circulation changes...

Japan Uncertainty Modelling Project

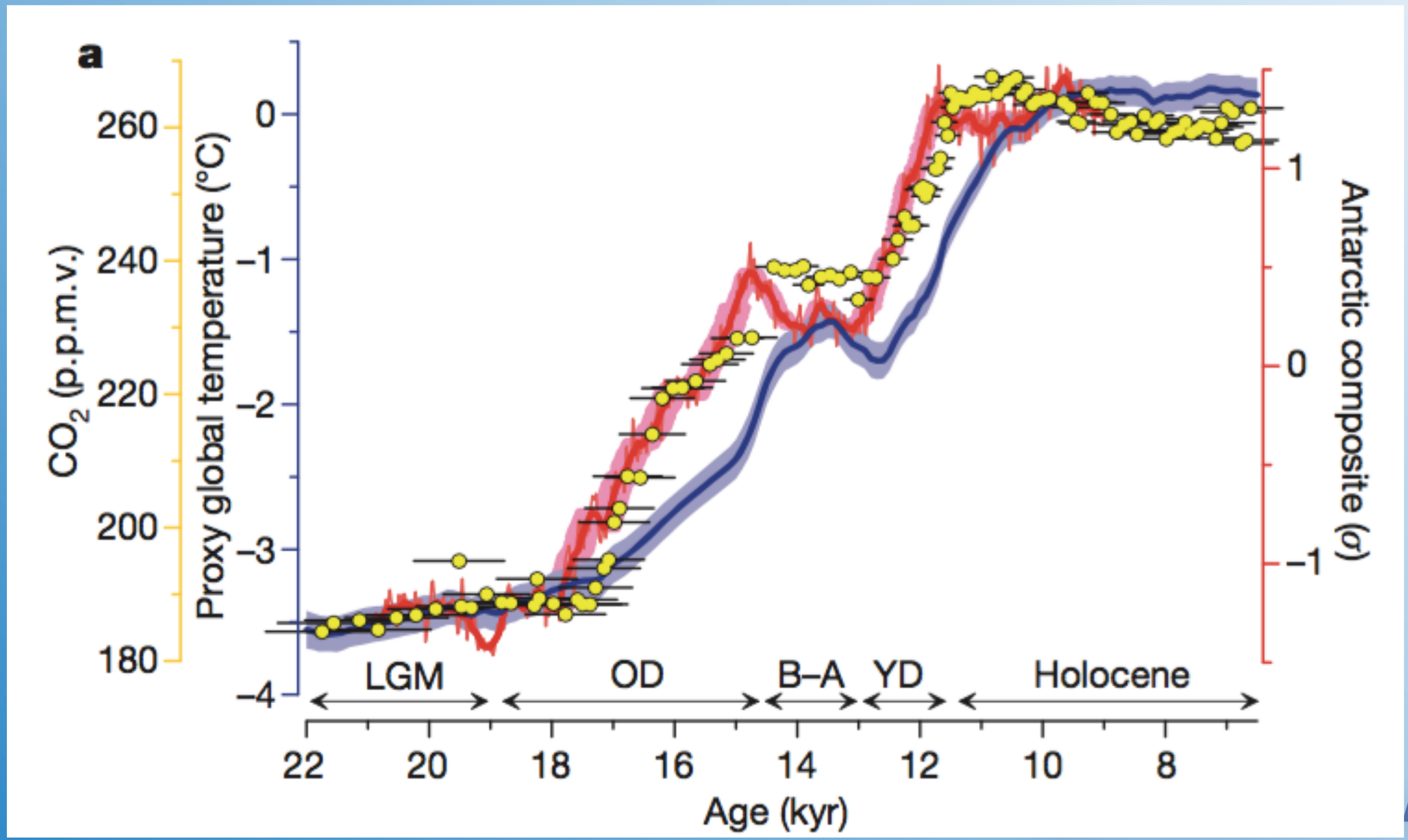
Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation

Jeremy D. Shakun[1,2], Peter U. Clark[3], Feng He[4], Shaun A. Marcott[3], Alan C. Mix[3], Zhengyu Liu[4,5,6], Bette Otto-Bliesner[7], Andreas Schmittner[3] & Edouard Bard[8]

