Climate Dynamics (2002) 19: 371–381
DOI 10.1007/s00382-002-0241-0

J.C. Hargreaves · J.D. Annan

Assimilation of paleo-data in a simple Earth system model

Received: 30 August 2001 / Accepted: 18 March 2002 / Published online: 13 June 2002

Abstract We describe the use of a Monte Carlo Markov Chain (MCMC) method based on Bayes' Theorem and the Metropolis-Hastings algorithm for estimation of model parameters in a climate model. We use the model of Saltzman and Maasch (1990). This is a computationally simple model, but with seven free parameters and substantial non-linearity it would be difficult to tune with commonly used data assimilation methods. When forced with solar radiation, the model can reproduce mean ocean temperature, atmospheric CO_2 concentration and global ice volume reasonably well over the last 500 ka. The MCMC method samples the multivariate probability density function of model parameters, which makes it a powerful tool for estimating not only parameter values but also for calculating the model's sensitivity to each parameter. A major attraction of the method is the simplicity and the ease of the implementation of the algorithm. We have used cross-validation to show that the model forecast for the next 50–100 ka is of similar accuracy to the hindcast over the last 500 ka. The model forecasts an immediate cooling of the Earth, with the next glacial maximum in around 60 ka. An anthropogenic pulse of CO_2 has a short-term effect but does not influence the model prediction beyond 30 ka. Beyond 100 ka into the future, the model ensemble diverges widely, indicating that there is insufficient information in the data which we have used to determine the longer term evolution of the Earth's climate.

J.C. Hargreaves (✉) · J.D. Annan
Proudman Oceanographic Laboratory,
Bidston Observatory Prenton, Merseyside,
CH43 7RA, United Kingdom
E-mail: jules@jamstec.go.jp

J.C. Hargreaves · J.D. Annan
Frontier Research System for Global Change,
3173-25 Showa-machi, Kanazawa-ku,
Yokohama, Kanagawa 236-0001, Japan

1 Introduction

A considerable amount of information about the past climate of the Earth is available. The climate proxies obtained from ice cores and marine, coastal and terrestrial sediment cores now provide information on a wide range of physical, chemical and biological processes. These data have the potential to aid greatly in the improvement of existing models of the Earth's climate by constraining the values of poorly known parameters on which a range of climate models depend. However, for complex numerical models, many well known parameter estimation methods cannot easily be implemented, due in part to the massive computational demands of optimal assimilation schemes, and in part to technical difficulties such as extreme non-linearity and discontinuities which may prohibit or severely hinder the application of classical methods.

We address the problem of modelling the Earth's climate on the glacial-interglacial time scale. Any model designed to simulate a number of ice age cycles, must, necessarily, be numerically very simple in comparison with the highly sophisticated and computationally demanding coupled ocean–atmosphere GCMs which are widely used today to study the Earth's climate over time scales of up to 100 years. Several computationally simple models have been developed and used to give hindcasts and forecasts of the Earth's climate (e.g. Imbrie and Imbrie 1980; Saltzman and Verbitsky 1995; Paillard 2001). These have, however, generally been somewhat hampered by the absence of an optimal method of parameter estimation and a lack of objective validation and error estimation, which makes it hard to assess their reliability and accuracy. We present here a method by which these questions can be addressed. We use the very simple prognostic model of Saltzman and Maasch (1990, hereafter SM). The model is highly parametrised and computationally very fast, lending itself well to a flexible but computationally demanding data assimilation

scheme such as the Monte Carlo Markov Chain method described by Harmon and Challoner (1997). The advantage of using such an assimilation scheme is that it can be used to generate a large amount of information about the model behaviour, and it is powerful enough to generate the optimal (probabilistic) solution even when applied to a highly nonlinear model. It is also extremely simple to implement, in contrast to (for instance) variational methods based on an adjoint model.

Using this method it is possible to derive the joint probability density function of the model parameters, which indicates how well the model can be constrained by the data. In addition we have used the method in combination with a cross validation technique to investigate the accuracy of the model predictions.

The rest of the study is organised as follows. In Sect. 2 and 3 respectively the model and the assimilation schemes are described. Following this the results of the assimilation are described with subsections on the overall results from the ensemble; the probability densities of the parameters; and the analysis of the forecasts. These sections are followed by the conclusions.

2 The model

The model of SM is a one point global model with three variables and seven free parameters which has been used to simulate the glacial–interglacial cycling over the last 0.5 Ma. Fundamental to the model is the assumption that all three variables (ice mass, atmospheric CO_2 concentration, and mean ocean temperature) can be written as a sum of an average value over scales of the order of millions of years ($\bar{X}_i$), plus an average over hundreds of years (X_i'). Therefore each variable (X_i) can be represented by the form

$$X_i = \bar{X}_i + X_i' \quad . \tag{1}$$

For modelling climate variation over the last 0.5 Ma, $\bar{X}_i$ is taken as constant, but extending the model for more than 0.5 Ma would require assumptions be made regarding the slow variation of $\bar{X}_i$.

The model is driven by the changes in summer insolation (Milankovitch forcing) at high latitudes caused by precession of the Earth's axis, variation in the eccentricity of the Earth's orbit and variation of the angle of tilt of the Earth with respect to its orbit. The precession is caused by the Sun and the other two variations are caused by the gravitational influence of the other planets in the solar system. The algorithm used was that of Berger (1978).

SM use previous analyses and various assumptions to postulate basic forms for the temporal derivatives of the three model variables, thus producing the three equations which can be integrated through time.

The basic assumption for the ice mass equations is that variations in ice mass depend on the mean summer surface temperature at high latitudes. The functional form chosen is based on the results of sensitivity analyses with general circulation models and statistical–dynamical models. The form of the ocean temperature equation is based on the assumption that the dominant influence on global mean water temperature is the ice volume in polar regions. The rate of temperature change is assumed to be linear with ice volume. The CO_2 equation is the most complex of the three equations being based on previous work (Saltzman and Maasch 1988). The relatively complex functional relationship contains six constants and is nonlinear in terms of temperature and CO_2.

The three equations are reproduced from the SM paper below as Eqs. (2), (3) and (4). I, μ, θ are the model variables, ice, CO_2 and ocean temperature respectively.

