Parameter estimation for energy balance models with memory

Lionel Roques, Mickaël D. Chekroun, Michel Cristofol, Samuel Soubeyrand, Michael Ghil

Parameter estimation for energy balance models with memory

By Lionel Roques[1,*], Mickaël D. Chekroun[2], Michel Cristofol[3], Samuel Soubeyrand[1] and Michael Ghil[2,4]

[1] *UR 546 Biostatistique et Processus Spatiaux, INRA, F-84000 Avignon, France*

[2] *Department of Atmospheric & Oceanic Sciences and Institute of Geophysics & Planetary Physics, University of California, Los Angeles, CA 90095-1565, USA*

[3] *Université Aix-Marseille, LATP, F-13397 Marseille, France*

[4] *Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL), Ecole Normale Supérieure, Paris, France*

[*] *lionel.roques@paca.inra.fr*

We study parameter estimation for one-dimensional energy balance models with memory (EBMMs) given localized and noisy temperature measurements. Our results apply to a wide range of nonlinear, parabolic partial differential equations (PDEs) with integral memory terms. First, we show that a space-dependent parameter can be determined uniquely everywhere in the PDE's domain of definition $\mathcal{D}$, using only temperature information in a small subdomain $\mathcal{E} \subset \mathcal{D}$. This result is valid only when the data correspond to exact measurements of the temperature.

We propose a method for estimating a model parameter of the EBMM using more realistic, error-contaminated temperature data derived, for example, from ice cores or marine-sediment cores. Our approach is based on a so-called mechanistic-statistical model, which combines a deterministic EBMM with a statistical model of the observation process. Estimating a parameter in this setting is especially challenging because the observation process induces a strong loss of information. Aside from the noise contained in past temperature measurements, an additional error is induced by the age-dating method, whose accuracy tends to decrease with a sample's remoteness in time. Using a Bayesian approach, we show that obtaining an accurate parameter estimate is still possible in certain cases.

Keywords: age dating; Bayesian inference; energy balance model; inverse problem; mechanistic-statistical model; memory effects

1. Introduction

Energy balance models (EBMs) are among the simplest climate models. They were introduced almost simultaneously by Budyko [1] and Sellers [2]. Because of their simplicity, these models are easy to understand and facilitate both analytical and numerical studies of climate sensitivity. A key feature of these models is that they eliminate the climate's dependence on the wind field and thus have only one dependent variable: the Earth's near-surface air temperature T.

In the most elementary, zero-dimensional (0-D) EBMs, the temperature T is globally averaged and thus depends on time alone, and not on any space variables. In this simplifi-

cation, the Earth's global temperature still depends on such parameters as the planetary albedo, i.e. the globally averaged reflectivity of the land and sea surface. Mathematically, the temperature T satisfies an ordinary differential equation (ODE) with dT/dt determined by the balance between the incoming shortwave (solar) radiation and the outgoing longwave (terrestrial) radiation emitted in response to the former. EBMs thus model radiative equilibrium [3]; the word "equilibrium" here does not refer to the thermodynamic concept thereof, only to one or more steady-state solutions of the EBM.

In this paper, we deal with EBMs that incorporate an explicit spatial dependence of T on a meridional variable x. It is natural to consider reaction-diffusion equations that take into account pole-to-equator temperature gradients, while respecting overall radiative balance. In the one-dimensional (1-D) models we consider here, the temperature T is assumed to depend on space only through a latitudinal variable $x \in (0, 1)$ that may either be sine(latitude) [4, 5] or a scaled latitude [6].

For such 1-D models of the Sellers type, Ghil [6] has shown, using numerical continuation methods, that the model had multiple steady states lying along an S-shaped bifurcation curve with respect to a normalized parameter measuring solar radiation. These numerical results were rigorously justified and generalized to the case of two-dimensional (2-D) EBMs — with possibly discontinuous nonlinearities — on compact Riemannian manifolds without boundary; see [7] and references therein, as well as [8, 9] for other results on 2-D EBMs.

Thus, the extension of the ODE picture to the partial differential equation (PDE) setting supported the idea of a "deterministic skeleton" — namely the S-shaped bifurcation diagram — underlying the cold and warm phases of the glaciation cycles observed in isotopic and microfaunal proxy records of Quaternary climate [3]. When combined with the concept of stochastic resonance [10], the coexistence of multiple equilibria helped provide one of several qualitative explanations of these cycles; see [11] for a mathematical treatment in the framework of stochastic differential equations.

The behavior of EBMs depends, therefore, on the precise form of the ODE or PDE model's right-hand side (rhs) and on the values of their coefficients and parameters. The EBMs we treat are semi-linear, i.e., the nonlinearities only involve the temperature itself, and not its derivatives. In this setting, we refer to a *parameter* as a constant that intervenes in an ODE's or PDE's rhs, and to a *coefficient* as a function of space in the rhs. Furthermore, with suitable tuning of their parameters, EBMs that resolve the Earth's land-sea geography and are forced by the seasonal insolation cycle have been shown to mimic, to a certain extent, the observed zonal temperatures for the observed present climate [12]. Once EBMs are fitted to the observations [2, 6] or to the simulated climate of general circulation models (GCMs) [12], they can be used to estimate the temporal response patterns to various forcing scenarios; such a methodology is of particular interest in the detection and attribution of climate change [13].

Unfortunately, in practice, the model coefficients and parameters cannot be measured directly, because they generally result from the intertwined effects of several physical processes that are not adequately resolved by an EBM, and substantial uncertainty still prevails in estimating these parameters when fitting EBMs to either observational or simulation data [13]. The investigation of robust and efficient methods of parameter estimation for EBMs is, therefore, of considerable practical interest. Still, reaction-diffusion EBMs — in the absence of forcing or delays — are too simplistic with respect to the natural variability of climate [3, 14, 15]. For instance, such EBMs do not support the

spatio-temporal chaos characteristics of climate dynamics; see [3, 14, 16] and references therein.