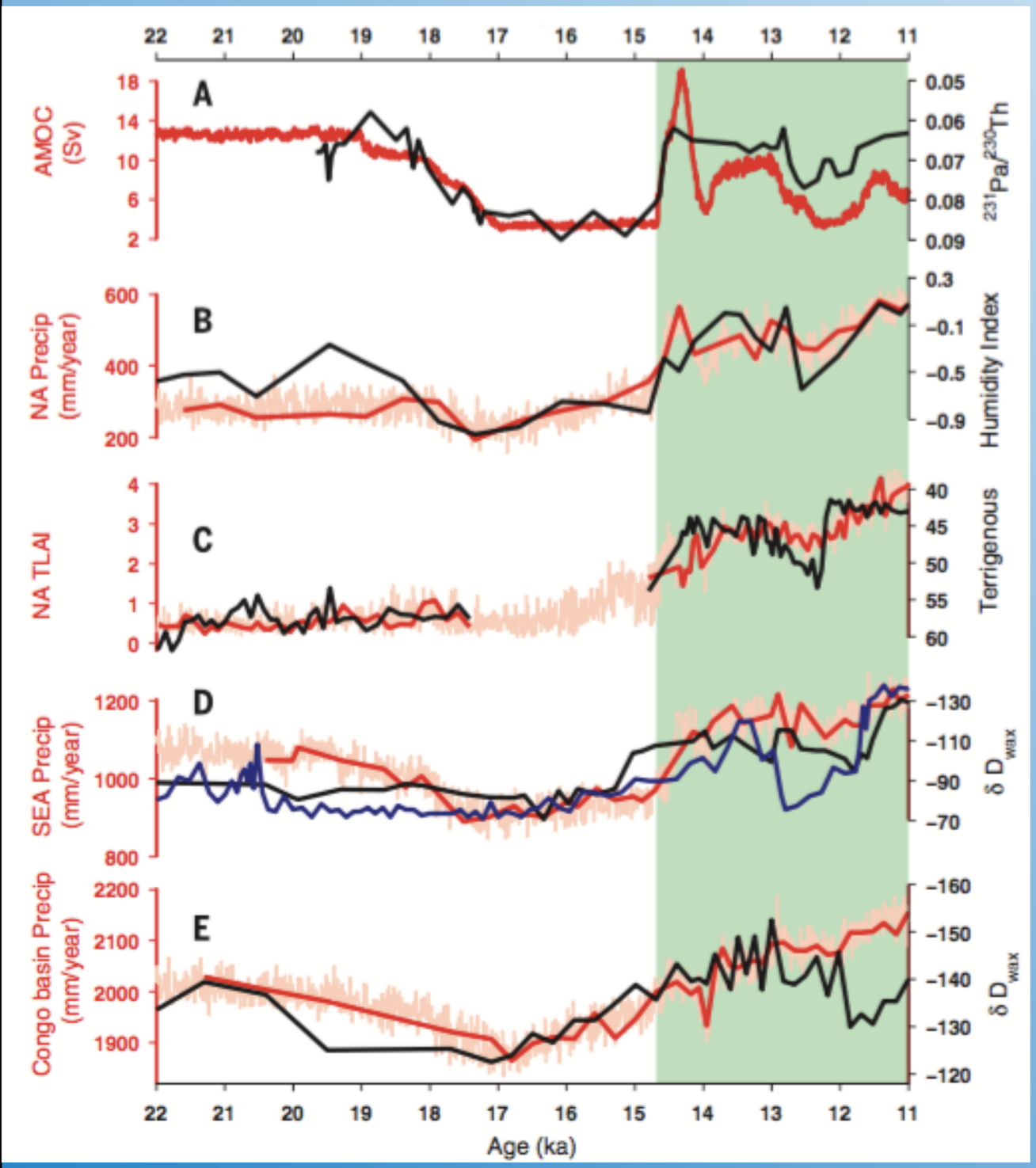


Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation

Jeremy D. Shakun[1,2], Peter U. Clark[3], Feng He[4], Shaun A. Marcott[3], Alan C. Mix[3], Zhengyu Liu[4,5,6], Bette Otto-Bliesner[7], Andreas Schmittner[3] & Edouard Bard[8]



Model simulations



- Otto-Bliesner et al 2014
- More state of the art GCM sims to come (CMIP6)

Other things

- Ensembles of cheaper models (and can do more?)
- Timing/magnitude errors in forcings, data

Progress so far...

Japan Uncertainty Modelling Project