$$\frac{dI}{dt} = \alpha_1 - \alpha_2 c\mu - \alpha_3 I - \alpha_2 \kappa_\theta \theta - \alpha_2 \kappa_\theta R(t) \tag{2}$$

$$\frac{d\mu}{dt} = \beta_1 - (\beta_2 - \beta_3\theta + \beta_4\theta^2)\mu - (\beta_5 - \beta_6\theta)\theta \tag{3}$$

$$\frac{d\theta}{dt} = \gamma_1 - \gamma_2 I - \gamma_3 \theta \tag{4}$$

$\kappa_\theta = (1/b)(\delta\tilde{\tau}/\delta\theta)$,where $\tilde{\tau}$ is the zonal mean surface air temperature at high latitude in summer. b and c are constants in the functional relationship between $\tilde{\tau}$ and μ, and their values have been estimated by modelling experiments. See SM for more details of the derivations. The constants $\alpha_{1,2,3}$, $\beta_{1,2,3,4,5,6}$ and $\gamma_{1,2,3}$ can be distilled into seven unknowns. After variable substitutions and normalisation SM produced the following set of equations:

$$\dot{X} = -X - Y - vZ - uR(t^*) \tag{5}$$

$$\dot{Y} = -pZ + rY + sZ^2 - wYZ - Z^2Y \tag{6}$$

$$\dot{Z} = -q(X + Z) \tag{7}$$

where the model variables are non-dimensional proxies for ice mass (X), CO_2 concentration (Y), and deep ocean temperature (Z); R is the non-dimensional variation of annual solar radiation at 65°N; $(\dot{\cdot}) = d()/dt^*$ and t^* is a rescaled time variable ($1000t^* = t$ where t is time in years). The seven coefficients p, q, r, s, u, v and w are all constants, defined by SM as 1.0, 2.5, 0.9, 1.0, 0.6, 0.2 and 0.5 respectively. The initial conditions for X, Y and Z are taken to be –1.0, 0.2 and 1.0 respectively. The model is integrated with a time step of 100 years, for which numerical errors are insignificant. As shown by SM these values produce reasonable agreement with ice mass data over last 0.5 Ma (taken from the SPECMAP data set (Prentice and Mathews 1988)) and CO_2 concentration measurements over the last 0.175 Ma (from the Vostok ice core). In fact the modelled CO_2 concentration agrees fairly well with the more recent extension of the Vostok core to 420 ka (Petit et al. 1999), which is one reason why we selected this particular model for our experiments out of the family of similar simple models proposed by Saltzman and various co-authors (e.g. Saltzman and Sutera 1987; Saltzman and Verbitsky 1993). For example, the model output generated by Saltzman and Verbitsky (1995) does not appear to compare so well with the CO_2 data which it anticipated. Other independently devised models of comparable complexity such as that of Imbrie and Imbrie (1980) or Paillard (1998) could equally be treated by the techniques we describe here. The review of Paillard (2001) compares hindcasts and forecasts from these two models and one other, but with an absence of statistical analysis or confidence intervals it is hard to know how much credibility they have. For consistency with SM we used the same proxy data for ice mass and deep ocean temperature that they described but we also used the recently extended Vostok ice core for the CO_2 data. These data do not perhaps all represent the state of the art (Shackleton 2000), but are nevertheless sufficient for our purpose of investigating the potential value of data assimilation techniques in simple climate models.

3 The data assimilation technique

The general goal of data assimilation is to use data in order to improve the performance of numerical models. In this context, that means finding parameter values (and initial conditions) for which the model output most closely resembles observational data over the last 0.5 Ma. There are several well established parameter optimisation techniques for this kind of problem based on variational methods (the use of an adjoint model is common), and at least in the case of a well posed linear problem, there is a unique solution which all 'optimal' assimilation methods will find (e.g. Lorenc

1986). However, most data assimilation techniques are designed for linear dynamics, and may fail to work correctly in the nonlinear case (Gauthier 1992; Miller et al. 1994). Moreover, since our data have errors and the truth is imperfectly known, the problem should be considered as a probabilistic one. That is, the correct solution is not a unique set of parameters but rather a set of parameter estimates along with confidence intervals (and perhaps even a covariance matrix). Standard variational methods do not naturally produce this probabilistic information, instead merely generating the optimum parameter values. They also require significant effort to implement. In contrast, the Monte Carlo Markov Chain (MCMC) method based on the Metropolis-Hastings algorithm is simple to implement, very flexible (including the ability to handle nonlinear models) and naturally produces an ensemble of parameter estimates from which statistical properties can readily be derived and probabilistic forecasts made. The substantial handicap of this method, which restricts its application to the computationally cheapest of models, is the massive computational demand which it imposes. However, this is not a significant restriction for this work since the SM model is very fast.

We have implemented the MCMC technique as described in Harmon and Challoner (1997). A more thorough theoretical background to the algorithm is given by Chib and Greenberg (1995). The goal of the technique is to sample the joint posterior probability density function (PDF) of the parameter space given the observations and any prior information concerning parameter values. This is achieved by performing a random walk through parameter space using the Metropolis-Hastings algorithm (Metropolis et al. 1953) described later. This random walk has the important and useful property that the density of parameter vectors visited during the walk converges to the joint PDF.

Harmon and Challoner (1997) experimented with the ecosystem model of Fasham et al. (1990), in which biological equations (with five or ten rate parameters to be estimated via the MCMC method) are forced by (and strongly tied to) a seasonal physical cycle in the ocean. Our implementation of the MCMC technique is very similar to their method which they describe fully and clearly, so only a brief summary is given below. We allowed the initial conditions (i.e. $X(t)$, $Y(t)$ and $Z(t)$ for $t = 0.5$ Ma BP) to vary, along with the seven free parameters in the model, giving a total of ten degrees of freedom. The random walk was continued until the parameter estimates had converged, which required around 3×10^7 iterations. This is rather more than the 10^6 integrations that Harmon and Challoner (1997) used, but the model we used exhibits much more complex behaviour than their's in that it generates five successive cycles which are not timed by the external forcing but are largely internally generated. We also used real data in contrast to their 'identical twin' set up (in which a synthetic data set is generated by an earlier model run, with random noise added to simulate observational error, for the purpose of testing the assimilation technique). The identical twin technique is widely used as a first test of assimilation methods, but is rather a weak test since it is virtually guaranteed that a good solution exists (there is no model error, and the observational errors are known precisely). However, in our example (as in most practical applications), the model remains imperfect even after tuning, and the implications of this are discussed below. It is of course clear from the large number of iterations required that this assimilation technique is only suitable for the least computationally demanding models. However the simplicity of implementation makes it an attractive option for tuning these small models, and it should be noted that the number of model runs performed is many orders of magnitude lower than would be necessary for a brute force attempt to map out parameter space. A multifactorial experiment with only 10 values per parameter would require a very costly 10^{10} model runs and still be inadequate for exploring parameter space.

The MCMC method is based on a Bayesian approach to parameter estimation. Given a data set $\mathbf{X}$ and a model which requires a set of parameters Φ, our goal is to obtain the conditional probability density function $f(\Phi|\mathbf{X})$ of the parameters given the data. Bayes' Theorem tell us that this can be expressed in terms of the probability density function (or likelihood) of the data given the model parameters $f(\mathbf{X}|\Phi)$ and any prior knowledge of the parameters $f(\Phi)$:

$$f(\Phi|\mathbf{X}) \propto f(\mathbf{X}|\Phi) f(\Phi) \tag{8}$$

Working in log-likelihoods for convenience, the likelihood on the right hand side of this equation can be readily calculated via

$$\ln f(\mathbf{X}|\Phi) = -\frac{1}{2}\sum_i \frac{(x_i - x_i')^2}{\sigma_i^2} + K \tag{9}$$

where the sum is over observations x_i and corresponding model output x'_i. σ_i are the observational errors (assumed Gaussian and independent). K is an unknown constant which is not required by the algorithm. This sum also defines the cost function which measures the quality of the fit of the model to the data.