Bhattacharya *et al.* [17] proposed an interesting generalization of reaction-diffusion EBMs. In this generalization, they introduced a delay mechanism in order to take into account the long response times of the ice sheets to temperature changes. In particular, they assumed that the albedo function depends not only on the present temperature but also on past temperatures through a memory function that modifies the shortwave radiation terms. The delay effects, combined with the model's nonlinearity, induce chaotic behavior with low-frequency variability [18]: the S-shaped bifurcation diagram of steady states — obtained in the "no-memory" case — is preserved, although the stability of the individual branches may be modified; see [19] for a rigorous treatment.

More general EBMs with memory (EBMMs) were subsequently considered in the literature. In these EBMMs, several physical mechanisms active in the model, and not just the albedo, are typically assumed to depend on a weighted linear combination of past temperatures, given by $H(t,x,T) := \int_{-\tau}^{0} \beta(s,x)\, T(t+s,x)\, ds$, for all $t > 0$, and $x \in (0,1)$, where $\beta(s,x)$ is the *memory kernel*; see §2 below for further details and [20] or Diaz *et al.* [21] in this volume for a recent survey.

In summary, EBMMs are still simple climate models but with richer dynamics than the one exhibited by classical EBMs, while preserving several key features and advantages of the latter. EBMMs may thus play an interesting role in the detection and attribution problem, for instance, when fitted to modern instrumental observations and to the simulation results of IPCC-class GCMs, or in the modeling of climates of the past when fitted to proxy records from the geological era of interest. It follows that the successful application of EBMMs to these two problem areas depends on a proper analysis and solution of the associated inverse problem. Such an analysis appears to be missing in the literature and we propose to solve at least some EBMM-related inverse problems of interest in the present article.

Recall that the *inverse* approach consists of determining unknown coefficients by using measurements of the temperature field $T(t,x)$. If $T(t,x)$ is known at all times $t \geq 0$ and at all points $x \in (0,1)$, it is plausible that all model coefficients can be determined, subject to reasonable mathematical assumptions about the PDE and its solutions. However, climate data are generally local and noisy, that is, $T(t,x)$ can only be measured at some points or in some (often small) subregions of the domain $(0,1)$, and the measurements are subject to error. In this paper, we study whether spatially variable coefficients of EBMM models can be either uniquely determined in the whole space domain $(0,1)$, by using only exact but local information about the temperature field, or be estimated with sufficient accuracy by using local and noisy measurements of the temperature.

The paper is organized as follows. In §2, we summarize the equations that govern EBMs, from the basic ODE model to general EBMMs. In §3, we obtain a uniqueness result for space-dependent coefficients of general 1-D EBMMs. These coefficients may be related to the albedo function in the solar-radiation term or the "greyness" function in the terrestrial-radiation term. Our result shows that a space-dependent coefficient can be determined uniquely in theory on the whole domain $\mathcal{D} = (0,1)$ by using local information about the temperature T that is available only on a small subdomain $\mathcal{E}$ of $(0,1)$ and during a short time period.

In §4, we propose a method for estimating space-dependent coefficients of EBMMs, based upon realistic data. The type of data we are interested in includes ice cores, continental loess or pollen records or deep-sea sediments [3, 14]. In the present, theoretical

setting, these all correspond to noisy measurements of past temperature at some localized geographical sites. Our approach is based on a *mechanistic-statistical model* [22, 23, 24], which combines a deterministic EBMM with a statistical model for the observation process. Estimating the EBMM's parameters is especially challenging because the observation process induces a strong loss of information: In addition to the noise in the past temperature measurements, an error is induced by the age-dating method, the accuracy of which tends to decrease as samples are extracted from earlier epochs in the record [25, 26]. Still, using a Bayesian approach, we show that obtaining an accurate estimate of the parameters is still possible in certain cases. The paper concludes with a discussion section, and several appendices provide technical details.

2. Energy balance models (EBMs)

Classical EBMs

One of the simplest 0-D EBMs,

$$c\,\frac{dT(t)}{dt} = \mu\,Q[1 - a(T(t))] - g(T(t)), \tag{2.1}$$

has been analyzed in [3, 27]. Here, c is a global thermal inertia coefficient and Q is the "solar constant," i.e. the mean insolation per unit area at the top of the atmosphere, and $\mu\,Q$ corresponds to present-day radiation conditions for $\mu = 1$. Other values of the parameter μ thus correspond to changes in these conditions and lead to changes in the number and stability of the model's equilibria [28, 29].

The functions $a(T(t))$ and $g(T(t))$ in the incoming and outgoing radiation terms of Eq. (2.1), respectively, correspond to the albedo function and to the greyness function; the latter models the difference between black-body radiation σT^4 — with σ the Stefan-Boltzmann constant — and the actually observed terrestrial one. In practice, it is hard to formulate realistic functions a and g. In particular, the albedo function may depend on the vegetation cover of the Earth's land surface, on the presence of clouds [17], and on the wave distribution on the ocean surface [30].

Several functional forms of a and g are available in the literature. In Sellers-type models [2], the albedo is a piecewise-linear ramp function of the form:

$$\begin{cases} a(T) = a_0 & \text{for } T \leq T_1, \\ a(T) = a_0 + (a_1 - a_0)\,(T - T_1)/(T_2 - T_1) & \text{for } T \in (T_1, T_2), \\ a(T) = a_1 & \text{for } T \geq T_2, \end{cases} \tag{2.2}$$

where $a_0 > a_1 > 0$ and $T_1 < T_2$ are all constants. Other albedo functions have been proposed for the study of more local problems, in restricted temperature ranges; for instance, Fraedrich [28] proposed linear or quadratic functions of the form:

$$a(T) = a_1 - a_2\,T \text{ or } a(T) = b_1 - b_2\,T^2,$$

where a_1, a_2, b_1, $b_2 > 0$.

