~~What drives the glacial-interglacial cycle? A Bayesian approach to a long-standing model selection problem~~
How far can fancy Monte Carlo methods take us?

Jake Carson Michel Crucifix Simon Preston
Richard Wilkinson

Universities of Warwick, Louvain, Nottingham and Sheffield

PEN 2015

Outline

Most simple models of the [...] glacial cycles have at least four degrees of freedom [parameters], and some have as many as twelve. Unsurprisingly [...this is] insufficient to distinguish between the skill of the various models (Roe and Allen 1999)

Outline

> *Most simple models of the [...] glacial cycles have at least four degrees of freedom [parameters], and some have as many as twelve. Unsurprisingly [...this is] insufficient to distinguish between the skill of the various models (Roe and Allen 1999)*

Our focus is on statistical computation:

- Can we combine a 'simulator', proxy model, and data in a Bayesian analysis?
  - Cf. Johannes Werner, Jonty Rougier, Andrew Parnell's talks/posters
- What can we learn?
- Does it matter if we cut feedbacks between climate and age, fitting each component independently?
- What can happen if we ignore dating uncertainties?

Outline

Most simple models of the [...] glacial cycles have at least four degrees of freedom [parameters], and some have as many as twelve. Unsurprisingly [...this is] insufficient to distinguish between the skill of the various models (Roe and Allen 1999)

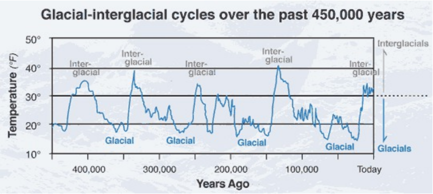
Our focus is on statistical computation:

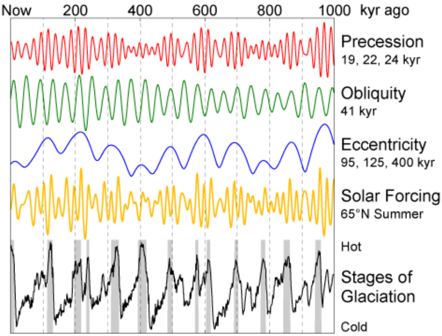
- Can we combine a 'simulator', proxy model, and data in a Bayesian analysis?
  - Cf. Johannes Werner, Jonty Rougier, Andrew Parnell's talks/posters
- What can we learn?
- Does it matter if we cut feedbacks between climate and age, fitting each component independently?
- What can happen if we ignore dating uncertainties?

Essentially a demonstration of recent Monte Carlo methodology (SMC^2, PMCMC), and GPU computation.

Many aspects of the modelling could be improved, and be incorporated within this framework.

Glacial-Interglacial cycle



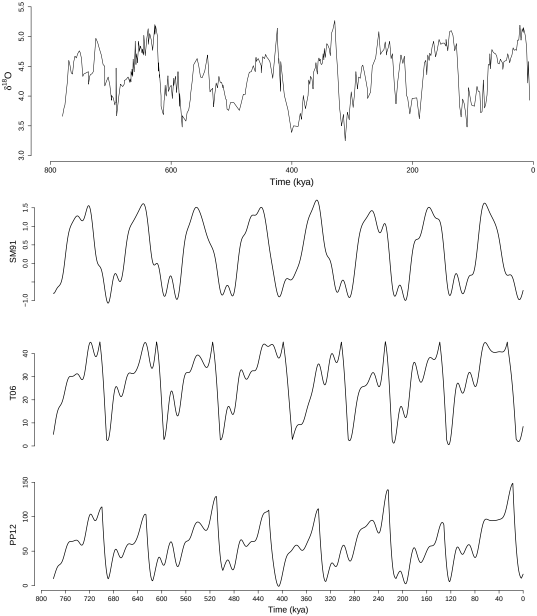


Eccentricity: orbital departure from a circle, controls duration of the seasons

Obliquity: axial tilt, controls amplitude of seasonal cycle

Precession: variation in Earth's axis of rotation, affects difference between seasons

Which is the better model?