The Metropolis-Hastings algorithm generates a chain of parameter sets $\{\Phi_i\}$, $i = 1, N$ of arbitrary length. The algorithm performs a random walk through parameter space based upon a simple iterative procedure. Each iteration consists of a trial step, in which a new set of parameters are generated, followed by an acceptance–rejection step, in which the decision is made whether to move to the new set of parameters, or to stay with the current set. Given a current set of parameters Φ_i, a new trial set of parameters Φ' is generated by adding a small random perturbation to each of the elements of Φ_i. The acceptance–rejection step is performed by calculating the ratio of probability densities $\frac{f(\Phi'|\mathbf{X})}{f(\Phi_i|\mathbf{X})}$ (using the equation described earlier) and comparing this number to a further single random sample drawn from the uniform distribution over the interval [0, 1]. If the probability ratio exceeds this last random sample, then Φ_{i+1} is set equal to Φ' (acceptance), otherwise Φ_{i+1} is set equal to Φ_i (rejection). This formulation implies that if the random perturbation generates an improved set of parameters, this is always chosen as the next step in the chain, while a worse set of parameters is accepted with a nonzero probability which decreases as the model-data misfit increases. The iteration is continued until the distribution of $\{\Phi_i\}$ has converged sufficiently. Each iteration requires a single model run (to calculate the probability density), which dominates the cost of the algorithm. When convergence has been achieved, the sequence $\{\Phi_i\}$ samples the joint PDF of parameters.

The time taken to converge can vary widely between implementations, and is hard to predict in advance. The selection of the random trial perturbation can be a critical step in determining the efficiency of the algorithm, since large perturbations will almost always be rejected and small perturbations will take too long to explore parameter space even when they are usually accepted. A few preliminary experiments were used to adjust the typical perturbation size so as to ensure an acceptance rate of around 25%, this being a generally recommended figure. This is the only detail of the algorithm which requires any problem specific tuning, and the method is otherwise very straightforward to implement.

This assimilation technique is quite similar to simulated annealing (Pincus 1970; Azencott 1992) which has also been applied to a variety of geophysical problems (Bennet and Chua 1994; Evensen 1994; Matear 1995). However the MCMC approach starts from the basic premise that, rather than searching for a unique optimum as simulated annealing (in its standard form) does, it is more appropriate to consider the problem as intrinsically probabilistic in nature and to work with the PDF of parameter values which are consistent with both the data and our prior knowledge.

4 Results

4.1 Ensemble results

The results from the Monte Carlo ensemble are summarised in Fig. 1. In this experiment, all the data were assimilated, and for interest the model was integrated forward beyond the present day for 250 ka. The parameter values derived by SM were used as the first guess for the run (see Table 1 for the parameter values

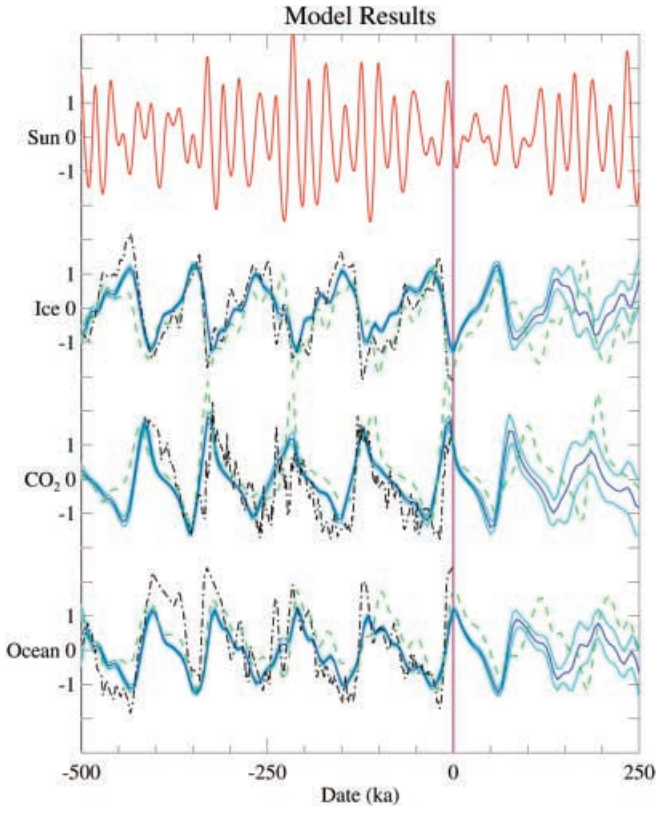

Fig. 1. Run with 500 ka of data assimilated. The *vertical magenta line* shows when assimilation was halted. The *red line* at the *top* is the solar insolation forcing. The *green dashed lines* are the results from using the SM parameters. The *black dot–dashed lines* are the data. The *dark blue lines* show the mean of the ensemble and the *light blue lines* show the one standard deviations of the ensemble

and cost for this run). The figure shows the run with the SM parameters (green dashed lines), the data (black dot–dashed lines), the ensemble mean (dark blue lines) and the one standard deviation width of the ensemble (light blue lines) for each of the three model variables.

Several conclusions can be drawn. Most obviously, the model ensemble provides a reasonable fit to the large time scale variation of the data. However, the model output is far too smooth to fit the fine detail accurately.

Over the forecast region, it can be seen that the ensemble width increases, indicating that as we look further into the future, the prediction becomes less certain. In theory, the width of the ensemble should be comparable with the model-data misfit, but this is not the case in Fig. 1; the ensemble is much too narrow. The fundamental reason for this discrepancy is that the assimilation technique (in common with many others, when the model is used as a 'strong constraint') makes the assumption that the model is structurally perfect, when in fact this is clearly far from the case. This erroneous assumption causes false confidence in the ensemble. Therefore the ensemble width cannot be used directly as a measure of expected error but can still indicate qualitatively where and how the predictability changes. The ensemble does indicate an immediate and rapid cooling of the Earth, with a new glacial maximum occurring around 60 ka into the future. At around 100 ka in the future and beyond, the ensemble width is growing, indicating a loss of predictability and a degeneration of the forecast to 'climatology'. There is no return of skill, which might be expected if the climate system was directly controlled by specific events in solar

Table 1. Parameter values: first guess (SM values); mean and inter-quartile range of distributions; best fit run