The function g is positive and can take several forms as well [1, 2, 6, 17, 30]. To be consistent with an Earth assumed to be in radiative equilibrium, any choice of a and g has to allow for one or more steady-state solutions of Eq. (2.1) [3].

The 1-D reaction-diffusion model presented in [6, 30] has the form:

$$c(x)\,\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k(x)\,\frac{\partial T}{\partial x}\right) + \mu\,Q(x)[1 - a(x,T)] - g(x,T),$$

for $t > 0$ and $x \in (0,1)$. On the rhs of this PDE one distinguishes, in all generality, between the *diffusion term* that contains the two spatial derivatives, and the *reaction terms* that contain no derivatives; the latter express here the balance between incoming and outgoing radiation.

The space variable x corresponds to the scaled latitude $x = 2\phi/\pi$, where ϕ is the latitude; the model is symmetric about the Equator and so $x = 0$ and $x = 1$ correspond to the equator and to the North Pole, respectively. The function k stands for a space-dependent diffusion coefficient and $c(x) > 0$ is the zonally averaged heat capacity of the Earth; in practice, this "thermal inertia" is mostly due to the oceans. In this 1-D model, all of the coefficients can depend on the space variable x. The function $Q(x)$ now corresponds to the meridional distribution of incident solar radiation averaged over a calendar year [6, 17, 30].

EBMs with memory terms (EBMMs)

In models with memory terms, the temperature values at a given time t depend not only on the temperatures at time t but also on a weighted linear combination of past temperatures $T(t-s)$ over some range of $s > 0$. To be more precise, one considers a *history function* $H(t,x,T)$ given by

$$H(t,x,T) = \int_{-\tau}^{0} \beta(s,x)\,T(t+s,x)\,\mathrm{d}s, \text{ for all } t > 0 \text{ and } x \in (0,1), \tag{2.3}$$

over a past interval of length $\tau > 0$, with a memory kernel $\beta(s,x)$ that belongs to the functional space $L^1([-\tau,0];C([0,1]))$ and verifies:

$$\begin{cases} \int_{-\tau}^{0} \beta(s,\cdot)\,\mathrm{d}s \equiv 1, \\ \beta(s,\cdot) \geq 0 \text{ for all } s \in [-\tau,0]. \end{cases} \tag{2.4}$$

At this level of generality, 1-D EBMMs can be described by the equation:

$$c(x,H(t,x,T))\,\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k(x)\,\frac{\partial T}{\partial x}\right) + f(t,x,T,H(t,x,T)), \tag{2.5}$$

for $t > 0$ and $x \in (0,1)$.

Initial and boundary conditions

Because of the definition of H, the initial condition in such models has to be of the form:

$$T(s,x) = T_0(s,x), \text{ for } s \in [-\tau,0] \text{ and } x \in [0,1], \tag{2.6}$$

for some Lipschtiz-continous function T_0 defined on $[-\tau,0] \times [0,1]$.

Following Bhattacharya *et al.* [17], the boundary conditions are of Neumann's type:

$$\frac{\partial T}{\partial x}(t,0) = \frac{\partial T}{\partial x}(t,1) = 0 \text{ for } t \geq 0. \tag{2.7}$$

Note that the assumption (2.7) also applies at $t = 0$. This compatibility condition is required to obtain the time continuity of the solution T of (2.5) up to $t = 0$.

When c does not depend on H, the use of the theory of semigroups typically allows a recasting of Eq. (2.5), together with the initial and boundary conditions (2.6) and (2.7), as an abstract *semilinear functional evolution equation* [31]. The solutions of such an equation are typically understood, in a weak or strong sense, within some appropriate Banach space whose properties depend on the assumptions that hold for f, c and k. These assumptions are discussed in Appendix A.

For instance, when $k \equiv 1$, the problem (2.5)—(2.7) can be described by

$$\frac{dT}{dt} = AT + F(T, T_t), \quad t \geq 0;$$

the operator A is the realization on a Banach space X of the Neumann-Laplacian [32], F is an X-valued (usually) nonlinear mapping defined on $X \times C([-\tau, 0]; X)$ and $T_t \in C([-\tau, 0]; X)$ is defined by $T_t(s) = T(t + s)$ for $s \in [-\tau, 0]$, as in [31].

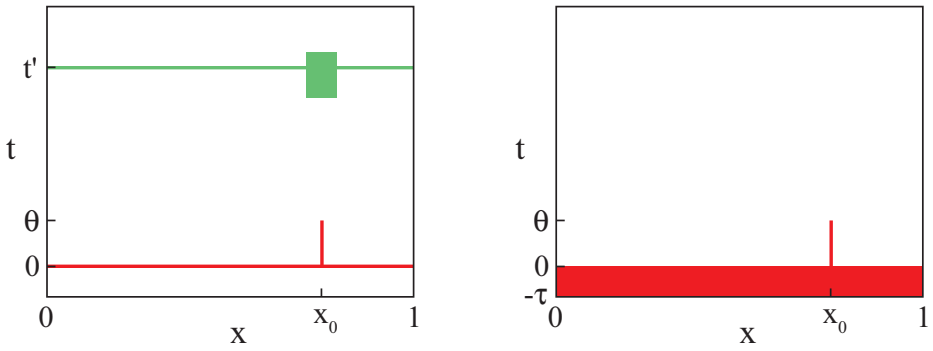
When c is a genuine function of H, the problem (2.5)—(2.7) becomes a *quasilinear evolution equation with memory*, and questions of existence, uniqueness, and continuous dependence typically require greater care; see [8, 20] for EBMMs on the Euclidean unit sphere in $\mathbb{R}^3$, and [33] for other quasilinear parabolic functional evolution equations. In Appendix B, we prove — using appropriately chosen comparison principles [34] — that the local existence and uniqueness of classical solutions in time holds for the problem (2.5)–(2.7), under fairly general conditions on the functions f, k, and c.