Models

Use conceptual models based on a few hypothesised relationships that capture some aspect of the climate system, driven by some aspect of the solar forcing

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = g(X_t, \theta) + F(t, \gamma)$$

where $\gamma = (\gamma_P, \gamma_C, \gamma_E)$ controls the combination of precession, obliquity and eccentricity.

$X_t \in \mathbb{R}^p$ denotes the state of the climate at time t.

Models

Use conceptual models based on a few hypothesised relationships that capture some aspect of the climate system, driven by some aspect of the solar forcing

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = g(X_t, \theta) + F(t, \gamma)$$

where $\gamma = (\gamma_P, \gamma_C, \gamma_E)$ controls the combination of precession, obliquity and eccentricity.
$X_t \in \mathbb{R}^p$ denotes the state of the climate at time t.

Embed these simulators within a state space model relating climate to observations

$$\begin{aligned} \mathrm{d}X_t &= g(X_t, \theta)\mathrm{d}t + F(t, \gamma)\mathrm{d}t + \Sigma \mathrm{dW} \\ Y_t &= d + sX_{1,t} + \epsilon_t \end{aligned}$$

The models have 10–15 parameters that need to be estimated

The statistical quantities we would like to calculate are

- Climate reconstruction (filtering)

$$\pi(x_{1:T}|y_{1:T}, \theta, \mathcal{M})$$

where $x_{1:T} = (x_1, \ldots, x_T)$

The statistical quantities we would like to calculate are

- Climate reconstruction (filtering)

$$\pi(x_{1:T}|y_{1:T}, \theta, \mathcal{M})$$

  where $x_{1:T} = (x_1, \ldots, x_T)$

- Model calibration (marginal parameter posterior)

$$\pi(\theta|y_{1:T}, \mathcal{M})$$

The statistical quantities we would like to calculate are

- Climate reconstruction (filtering)

$$\pi(x_{1:T}|y_{1:T},\theta,\mathcal{M})$$

  where $x_{1:T} = (x_1,\dots,x_T)$
- Model calibration (marginal parameter posterior)

$$\pi(\theta|y_{1:T},\mathcal{M})$$

- Model selection (model evidence/Bayes factors)

$$\pi(y_{1:T}|\mathcal{M})$$

The statistical quantities we would like to calculate are

- Climate reconstruction (filtering)

$$\pi(x_{1:T}|y_{1:T}, \theta, \mathcal{M})$$

  where $x_{1:T} = (x_1, \ldots, x_T)$
- Model calibration (marginal parameter posterior)

$$\pi(\theta|y_{1:T}, \mathcal{M})$$

- Model selection (model evidence/Bayes factors)

$$\pi(y_{1:T}|\mathcal{M})$$

These are progressively more difficult to calculate. Moreover,

$$\pi(X_{t+1}|X_t, \theta, \mathcal{M})$$

is unknown ruling out many Monte Carlo approaches.

Bayes factors

Consider comparing two models, $\mathcal{M}_1$ and $\mathcal{M}_2$. Bayes factors (BF) are the Bayesian approach to model selection.

The log Bayes factor is the difference between the log-likelihoods for two models

$$\log(BF) = \log \pi(y_{1:T} \mid \mathcal{M}_1) - \log \pi(y_{1:T} \mid \mathcal{M}_2)$$

$$\text{posterior odds} = \text{prior odds} \times \text{Bayes factor}$$

$\log(BF)$ range	$\mathbb{P}(\mathcal{M}_1 \vert D)$	Interpretation
> 5	0.99 - 1	V. strong evidence in favour of model $\mathcal{M}_1$
3 to 5	0.95 - 0.99	Strong evidence in favour of model $\mathcal{M}_1$
-3 to 3	0.05 - 0.95	Grey zone
-3 to -5	0.01 - 0.05	Strong evidence in favour of $\mathcal{M}_2$
< -5	0 - 0.01	V. strong evidence in favour of model $\mathcal{M}_2$

Results: synthetic data

We generate simulated data from SM91, using both the astronomically forced and unforced version of the model.

- 800 data points over the last 800kyr
- Realistic measurement and discrepancy variances used

Can we infer the parameter values used? And the model used?