Parameter	SM value	Results from parameter distributions				
		Mean	25%	50%	75%	Best fit
X_0	–1.0	–0.77	–0.91	–0.74	–0.60	–0.62
Y_0	0.2	0.28	0.16	0.25	0.38	0.11
Z_0	1.0	0.43	–0.17	0.46	1.03	–0.36
p	1.0	0.79	0.72	0.77	0.84	0.82
q	2.5	4.97	4.24	4.83	5.62	4.24
r	0.9	0.90	0.86	0.89	0.94	0.95
s	1.0	0.57	0.53	0.57	0.62	0.53
u	0.6	0.30	0.27	0.30	0.32	0.32
v	0.2	0.07	0.00	0.07	0.14	0.02
w	0.5	0.83	0.75	0.83	0.91	0.66
α_1(kg a^{-1} × 10^{16})	1.63	0.51	1.46	1.52	1.58	1.46
α_2 (kg a^{-1} × 10^{16})	1.24	1.24	–	–	–	1.24
α_3 (a^{-1} × 10^{-4})	1.00	1.00	–	–	–	1.00
$\beta_1 + \bar{\mathbf{F}}^*_\mu$ (ppm a^{-1} × 10^{-1})	10.6	9.60	9.49	9.60	9.71	9.80
β_2(a^{-1} × 10^{-3})	4.27	4.04	3.99	4.04	4.09	4.16
β_3(ppm(°C a)$^{-1}$ × 10^{-3})	1.89	1.84	1.83	1.84	1.85	1.87
β_4((°C a)$^{-1}$ × 10^{-4})	2.04	2.04	–	–	–	2.04
β_5(ppm(°C a)$^{-1}$ × 10^{-1})	4.61	4.30	4.26	4.29	4.32	4.33
β_6((°C^2 a)$^{-1}$ × 10^{-2})	4.92	4.70	4.67	4.69	4.72	4.67
γ_1(C a^{-1} × 10^{-3})	1.99	3.97	3.38	3.85	4.48	2.28
γ_2(°C(kg a)$^{-1}$ × 10^{-23})	1.35	2.68	2.28	2.60	3.03	2.28
γ_3(a^{-1} × 10^{-4})	2.50	4.98	4.24	4.83	5.62	4.24
κ_θ(°C^{-1} × 10^{-2})	3.01	1.04	0.08	1.10	2.15	0.27
κ_R(m^2 W^{-1} × 10^{-3})	3.51	5.22	4.75	5.24	5.68	5.63
Cost	473	284	233	241	300	226

forcing. The forecasting ability of the model is considered further in Sect. 4.3.

4.2 The probability densities of the parameters

In their paper SM estimated the parameter values (p, q, r, s, u, v, w) using various physically based arguments and results from other models. No formal error estimates are provided either for these parameter estimates or for the parameters of the underlying physical equations which are subsequently calculated by inverting the normalisation and variable substitutions. There is, therefore, little discussion as to what extent the parameters could be allowed to vary while still providing reasonable consistency with the data and no

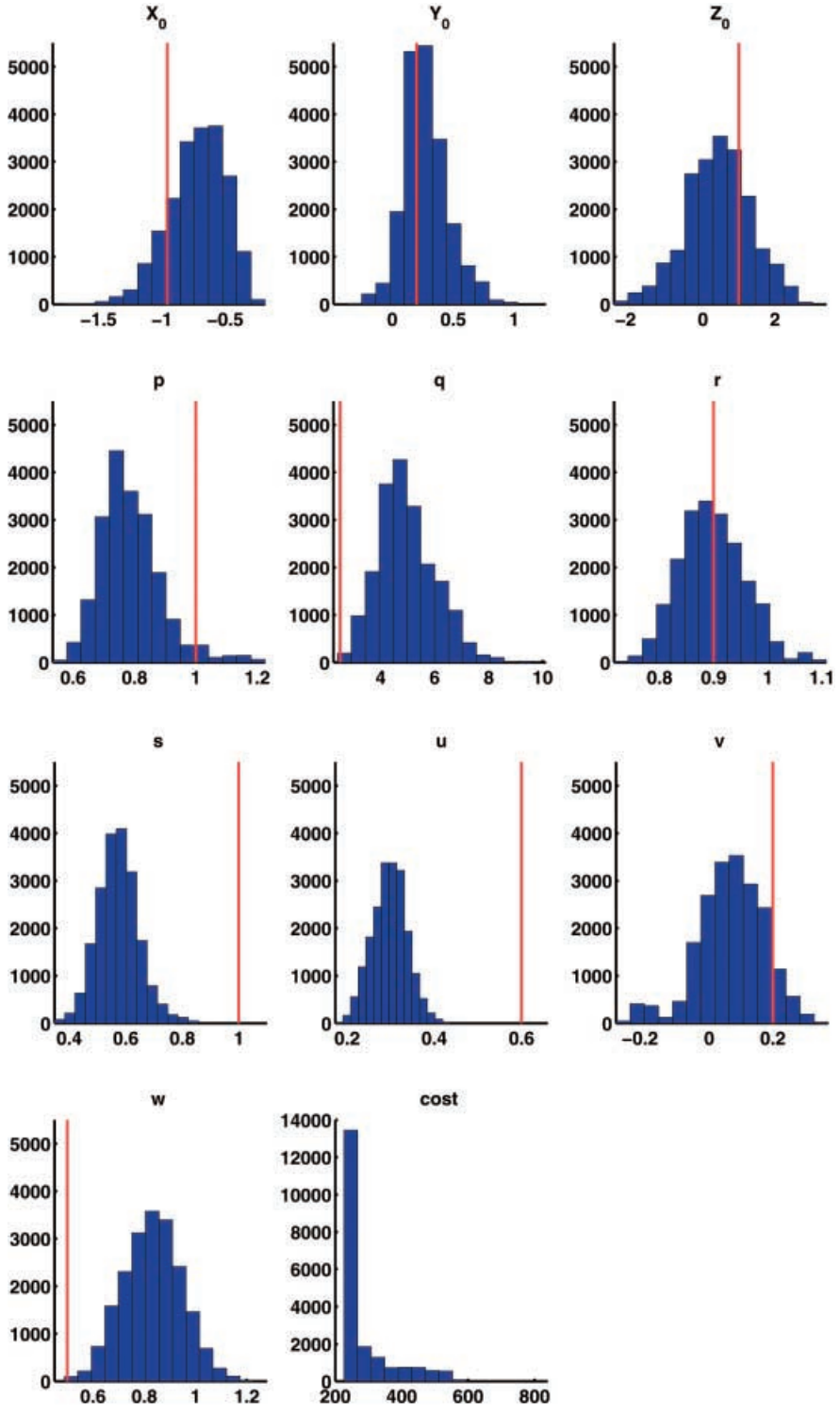

Fig. 2. Distributions of parameters p, q, r, s, u, v, w and of the cost of the ensemble of runs. The SM values are shown in *red*

indication of which parameters the model is more sensitive to.

The MCMC method samples the joint PDF of the parameters so the results obtained here make possible an assessment of how well the physical quantities in the original equations are constrained by the model and data. We do this here by looking at the distributions of each of the model parameters.