3. Unique coefficient determination for exact data

For reaction-diffusion equations without a memory term, uniqueness results in inverse problems are generally obtained using the method of Carleman estimates [35]. This method requires, among other measurements, knowledge of the solution $T(t', x)$ of the equation at some time $t' > 0$ and for all x in the domain $(0, 1)$ [37, 38, 39]; see figure 1 (a). For polynomial reaction terms, more recent approaches [40, 41] lead to uniqueness results for one or several coefficients, under the assumption that T and its first spatial derivative are known at a *single point* x_0 in $(0, 1)$, for all t in a small interval $(0, \theta)$, and that the initial data $T(0, x)$ is known over the entire interval $(0, 1)$. Thus, one of the main differences between these two approaches is that, in the method presented in [40, 41], the measurement of the solution at time t' is replaced by the knowledge of the initial condition; compare the green and red regions in figure 1 (a). Contrarily to the measurement at time t', the initial condition does not contain any information on the coefficient to be determined. Thus, this approach shows that coefficients of reaction-diffusion equations without a memory term can be uniquely determined using only local information about the solution T. In this study, we will show that comparable results can be obtained for EBMMs with memory terms and nonlinear reaction terms; see figure 1 (b).

We assume here that the function f in Eq. (2.5) has the form:

$$f = f_\alpha(t, x, T, H(t, x, T)) = f_1(t, x, T, H(t, x, T)) + \alpha(x)\, f_2(t, T, H(t, x, T)), \tag{3.1}$$



(a) (b)

Figure 1. Space-time measurement regions which can lead to unique coefficient determination. (a) For reaction-diffusion equations without a memory term: In green, the region that is required when the method of Carleman estimates is used. In red, the region used in the method developed in [40, 41]. (b) For EBMMs: the region that is used in this study.

and that the function $x \mapsto \alpha(x)$ is not known. We refer to Appendix A for the exact assumptions on f_1 and f_2 used throughout this article. Our goal is to study which conditions allow one to determine the function α.

More precisely, consider a solution $\tilde{T}$ of Eq. (2.5), where $f = f_1 + \alpha f_2 \equiv f[\alpha]$ is replaced by $f[\tilde{\alpha}] = f_1 + \tilde{\alpha}\, f_2$. Assume that $\tilde{T}$ satisfies the same initial and boundary conditions — i.e., (2.6) and (2.7), respectively — as T. Our goal is to prove that for any subset of positive measure $\mathcal{E} \subset (0,1)$, and any $\theta > 0$,

$$\{\tilde{T}(t,x) = T(t,x) \text{ in } (0,\theta) \times \mathcal{E}\} \Rightarrow \{\tilde{\alpha}(x) \equiv \alpha(x) \text{ over } [0,1]\}. \tag{3.2}$$

In other words, the space-dependent coefficient $\alpha(x)$ is uniquely determined on $[0,1]$ by any measurement of T on a set $(0,\theta) \times \mathcal{E}$.

The results of the present section are obtained under an additional assumption on the kernel β of the history function H in Eqs. (2.3) and (2.4), namely that

$$\exists\, \delta > 0 \text{ s.t. } \beta(s,\cdot) \equiv 0 \text{ for all } s \in [-\delta, 0]. \tag{3.3}$$

Thus, very recent past temperatures are not taken into account in $H(t,x,T)$. This assumption is not very restrictive since, for small enough δ, the continuity of $t \mapsto T(t,x)$ implies that these temperatures are very close to $T(t,x)$, which is already taken into account as a variable of the reaction function f in (2.5).

Under assumption (3.3) and the regularity assumptions of Appendix A, we prove in Appendix B that the problem (2.5) with initial condition (2.6) and boundary condition (2.7) admits a unique classical solution, at least for small times $t < t^*$, and t^* does not depend on α. Global existence in time of solutions to Eqs. (2.5)–(2.7) clearly depends on the precise form of f; see [8] for global existence results in the case of EBMMs on the Euclidean unit sphere in $\mathbb{R}^3$. For the statement of the main result of this section and the numerical results of § 4, we only need the solution T to be defined for small times. Further existence results for parabolic equations with delay terms are given in refs. [31, 34].

Uniqueness result

We assume that the unknown coefficient $\alpha(x)$ of (3.1) lies in the function space:

$$\mathcal{M} := \{\psi \text{ is Lipschitz-continuous and piecewise analytic on } [0,1]\}. \tag{3.4}$$

Remarks

- A continuous function ψ is called piecewise analytic if there exist $n \geq 1$ and an increasing sequence $(\kappa_j)_{1\leq j\leq n}$ such that $\kappa_1 = 0$, $\kappa_n = 1$, and
  $\psi(x) = \sum_{j=1}^{n-1} \chi_{[\kappa_j,\kappa_{j+1})}(x)\varphi_j(x)$, for all $x \in (0,1)$;
  here φ_j are some analytic functions defined on the intervals $[\kappa_j, \kappa_{j+1}]$, and $\chi_{[\kappa_j,\kappa_{j+1})}$ are the characteristic functions of the intervals $[\kappa_j, \kappa_{j+1})$ for $j = 1, \ldots, n-1$.
- The assumption $\alpha \in \mathcal{M}$ is not very restrictive. For instance, the set of piecewise linear functions on $[0,1]$ is a subset of $\mathcal{M}$.