Results: synthetic data

We generate simulated data from SM91, using both the astronomically forced and unforced version of the model.

- 800 data points over the last 800kyr
- Realistic measurement and discrepancy variances used

Can we infer the parameter values used? And the model used?

Model		Dataset	
		SM91-unforced	SM91-forced
SM91	Forced		0
	Unforced		−52.4
T06	Forced		−24.7
	Unforced		−61.4
PP12	Forced		−52.5

- Strongest evidence for the true model found each time
- For the data generated from the forced model, the forced version of the wrong model is preferred.

Results: synthetic data

We generate simulated data from SM91, using both the astronomically forced and unforced version of the model.

- 800 data points over the last 800kyr
- Realistic measurement and discrepancy variances used

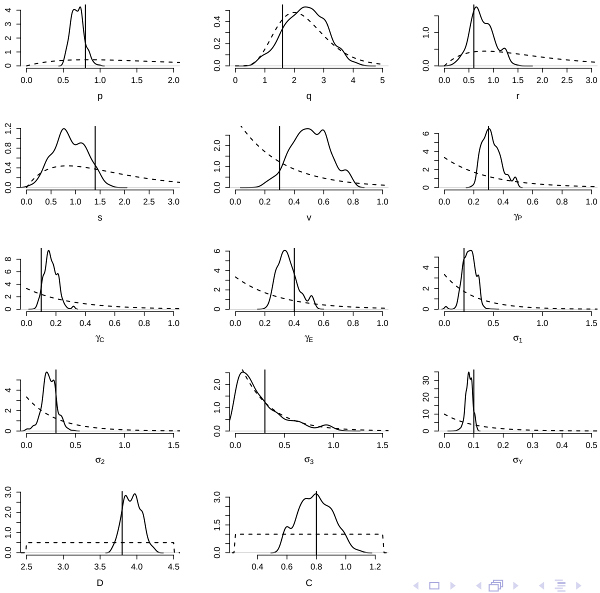
Can we infer the parameter values used? And the model used?

Model		Dataset	
		SM91-unforced	SM91-forced
SM91	Forced	−2.9	0
	Unforced	0	−52.4
T06	Forced	−21.8	−24.7
	Unforced	−18.3	−61.4
PP12	Forced	−49.6	−52.5

- Strongest evidence for the true model found each time
- For the data generated from the forced model, the forced version of the wrong model is preferred.

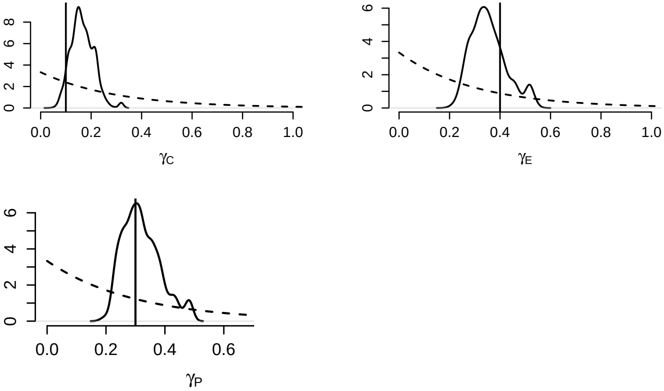
Results: synthetic data

True param = vertical line, solid line = posterior, dashed = prior



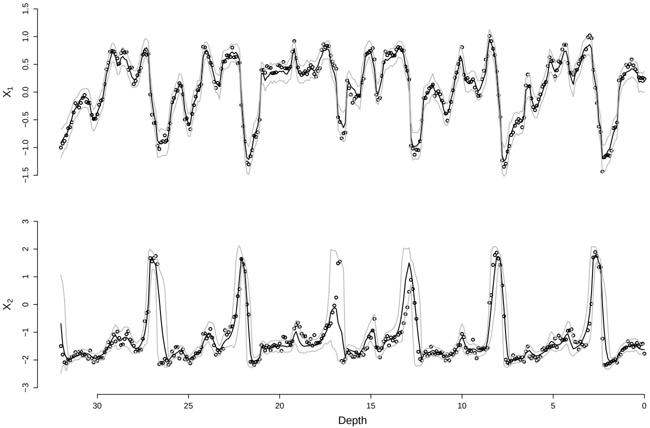
Results: synthetic data

$\gamma = (\gamma_P, \gamma_E, \gamma_C)$ controls the relative contribution of the three components of the orbital variations in the forcing.