In Fig. 2 we illustrate the PDF of the parameters obtained by recording every 1000th run of the last 20 million runs (ignoring the initial transient where the MCMC algorithm forgets the initial guess) and plotting

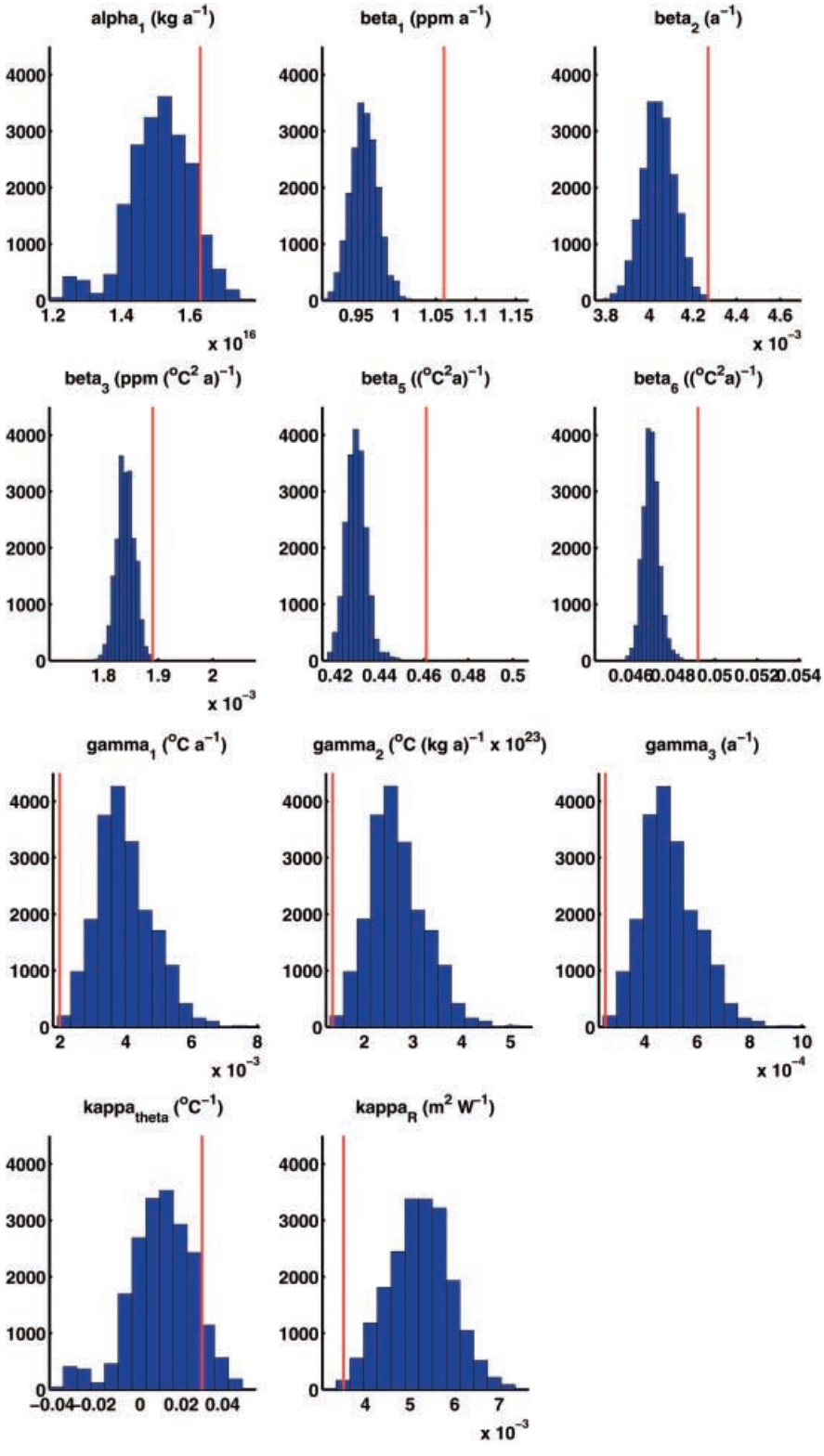

Fig. 3. Parameter distributions for the original equations. The SM values are shown in *red*

histograms of the parameter distributions so obtained. Just as the parameters p, q, r, s, u, v, w were derived from the original parameters in Eq. (2) (3) and (4) so the procedure can be reversed and the original parameters can be recalculated from our distributions (see the SM paper for the details of this straightforward algebraic exercise). Figure 3 shows the PDFs obtained by the MCMC method for the un-normalised parameters.

Table 1 compares the results of SM with the results from our ensemble of runs and the best fit run of the ensemble. For consistency throughout this study we recalculated the SM original parameter values using the constants and scaling parameters as quoted in the SM paper. Apart from a couple of typographical errors in SM [the values of β_3 and β_6 in Eq. (38) have been exchanged, and the last part of SM Equation (36) should read $v = (\kappa_\theta/c)\ (\lambda_3\lambda_2^{-1})$] there remain only small differences in the values of the un-normalised parameter values we obtain compared to those quoted in SM. These differences are sufficiently small that it is possible that the use of more significant figures in the constants or scaling parameters than those quoted by SM in their paper may be the cause.

The inter-quartile ranges (IQRs) of the parameter distributions are shown in Table 1. IQR is preferred over standard deviation as a more robust statistic for non-Gaussian variables.

X_0, Y_0, and Z_0 are the starting positions for 500 ka ago. These values were not quoted in the SM paper since they are not required for derivation of the physical parameters. They do, however, provide extra degrees of freedom for the model so are quoted here. The SM values were obtained from Saltzman (personal communication 1999).

SM state that all the parameters (p, q, r, s, u, v and w) should be positive. In our calculations we did not force this condition but allowed all the parameters the freedom to be negative. As shown in Fig. 2 only parameter v shows any negative values. A negative value of v in Eq. (5) implies, rather implausibly, that a cold ocean causes ice to melt. About one fifth of the runs lie in this range. This result can be understood by considering the mathematical consequences of Eq. (5) and (7). Looking at Eq. (7) and noting that, for our results, parameter q is always large (see Table 1 and Fig. 2), it can be seen that the ocean temperature Z relaxes rather rapidly to $-X$, i.e. the ocean temperature and ice mass have approximately equal and opposite magnitudes throughout the model integration (this can also be confirmed by examining the various plots of model outputs which are shown). This means that Eq. (5) can be approximated by

$$\dot{X} \simeq -(1+v)X - Y - uR(t^*) \tag{10}$$

and we see that the influence of v is merely to adjust the rate at which X relaxes to its mean, with the overall rate varying between 1 and 1.14 across v's inter-quartile range and dropping slightly below 1 for negative v. So, even though a negative value of v would imply an implausible physical feedback, this model is rather insensitive to this effect.

The other parameter distributions are skewed with truncation above zero. This is an encouraging result for the model validity. X_0, Y_0 and Z_0 may be either positive or negative since these are the initial conditions for the model run. Of these, Z_0 shows a particularly large range of variability in the model results so its value is not well defined by the model and data. This is explained again by the rapid relaxation of Z to $-X$, combined with the lack of data for Z in the early part of the model run.

The last histogram in Fig. 2 which shows the range of the cost function (normalised sum of squared errors) over the model runs is of interest since it demonstrates that more than 50% of the model runs are in the area of very low cost, below about 250. This means that the model is spending the majority of the runs close to minima of the cost function, even though the parameter values are varying widely. For comparison the run using the SM parameters gives a cost of 473.