Theorem 3.1. *Let α and $\tilde{\alpha}$ satisfy assumption* (3.4). *Let T and $\tilde{T}$ be the solutions of Eqs.* (2.5)–(2.7) *on $[0,t^*) \times [0,1]$, with $f = f[\alpha]$ and $f = f[\tilde{\alpha}]$, respectively, cf. Eq.* (3.1). *Assume that there exist $x_0 \in [0,1]$ and $\theta \in (0,t^*)$ such that*

$$\begin{cases} T(t,x_0) &= \tilde{T}(t,x_0), \\ \dfrac{\partial T}{\partial x}(t,x_0) &= \dfrac{\partial \tilde{T}}{\partial x}(t,x_0), \end{cases} \quad \text{for all } t \in (0,\theta). \tag{3.5}$$

Then $\alpha \equiv \tilde{\alpha}$ on $[0,1]$ and consequently $T \equiv \tilde{T}$ on $[0,t^*) \times [0,1]$.

The proof of Theorem 3.1 is given in Appendix C. This result means that the space-dependent coefficient $\alpha(x)$ is uniquely determined on $[0,1]$ by any measurement of T and $\partial T/\partial x$ at a single point x_0 during the time period $(0,\theta)$. In particular, the assertion (3.2) is a consequence of this result.

4. Parameter estimation for noisy measurement histories

In the previous section, we showed that space-dependent coefficients of general EBMMs can be uniquely determined by local measurements of the temperature T. In this section, we show that it is possible to estimate such coefficients based upon realistic data, which are typically noisy and incomplete, in both space and time.

The available data, derived for instance from ice cores, continental records or deep-sea sediments, generally correspond to noisy measurements of the past temperature at some localized geographical positions. Such data contain two sources of uncertainty: (i) in the value of the measured temperature; and (ii) in the accuracy of the dating, which tends to decrease as samples are derived from earlier time points [25]. We propose a method for parameter estimation using such noisy measurements.

Our approach uses a mechanistic-statistical model. The idea is to build a model that links a mechanistic vision of the studied phenomenon — in our case through the EBMM — with data collected from the observation of this phenomenon [22, 23, 24]. Estimating the parameters of the mechanistic part of the model is challenging when the observation process induces a strong loss of information, because of severely subsampled observations and substantial noise, as is the case here.

An EBMM with unknown coefficient

We assume that the temperatures are given by an EBMM corresponding to a particular case of Eq. (2.5), with:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \alpha(x)\,(1 - a(T)) - q_0 - q_1\,T - \left(\frac{1}{\tau}\int_{-\tau}^{0} T(t+s)\,\mathrm{d}s\right)^3, \tag{4.1}$$

for $t > 0$ and $x \in (0,1)$. The function a is a ramp function as in (2.2) and $\alpha(x)$ is the coefficient to be estimated. We carried out our computations for two values of the delay parameter, $\tau = 0.2$ ky and $\tau = 0.7$ ky, during the time interval $0 \le t \le 5$ ky, where 1 ky = 1000 years. We refer to the subsection titled "Numerical results" below for the precise numerical setting of Eq. (4.1).

A statistical model of the observation process

We assume that data are available at several locations S_k, $k = 1, 2, 3$. At each location k, we denote by $t_1, \ldots, t_I$, where $I = 50$ in our simulations, the sequence of decreasing epochs $\theta = 5 \ge t_1 > \ldots > t_I \ge -\tau$, at which the temperature $T(t_i, S_k)$ is measured, based on laboratory sampling of a given ice core, say, extracted at location S_k. The times t_i do depend, in general, on the location S_k, but the label k has been dropped for the sake of concision and clarity. Let $Y_k(t_i)$ denote the measure of the temperature $T(t_i, S_k)$.

The uncertainty due to the age dating approximation implies that $Y_k(t_i)$ is actually a measure of the temperature $T(s(t_i), S_k)$, where $s(t)$ is a function deforming the time scale that can vary with k. This function is the result of errors in the chronostratigraphy, i. e. in the *age-depth plot*, of a given core [36]. Furthermore, the uncertainty in measuring the temperature value implies that $T(s(t_i), S_k)$ contains noise as well.

Our model for the observation process has to take into account these two sources of uncertainty. First, given $s(t_i)$, $i = 1, \ldots, I$, the observed variables $Y_k(t_i)$ are assumed to be conditionally drawn from independent Gaussian distributions $\mathcal{N}$:

$$Y_k(t_i) \mid s(t_i) \sim \text{indep. } \mathcal{N}\left\{T(s(t_i), S_k), \sigma^2\right\}, \tag{4.2}$$

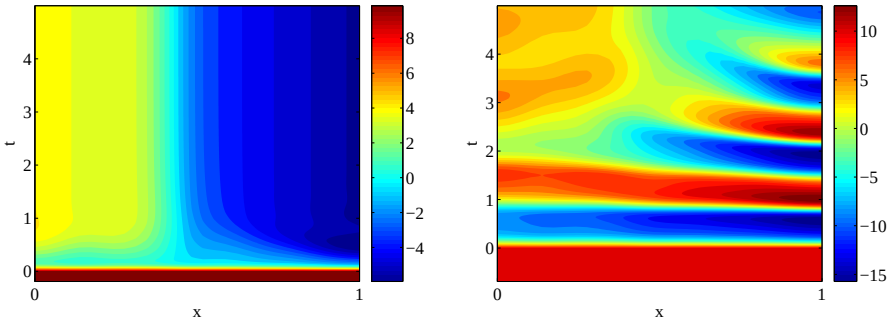
where σ^2 is the variance of the noise in the temperature measurements.