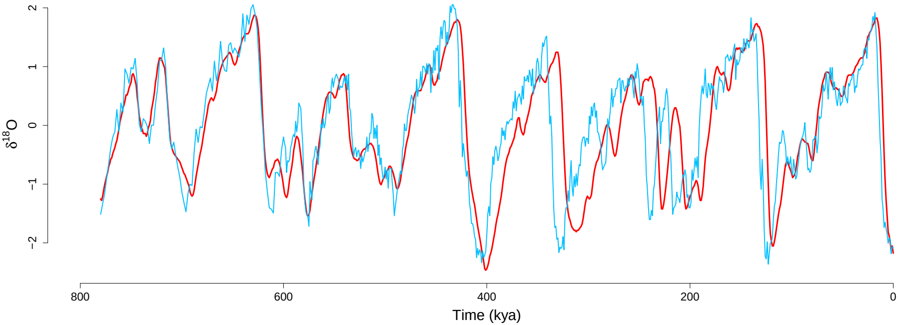


Results: synthetic data - climate reconstruction

Dots = truth, black line = estimate, grey = 95% CI



Results: ODP677



We use the ODP677 stack (a composite record from multiple cores), which has been dated by two authors:

- Lisiecki and Raymo (2005) used orbital tuning
- Huybers 2007 used a depth-derived age model (no orbital tuning)

Do we get the same results using the two dating estimates?

Results: ODP677

Model		Dataset	
		H07(unforced)	LR04(forced)
SM91	Forced	−8.4	
	Unforced	−3.9	
T06	Forced	−6.2	
	Unforced	0	
PP12	Forced	−13.9	

- Using Huybers' non-orbitally tuned data, we find evidence in favour of the unforced T06 model.

Results: ODP677

Model		Dataset	
		H07(unforced)	LR04(forced)
SM91	Forced	−8.4	−14.3
	Unforced	−3.9	−37.0
T06	Forced	−6.2	−10.6
	Unforced	0	−29.4
PP12	Forced	−13.9	0

- Using Huybers' non-orbitally tuned data, we find evidence in favour of the unforced T06 model.
- Using Lisiecki's orbitally tuned data, we find strong evidence for forced models (PP12)

The dating method applied changes the answer
Orbitally tuned data leads us to strongly prefer the orbitally tuned version of each model (and vice versa)

The age model used to date the stack (often taken as a given) has a strong effect on model selection conclusions

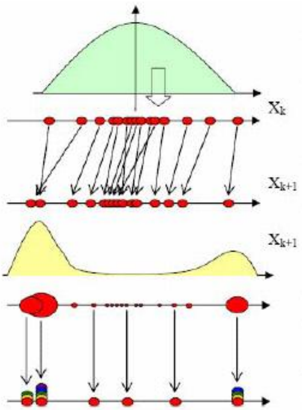
Computational details

Sequential Monte Carlo (SMC) methods are the natural approach for finding the filtering distributions $\pi(x_{1:T}|y_{1:T}, \theta)$

- Represent all distributions by collection of weighted particles $\{x^{(i)}, w^{(i)}\}$, e.g.,

$$p(x) \approx \sum w_0^{(i)} \delta_{x^{(i)}}(x)$$

- Sequentially build up approximation to $\pi(x_{1:t}|y_{1:t}, \theta)$ one step at a time.



Parameter estimation

SMC provides an unbiased estimate of the marginal likelihood

$$\pi(y_{1:T}|\theta) = \pi(y_1|\theta) \prod_{t=2}^{T} \pi(y_t|y_{1:t-1}, \theta)$$

when we substitute the estimate

$$\tilde{\pi}(y_t|y_{1:t-1}, \theta) = \frac{1}{M} \sum w_t^n$$

for $\pi(y_t|y_{1:t-1}, \theta)$.