The seven parameters and three initial conditions define a cost function in 10 dimensions which, due to the high nonlinearity of the problem, can have a complex shape and many local minima. This is difficult to visualise, but a typical example is given in Fig. 4. This shows the cost of the model run with all values at the best fit value shown in Table 1 but p is varied over its inter-quartile range. The lowest minimum coincides with the best fit value in the Table, but this slice through the cost function has a complex shape with many local minima. Note that for any value of p in this inter-quartile range, there is a set of parameter values that has very low cost, so the joint PDF must be considered in order to draw meaningful conclusions about any particular parameter, and there is little point in calculating the 'best' value and

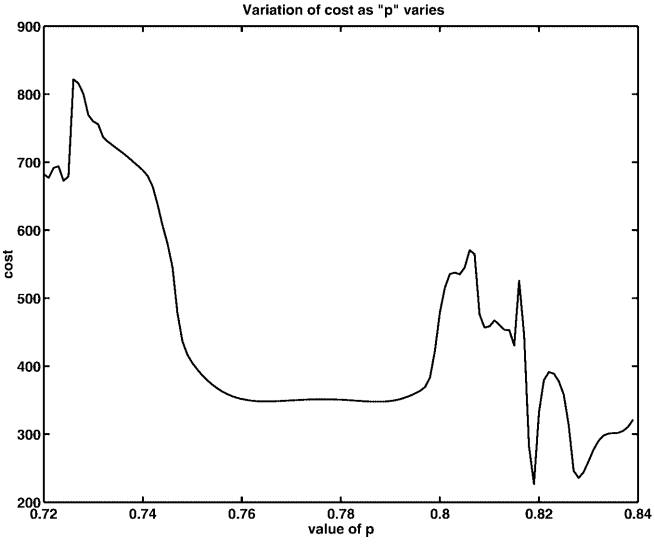

Fig. 4. Variation of cost as parameter 'p' varies through its IQR. All other parameters are set to their values at the best fit

IQR of a single parameter in isolation since these depend critically on the values of other model parameters, all of which are imperfectly known. Extending this to 10 dimensions makes it clear why sub-optimal methods of parameter estimation, or those based on linear approximations, may struggle to find good solutions.

Performing a model run using the mean values of each parameter does not generate output with a low cost. This is again because of the high complexity of the surface of the cost function, or equivalently, the extreme nonlinearity of the model. The output from the model run using mean parameter values differs greatly from the mean of the outputs of the model runs in the ensemble.

As mentioned, when calculating the parameters for the original (i.e. the un-normalised) equations we took the quoted SM values for all the constants apart from those seven parameters (p, q, r, s, u, v, w) calculated by the MCMC method. In order to produce parameter distributions which still sample the PDF it is necessary to transform the parameters for each run and then recombine them into distributions. The results from this procedure are shown in Fig. 3. The mean and IQR of the parameters are compared with the best fit solution and the SM original parameters in Table 1.

The secondary peak in α_1 and κ_θ is caused by the similar peak in negative values of v discussed previously, since the variation of these two parameters depends only on v. The variation in $\gamma_{1,2,3}$ all depend only on q which is why the distributions look very similar to each other, and to that of q itself. In contrast $\beta_{1,2,3,5,6}$ are functional combinations of r, w, p and s. β_3 depends only on w but the others all depend on combinations of the parameters which may explain why they appear to have a more Gaussian distribution (since a linear combination of non-Gaussian distributions usually tends towards a Gaussian shape).

Parameters α_1, $\beta_{1,2,3,5,6}$ are all quite tightly constrained although the original SM values still lie outside the IQRs except for β_6. Parameters $\gamma_{1,2,3}$ and $\kappa_{\theta,\mathrm{R}}$ are less well defined, and the means of the distributions diverge more greatly from the SM values. For example $1/\gamma_3$ is the time constant for the dissipation of the global properties of deep water which SM specified 'arbitrarily' to be 4 ka. Our IQR gives a value of between 1.8 and 2.4 ka. SM state that there has previously been debate over the positivity of γ_2. Our results indicate that a negative γ_2 is implausible given that this feature did not arise in any of the model runs. Note that we placed no specific bound upon this parameter. A positive value of γ_2 means that a rise in ice mass causes ocean temperature to decrease.

The two values κ_θ and κ_R are related to the change in air temperature associated with the change in ocean temperature and Milankovitch forcing respectively. If we assume that the parameter b is 18° as quoted by SM then the ratio of the change in air temperature to ocean temperature is not well constrained by the model with an IQR between 0.01 to 0.4, compared to the SM value of 0.6. With the same assumption of b, the ratio of the variation of air temperature change to Milankovitch forcing varies between 0.09 and 0.1 °C$(W\ m^2)^{-1}$ which is lower than the SM value of 0.2 °C$(m^2)^{-1}$.

4.3 Forecasting

4.3.1 Cross-validation testing

In general terms, the skill of a model forecast can often be directly measured by making the prediction, and then comparing to observations as they are made. By repeating over several forecasts (either through the passage of time or over different spatial sub-domains), a statistically reliable estimate of their accuracy can be calculated. However, this approach is impossible when forecasts are being made for many thousands of years into the future, and averaged over the entire Earth. In the identical twin scenario, it is also possible to generate many hypothetical data sets on which to test model performance, but we have already seen that our experimental setup is far removed from the identical twin case (ie the model contains substantial unknown structural errors) and so this approach is likely to be of limited use.

Instead, we use a standard cross-validation approach and test the ability of the model, when tuned on an initial subset of the data, to forecast the remainder. By using initial segments with lengths from 100 ka to 450 ka length (in 50 ka steps), a set of eight forecasts can be evaluated. Since the data are rather sparse in time and the model evolves quite smoothly, the RMS errors of the model are calculated in 50 ka width bins.

A typical output from the cross-validation exercise is shown in Fig. 5. In this example, only the data prior to 350 ka BP were assimilated. The model appears to generate a reasonable forecast for around 50 ka, and beyond 100 ka the fit between model and data is very poor. It is encouraging to see that the ensemble is also spreading by this point which indicates that the model forecast here is not well constrained by the data that was assimilated. This figure also shows that the model is not merely controlled by the Milankovitch insolation, as around the present day and beyond it is completely out of phase with the previously shown model output. Instead, the model contains its own internal dynamics which may be excited by the forcing but are not timed by it. Figure 1 in contrast shows the abilities of the model to hindcast data to which it is tuned (both by ourselves and SM) but contains no element of forecasting.