Second, we construct a model for $s(t_i)$:

$$s(t_i) = \theta - \sum_{j=1}^{i} \eta_j \text{ with } \eta_j \sim \text{indep. } \Gamma\left(\frac{t_{j-1} - t_j}{\kappa^2}, \kappa^2\right), \tag{4.3}$$

where Γ denotes the gamma distribution, κ^2 is a positive parameter that controls the shape of the distribution and $t_0 = \theta$. The expectation of $s(t)$ in Eq. (4.3) is t, and its variance increases as t moves further into the past. Another important feature of the model (4.3) is that it is order-preserving: if $t_i > t_j$, then $s(t_i) > s(t_j)$. It is indeed reasonable to assume that there is no uncertainty on the order of the estimated times $t_1, \ldots, t_I$.

Our model for the observation process is the combination of Eqs. (4.2) and (4.3). Using this model, we are able to compute the joint distribution of the observations $\mathcal{Y} := \{Y_k(t_i),\ i = 1, \ldots, I,\ k = 1, 2, 3\}$, conditioned on the "real" temperatures $\mathcal{T} := \{T(s(t_i), S_k),\ i = 1, \ldots, I,\ k = 1, 2, 3\}$. This joint distribution, denoted by $F(\mathcal{Y})$, is computed explicitly in Appendix D. If the temperatures $\mathcal{T} = \{T(s(t_i), S_k)\}$ are governed by



(a) $\tau = 0.2$ ky (b) $\tau = 0.7$ ky

Figure 2. The solution $T(t, x)$ of the EBMM model (4.1).

the EBMM (4.1), then they depend deterministically on the unknown coefficient α and the unknown initial data T_0 of Eq. (2.6). Thus, the conditional distribution $F(\mathcal{Y})$ of the observation process is equal to the likelihood $\mathcal{L}_{\mathcal{Y}}(\alpha, T_0)$ of the model parameterized by α and T_0.

Bayesian inference

Let T_0 correspond to the temperatures during the time window $[-\tau, 0]$ with prior distribution π_1. Let α denote the coefficient vector of our model with prior distribution π_2; we assume that the other coefficients and parameters are known. The posterior distribution $p(T_0, \alpha \mid \mathcal{Y})$ of T_0 and α is proportional to:

$$p(T_0, \alpha \mid \mathcal{Y}) \propto \mathcal{L}_{\mathcal{Y}}(\alpha, T_0)\pi_1(T_0)\pi_2(\alpha). \tag{4.4}$$

In the absence of further information, we assumed independent uniform prior distributions in a sufficiently large interval for the value of the parameter $\alpha(x)$ at each latitude x :

$$\pi_2(\alpha(x)) \sim U(0, 1000).$$

For the sake of simplicity, we assumed that the prior distribution of T_0 was a Dirac delta function:

$$\pi(T_0) \sim \delta_K, \tag{4.5}$$

where K is a constant obtained by averaging the observations $Y_k(t_i)$ over all negative times $t_i < 0$ and $k = 1, 2, 3$. This means that T_0 is constant in space and time and takes the value K. We draw a sample from the posterior distribution by a Markov chain Monte Carlo (MCMC) algorithm [42, 43]; see Appendix E for details.

Numerical results

For a given coefficient $\alpha(x)$, we computed the solution $T(t, x)$ of the EBMM (4.1) for $\tau = 0.2$ ky and $\tau = 0.7$ ky, emanating from the same constant initial history $T_0 = 10°\text{C}$ given for $[-\tau, 0]$ and $x \in [0, 1]$. The parameter values used in these computations for solving Eq. (4.1) are: $q_0 = 204$ and $q_1 = 1.73$ in the longwave-radiation function, and $a_0 = 0.62$, $a_1 = 0.25$, $T_1 = -10$, and $T_2 = 0$ in the albedo function defined by Eq. (2.2).

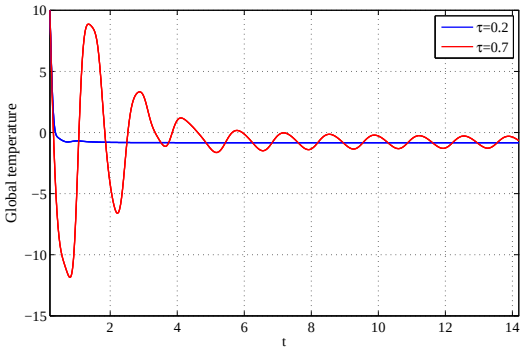


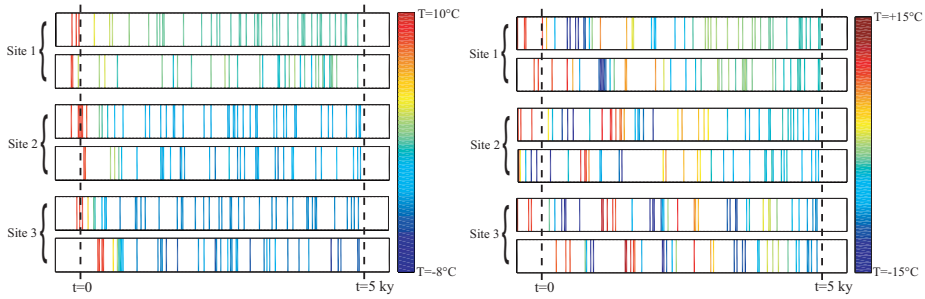
Figure 3. Time series of the global average temperature. The case $\tau = 0.7$ ky exhibits a transient with large variations that suggest the presence of nonnormal modes [47]; see Supplementary Information in [48] for a similar signature of such modes in a simpler model.