We can then use these estimates in a pseudo marginal scheme such as PMCMC (Andrieu *et al.* 2010) to estimate

$$\pi(\theta, x_{1:T}|y_{1:T})$$

and

$$\pi(\theta|y_{1:T})$$

SMC^2

We've found that SMC^2 (Chopin *et al.* 2011) works well for our problem

Basic idea:

- Introduce M parameter particles $\theta_1, \ldots, \theta_M$
- For $t = 1, \ldots, T$
  - For each θ_i run a particle filter targeting $\pi(X_{1:t}|y_{1:t}, \theta_i)$
  - Recalculate all the importance weights and resample if necessary

Note that to avoid particle degeneracy, it is still usually necessary to run a PMCMC sampler targeting $\pi(\theta, X_{1:t}|y_{1:t})$ at each resampling step.
We also use Brownian bridge proposals to guide the particles and reduce degeneracy.

This takes 3-4 days on a standard server, or 4-6 hours on a GPU (2500 processors), and took several months to code.

SMC^2

We've found that SMC^2 (Chopin *et al.* 2011) works well for our problem

Basic idea:

- Introduce M parameter particles $\theta_1, \ldots, \theta_M$
- For $t = 1, \ldots, T$
    - For each θ_i run a particle filter targeting $\pi(X_{1:t}|y_{1:t}, \theta_i)$
    - Recalculate all the importance weights and resample if necessary

Note that to avoid particle degeneracy, it is still usually necessary to run a PMCMC sampler targeting $\pi(\theta, X_{1:t}|y_{1:t})$ at each resampling step.
We also use Brownian bridge proposals to guide the particles and reduce degeneracy.

This takes 3-4 days on a standard server, or 4-6 hours on a GPU (2500 processors), and took several months to code.

- We used about 10^8 simulator runs!

Age model

Can we also quantify chronological uncertainty?

Age model

Can we also quantify chronological uncertainty?

Target

$$\pi(\theta, T_{1:N}, X_{1:N} | y_{1:N})$$

where $T_{1:N}$ are the times of the observation $Y_{1:N}$, which were previously taken as given.

- Inferring 2400 state values, 800 ages, 15 parameters

Age model

Can we also quantify chronological uncertainty?

Target

$$\pi(\theta, T_{1:N}, X_{1:N} | y_{1:N})$$

where $T_{1:N}$ are the times of the observation $Y_{1:N}$, which were previously taken as given.

- Inferring 2400 state values, 800 ages, 15 parameters

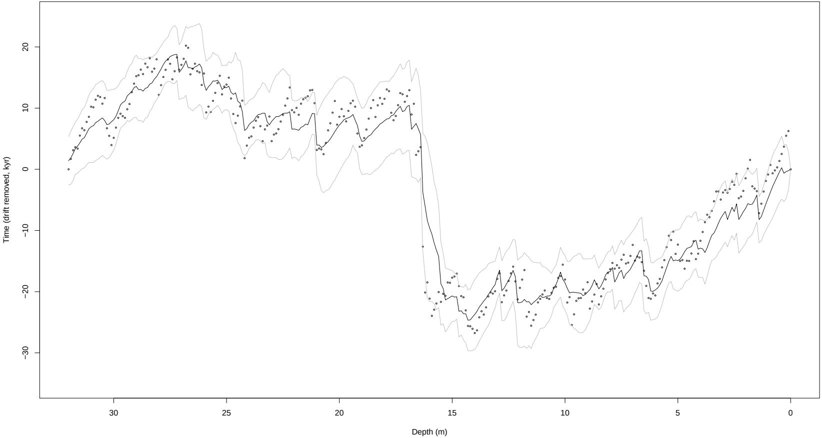
Propose a simple age model for sediment accumulation:

$$\mathrm{d}H = -\mu_s \mathrm{d}T + \sigma \mathrm{d}W$$

Where H is the depth in the core relating to time T

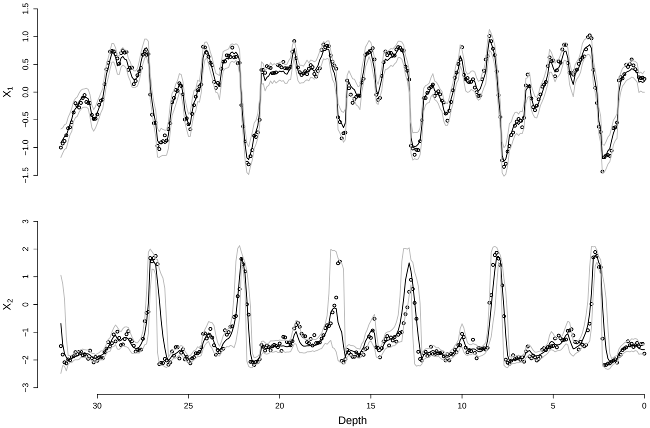
Simulation study results - age vs depth (trend removed)

Dots = truth, black line = estimate, grey = 95% CI



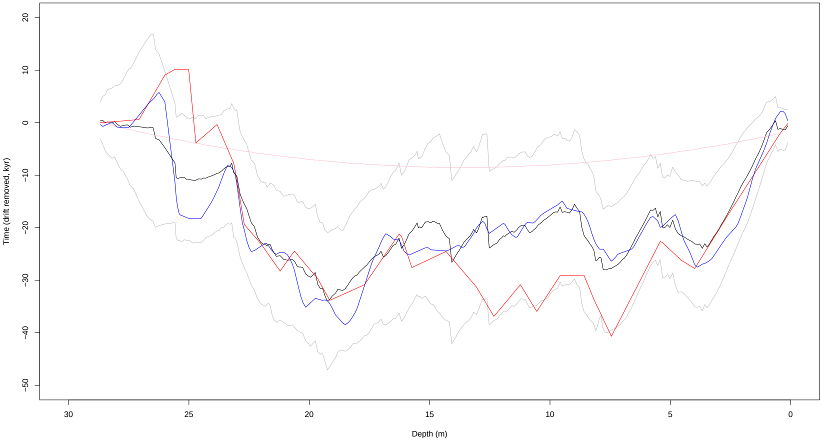
Simulation study results - climate reconstruction

Dots = truth, black line = estimate, grey = 95% CI



Results for ODP846 - age vs depth (trend removed)

Black = posterior mean, grey = 95%CI, red = H07, blue = LR04



It would be *difficult* to reproduce the method in LR04. We get a similar answer, using a single core, with complex joint uncertainty estimates.

Conclusions

- The data do contain enough information to discriminate between models and fit parameters
- Monte Carlo methodology and computer power are now sufficiently advanced that we can tackle the joint reconstruction, age model, and model selection problems for simple models
  - Proxy combination
  - but it remains computationally expensive ($\sim$ 1 week to compute per model here)
- Independently calculating age estimates and fitting climate models may be a bad idea (particularly if we ignore uncertainties)

Conclusions

- The data do contain enough information to discriminate between models and fit parameters
- Monte Carlo methodology and computer power are now sufficiently advanced that we can tackle the joint reconstruction, age model, and model selection problems for simple models
    - Proxy combination
    - but it remains computationally expensive ($\sim$ 1 week to compute per model here)
- Independently calculating age estimates and fitting climate models may be a bad idea (particularly if we ignore uncertainties)

Still to do/issues:

- ...

Conclusions

- The data do contain enough information to discriminate between models and fit parameters
- Monte Carlo methodology and computer power are now sufficiently advanced that we can tackle the joint reconstruction, age model, and model selection problems for simple models
  - Proxy combination
  - but it remains computationally expensive ($\sim$ 1 week to compute per model here)
- Independently calculating age estimates and fitting climate models may be a bad idea (particularly if we ignore uncertainties)

Still to do/issues:

- . . .

Thank you for listening!

Bayes factors

Advantages:

- Can provide evidence for and against a model
- Penalises for complexity (Occam's razor).
- Asymptotic consistency

Disadvantages

- Hard to calculate
- Sensitive to choice of prior
- Integrated likelihood may not be desirable treatment
  - predictive evaluation via scoring rules? (not p-values)