A summary of the nine cross-validation results are shown in Fig. 6. Each line indicates the error associated with a particular forecast interval as a function of the duration of data assimilated. So, for example, the blue circles indicate the RMS error when forecasting in the range 50–100 ka past the end of the assimilated data set (this cannot be calculated for the case in which 450 ka of

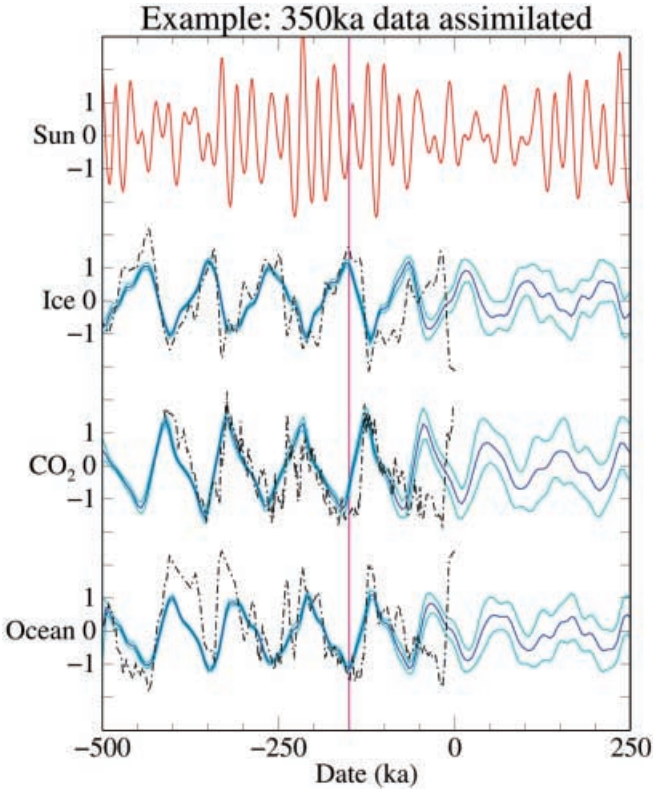

Fig. 5. 500 ka run with 350 ka of data assimilation. Key as for Fig. 1

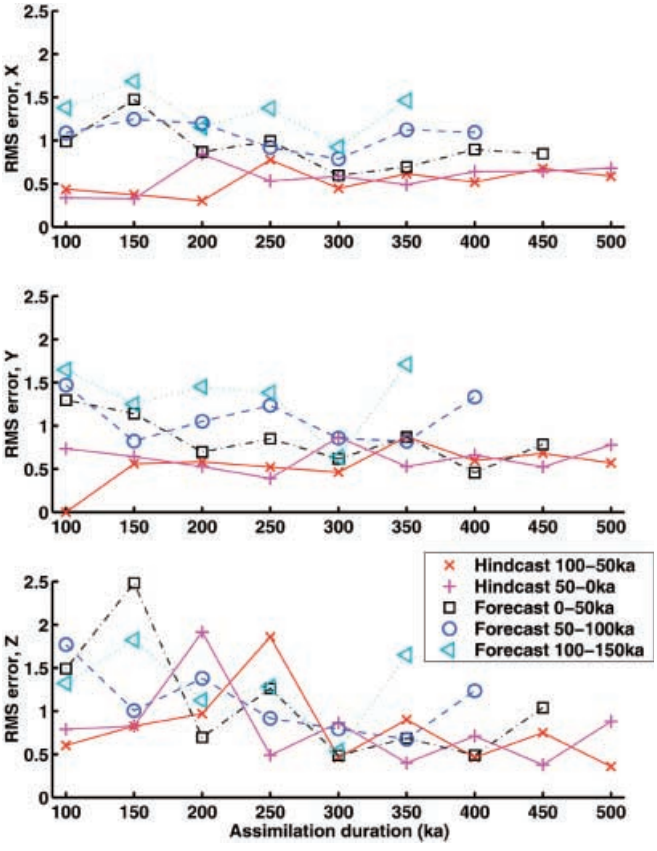

Fig. 6. RMS errors for hindcasts and forecasts when different amounts of data are assimilated

data were assimilated as this time interval is still in the future). Also shown for comparison are the RMS errors of the hindcasts. When very little data is assimilated, the forecast results are very unreliable (note that an RMS error of $\sqrt{2} \simeq 1.4$ indicates a complete lack of skill since both data and model forecast have a standard deviation of 1). As the volume of assimilated data increases, so does the forecast skill which is encouraging as it suggests that the parameters are being increasing tuned to 'good' values and not merely (over-)fitting the data which has been used. As the assimilation interval increases, the forecast skill for the 0–50 ka range appears to be converging to close to the hindcast skill, and there is also some indication of skill in the 50–100 ka range too. Longer range forecasts are clearly worse, although it is hard to evaluate their skill reliably since this can only be calculated when a short interval of data was assimilated. However the ensemble widths also increase beyond 100 ka forecasts which suggests that they will have low reliability.

This figure suggests that the skill of the forecast over 0–50 ka is hardly any worse than the hindcast, and there is some skill even as far as 100 ka into the future. Therefore, to the extent that we consider the hindcast in Fig. 1 to be a reasonable approximation to the truth, we should also believe that the forecast of the next glacial maximum is reliable.

The use of such simple climate models for generating forecasts is of course not new (e.g. Imbrie and Imbrie 1980; Saltzman and Verbitsky 1995), but in contrast to these previous papers, the use here of an optimal data assimilation method and an ensemble of model runs allows us for the first time to evaluate critically and objectively both the potential accuracy and also the useful duration of such forecasts.

4.3.2 Anthropogenic pulse

We have also integrated the model ensemble (generated by assimilating the entire data set) forward for 250 ka but added an extra source term to the atmospheric CO_2 equation corresponding to anthropogenic forcing. We have assumed that the anthropogenic influence consists of a large pulse of CO_2 generated by the burning of all available reserves of fossil fuels over the next few hundred years. The precise nature of the input does not influence our results, as we are not attempting to model short-term effects and the effects beyond one thousand years are ominated by the total carbon mass burnt (Lenton 2000).

The model forecasts are shown in Fig. 7. It can be seen that the initial transient decays rapidly, and has converged to the unperturbed case within 30 ka. The timing and strength of the next glacial maximum is not affected. By examining the dynamical equations, we can

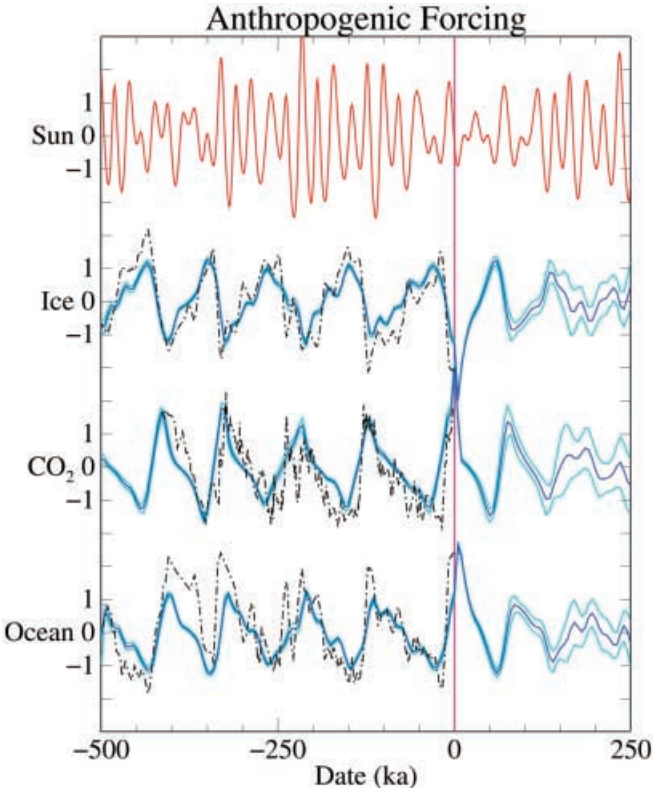

Fig. 7. Run with 500 ka of data assimilation and an anthropogenic pulse of CO_2. Key as for Fig. 1

explain why the timing of the next transition is so robust. With atmospheric CO_2 and ocean temperature both at peak values, the flux of CO_2 into the ocean is very large (according to the SM formulation, with either their original parameters or our optimised values). Although a warm ocean temperature would tend to reduce CO_2 solubility, their hypothesis (explored more fully in the earlier paper by Saltzman and Maasch 1988) is that changes in circulation patterns will offset and indeed outweigh this effect. When a large additional pulse of anthropogenic CO_2 is added, the dominant effect is to increase the negative flux of CO_2 until the unperturbed solution is reached.