When $\tau = 0.2$ ky, the solution of (4.1) exhibits small temporal variations that are quickly damped and a stable steady state is quickly reached (figure 2 (a)). For $\tau = 0.7$, a stable periodic orbit $T_p(t, x) = T_p(t+p, x)$ of period $p \simeq 1.33$ ky is reached asymptotically, but the transient is considerably longer and larger amplitudes persist for quite a while (figures 2 (b) and 3). Such large-amplitude variations that precede the setting in of the periodic solution T_p point to the presence of nonnormal modes associated with the linearization of Eq. (4.1) about the stable periodic solution T_p. Such nonnormal modes are associated with polynomial rather than exponential behavior in time and they are responsible for large transient energy growth in the linearized problem. This growth can be further amplified by nonlinear effects or by noise; see [44, 45, 46, 47] for many manifestations of such modes and the Supporting Information of [48] for an analogous situation in a simpler model.

At each one of the locations $S_1 = 0.5$, $S_2 = 0.7$ and $S_3 = 0.9$, we randomly drew 50 epochs t_i in the interval $(-\tau, 5)$ and recorded the temperatures $T(t_i, S_k)$ at these epochs and locations. Using our observation model (4.2, 4.3), we constructed noisy observations $Y_k(t_i)$ of these temperatures. The exact temperatures at the exact times are presented together with the measured values in figure 4. Observed temperatures $Y_k(t_i)$ for negative times were used to determine the prior distribution of T_0, cf. Eq. (4.5). Observed temperatures $Y_k(t_i)$ for positive times were used to make the posterior inferences.

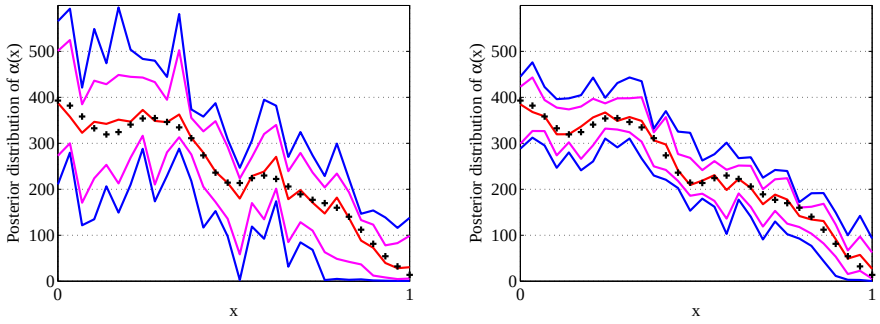
Figure 5 shows the marginal posterior quantiles of $\alpha(x)$, when $\tau = 0.2$ ky and $\tau = 0.7$ ky. In both panels, the median of the posterior distribution is quite close to the true values of the coefficient $\alpha(x)$. Moreover, the true values do lie between the first and last deciles of the distribution, for all values of x. Remember that observations were only made at three sites S_k, and that all three sites lie in the right half of the model domain, $S_k \in [0.5, 1]$. This restriction only seems to affect somewhat the estimation error of $\alpha(x)$ in the left half of the model domain, $x \in [0, 0.5]$, for $\tau = 0.2$ ky.

Overall, the main difference between figures 5(a) and (b) is that the marginal distribution is more variable in the case $\tau = 0.2$ ky, for all x, meaning that the insolation coefficient $\alpha(x)$ is harder to estimate in this case. Such a result is somewhat surprising because the larger and faster variations in the case $\tau = 0.7$ lead to larger measurement errors; this can be easily seen in figure 4. A possible explanation for this result is that the solution of (4.1) is much more sensitive to variations in the parameters when $\tau = 0.7$.



(a) $\tau = 0.2$ ky (b) $\tau = 0.7$ ky

Figure 4. Actual temperatures versus measured temperatures. At each site S_k, $k = 1, 2, 3$, the upper row corresponds to the actual temperatures at the actual times, while the lower row corresponds to the measured temperatures at estimated times.



(a) $\tau = 0.2$ ky (b) $\tau = 0.7$ ky

Figure 5. Marginal posterior quantiles of the parameter $\alpha(x)$. The red curve is the posterior median of $\alpha(x)$, the magenta curves the first and last deciles of its distribution, and the blue curves the first and last percentiles. The true values of $\alpha(x)$ are given by the symbol $+$.

In order to measure the sensitivity to parameter variations, we computed the average L^2-response $\overline{R}_\varepsilon$ of the model (4.1), over the time window $0 < t < 5$, to an additive perturbation z of the parameter α. The perturbation z is drawn from a Gaussian random field with standard deviation ε; see Appendix F for the precise definition of $\overline{R}_\varepsilon$.

The results are presented in figure 6 for $\tau = 0.2$ ky (blue curve) and for $\tau = 0.7$ ky (red curve). In both cases, the L^2-response increases linearly with ε, but the slope is more than twice as large in the latter case. As noticed above, the nonnormal modes are responsible for the large fluctuations observed over the window $0 < t < 5$, cf. figure 3. It is known that, in general, these modes are most sensitive to perturbations of the linearized problem; see [45, 47]. The simple L^2-response analysis we performed here confirms that in the presence of such modes, larger changes occur (on average) in the solutions of the nonlinear problem (4.1), when perturbing the coefficients of this EBMM.

5. Discussion

We have investigated the extent to which the coefficients of energy balance models with memory (EBMMs) can be determined or estimated based on localized measurements of

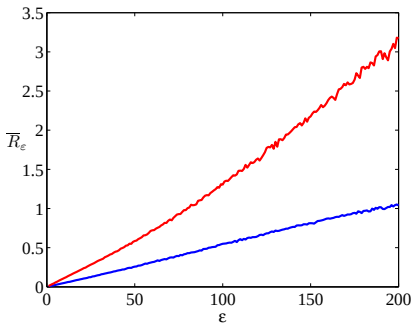


Figure 6. Average L^2-response $\overline{R}_\varepsilon$ of the model (4.1), over the time interval $t \in (0, 5)$, to random perturbations z sampled from a Gaussian random field $\mathcal{Z}$ with standard deviation ε. The blue curve corresponds to the case $\tau = 0.2$ ky and the red curve to $\tau = 0.7$ ky.

past temperatures. In the purely deterministic case of exact measurements, we obtained rigorous theoretical results on the determination of a coefficient in the EBMM's reaction term, while in the more realistic case of stochastically perturbed measurements, we obtained numerical results for the estimation of this coefficient. Both types of results can be extended to a general class of nonlinear parabolic PDEs with integral memory terms.