Of course these results should be taken with a degree of scepticism: the perturbation takes the model variables well outside the range over which they are calibrated, and at such extremes other processes which have not been included may start to dominate. For example, the derivation of the ice mass equation approximated $\tanh(0.004\mu)$ by 0.004μ, where μ is atmospheric CO_2 concentration in ppm. This approximation becomes increasingly inaccurate for $\mu > 350$ ppm. However, similar criticisms could probably be levelled at all efforts to predict the future climate, given the current unprecedented anthropogenic perturbation to the climate system, and this does not (and should not) stop people from trying. Moreover, at least this simple model has proved itself capable of hindcasting, and forecasting (via the cross-validation exercise) with a meaningful degree of skill over intervals of order 50–100 ka.

5 Conclusions

The long time series of geological data sets are potentially highly valuable for use in combination with models to advance the understanding of the evolution of the paleoclimate. We have presented a form of data assimilation based on Monte Carlo principles and implemented it for the three variable model of Saltzman and Maasch (1990) using the Vostok ice core and two SPECMAP cores as proxies for CO_2, ocean temperature and global ice mass respectively.

Optimal data assimilation techniques are generally highly CPU-intensive, and this method is no exception to the rule. For this reason, large-scale GCMs generally rely on more limited tuning and parameter estimation methods. The simple SM model provide a good test case for experimenting with advanced techniques, and provides some useful evidence on the plausible ranges of various poorly known parameters. One particular advantage of the method used here is the ease with which it can be implemented.

Despite the large number of free parameters in the SM model, the MCMC method converged successfully and was stable and well-behaved in application. The previous application of this technique by Harmon and Challoner (1997) had not used experimental observations but rather concentrated on identical twin experiments, so this result was encouraging in itself.

Since the MCMC method samples the PDFs of the model parameters, we have been able to rederive the physical parameters underlying the model and also calculate the extent to which these parameters are defined by the model.

The predictive capability of the model was investigated using cross-validation and we have shown that the model can make useful predictions for approximately 50–100 ka. The very short time scale events (< 1000 years) are not present in the model dynamics so remain unmodelled and therefore unpredictable, but the broad shape of the ice age fluctuations is replicated reasonably well. Our implementation uses the model as a strong constraint, but model error creates a significant problem in interpreting the ensemble width. The ensemble has 'false confidence' in itself, having insufficient width in comparison to the true errors. We intend to develop the method in the future to allow for model errors by considering the model dynamics as a weak constraint.

The model predicts a new glacial maximum to occur at around 60 ka into the future. The input of an anthropogenic pulse of CO_2 was found to not affect this event significantly, although it had a larger influence over the short term. However, such a large pulse does take the model out of the range of its calibrated applicability so this result is most probably more

illustrative of the stability of the model rather than the stability of the climate.

Acknowledgements We are grateful to NERC for supporting this research, and also John Shepherd, Hyungmoh Yih and two referees for helpful comments on the manuscript.

References

Azencott R (1992) Simulated annealing. John Wiley

Bennet AF, Chua BS (1994) Open-ocean modelling as an inverse problem: the primitive equations. Mon Weather Rev 122: 1326–1336

Berger A (1978) A simple algorithm to compute long term variations of daily or monthly insolation. Tech Rep 18, Universitè Catholique de Louvain, Belgium

Chib S, Greenberg E (1995) Understanding the Metropolis-Hastings algorithm. The American Statistician 49(4): 327–335

Evensen G (1994) Inverse methods and data assimilation in nonlinear ocean models. Physica D 77: 108–129

Fasham MJR, Ducklow HW, McKelvie SM (1990) A nitrogen-based model of plankton dynamics in the oceanic mixed layer. J Mar Res 48: 591–639

Gauthier P (1992) Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. Tellus 44A: 2–17

Harmon R, Challoner P (1997) A Markov chain Monte Carlo method for estimation and assimilation into models. Ecol Modell 101: 41–59

Imbrie J, Imbrie JZ (1980) Modelling the climatic response to orbital variations. Science 207: 943–953

Lenton TM (2000) Land and ocean carbon cycle feedback effects on global warming in a simple earth system model. Tellus Ser B Chem Phys Meteorol 52: 1159–1188

Lorenc AC (1986) Analysis methods for numerical weather prediction. Q J R Meteorological Soc 112: 1177–1194

Matear RJ (1995) Parameter optimisation and analysis of ecosystem models using simulated annealing: a case study at Station P. J Mar Res 53: 571–607

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machines. J Chem Phys 21: 1087–1091

Miller RN, Ghil M, Gauthiez F (1994) Advanced data assimilation in strongly nonlinear dynamical systems. J Atmos Sci 51(8): 1037–1056

Paillard D (1998) The timing of Pleistocene glaciations from a simple multiple-state model. Nature 391: 378–381

Paillard D (2001) Glacial cycles: towards a new paradigm. Rev Geophy 39(3): 325–346

Petit J, Jouzel J, Raynaud D, Barkov NI, Barnola J, Basile I, Bender M, Chappellaz J, Davis M, Delaygue G, Delmotte M, Kotlyakov VM, Legrand M, Legrand C, Lorius VYL, Pepin L, Ritz C, Saltzman E, Stievenard M (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature 399: 429–436

Pincus M (1970) A Monte Carlo method for the approximate solution of certain types of contrained optimisation problems. Operat Res 18: 1225–1228

Prentice ML, Mathews RK (1988) Cenozoic ice-volume history: development of a composite oxygen isotope record. Geology 16: 963–966

Saltzman B, Maasch KA (1988) Carbon cycle instability as a cause of the late Pleistocene ice age oscillations: modeling the asymmetric response. Global biogeoch Cycles 2: 177–185

Saltzman B, Maasch KA (1990) A first-order global model of late Cenozoic climate. Trans R Soc Edinburgh Earth Sci 81: 315–325

Saltzman B, Sutera A (1987) The mid-Quatenary climatic transition as the free response of a three-variable dynamical model. J Atmos Sci 44(1): 236–241

Saltzman B, Verbitsky M (1993) Multiple instabilities and modes of glacial rhythmicity in the Plio-Pleistocene: a general theory of late Cenozoic climatic change. Clim Dyn 9: 1–15

Saltzman B, Verbitsky M (1995) Predicting the Vostok CO_2 curve. Nature 377: 690–690

Shackleton N (2000) The 100,000 year ice-age cycle identified and found to lag temperature, carbon dioxide and orbital eccentricity. Science 289: 1897–1902