In particular, we showed in §3 that a space-dependent coefficient $\alpha(x)$ of a general EBMM can be uniquely determined using only local information about the temperature over an arbitrarily small subset $\mathcal{E}$ in space and during a short time window $0 < t < \theta$; see figure 1 (b). More precisely, given measurements of the temperature and of its first spatial derivative at a single site x_0 and for $t \in (0, \theta)$, and provided that the initial data — i.e., the temperature $T_0(t, x)$ for $-\tau \leq t \leq 0$ — be known, there is a unique coefficient $\alpha(x)$ of the model's reaction term that can lead to such measurements.

Note that, for this result to be true, the initial temperatures $T_0(t, x)$ need not be generated using the parameter $\alpha(x)$ and, therefore, do not contain any information on $\alpha(x)$. Thus, surprisingly, α is uniquely determined everywhere in the domain by information contained in purely local measurements. This type of uniqueness result was available for classical reaction-diffusion equations without memory terms, but is new, to the best of our knowledge, for parabolic equations with memory terms. Related results for hyperbolic equations with memory terms appear in [49].

In practice, temperature measurements are strongly contaminated by noise, as well as being heavily under-sampled in time and space. This situation prevails in particular when studying temperatures that precede modern-era, instrumental measurements, and one needs to use inverse methods for deducing them from proxy records preserved in ice cores, deep-sea sediments or continental records [3, 14, 50].

In §4, we treated the more realistic and difficult problem of estimating the unknown space-dependent coefficient of an EBMM in the presence of incomplete and error-contaminated measurements. Model solutions for the correct coefficient were shown in figures 2 and 3. The loss of information in the sampling and measurement process was modeled statistically by taking into account the two main problems that arise in the interpretation of such past-temperature data: (1) age-dating errors [25, 36]; and (2) converting the isotopic, microfaunal, dendrochronological or other measurements into past temperature changes [25, 50]. The two kinds of errors were visualized in figure 4. We obtained a mechanistic-statistical model for coefficient-inference by coupling this statistical

model with an EBMM, and subsequently showed how to estimate the EBMM's unknown coefficient by a Bayesian approach.

Our numerical results for the coefficient estimation in figure 5 are promising and we expect to apply them next to real observational data as well as to GCM model simulations. The results might be valuable for the detection-and-attribution problem of modern climate change, cf. [13, 51]. It is especially notable that our method still allows for fairly accurate parameter estimation when the solution of the EBMM exhibits significant spatial and temporal fluctuations, as noted during the transient regime for the case $\tau = 0.7$ of figure 3. Our numerical results even suggest that, although the age-dating errors lead to measurement errors that are larger overall in the presence of such temporal fluctuations, parameter estimation is more accurate in those cases in which the model solution is more sensitive to variations in the parameters, cf. figure 6. This sensitivity to parameter variations may be attributed to the presence of nonnormal modes [46, 47].

This empirical fact has been noted in other parameter estimation problems. Considerable experience with sequential estimation has accumulated over the last three decades in meteorology, oceanography and climate dynamics [52, 53, 54]. In this area of data assimilation, for instance, it is known that a model's strong dynamic variability may help reduce errors during the estimation procedure, whereas in a regime that corresponds to exponential convergence towards a steady state serious limitations in error reduction may arise [55]. We have shown in this study that the model's response to random parameter variations, as illustrated in [48], may help quantify these heuristic considerations and provide further insight into the estimation algorithm's performance.

More work is needed, however, in order to clarify the extent to which spatial and temporal fluctuations of the solution, along with the presence of nonnormal modes, might lead to more accurate parameter estimation when the observational data are subject to age-dating errors. One could thus test the mechanistic-statistical method developed here on weakly mixing chaotic systems, which seem to exhibit a linear response with small slope to parameter variations [48], and compare it with the results for strongly mixing chaotic systems, such as the Lorenz [56] model.

Finally, the mechanistic-statistical approach developed in this paper could be adapted to other types of observations, as well as to other types of PDE models, such as stochastic PDEs of PDEs with memory terms associated with the diffusion operator; the latter arise in viscoelastic models or in heat conduction in a material with memory [57]).

The authors are grateful to D. Kondrashov for helpful discussions. This study was supported by the U.S. National Science Foundation grant DMS-1049253 and the French "Agence Nationale de la Recherche" within the projects PREFERED and URTICLIM.

Appendix A. Assumptions on c, k, f and α

The heat capacity c and the diffusion coefficient k satisfy:

$$(x, u) \mapsto c(x, u) \in C^1([0, 1] \times \mathbb{R}) \text{ and } c > 0 \text{ on } [0, 1], \tag{A 1}$$

and

$$x \mapsto k(x) \in C^2([0, 1]) \text{ and } k > 0 \text{ on } [0, 1]. \tag{A 2}$$

The functions f_1 and f_2 in (3.1) verify:

$$f_1(t, x, u, v) \in C^1(\mathbb{R}_+ \times [0, 1] \times \mathbb{R} \times \mathbb{R}) \tag{A 3}$$

and

$$f_2(t,u,v) \in C^1(\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}). \tag{A 4}$$

Furthermore, we assume that

$$f_2(0, T_0(0,x), H(0,x,T_0(0,x))) \neq 0 \text{ for all } x \in [0,1]. \tag{A 5}$$

The parameter α is assumed to be Lipschitz-continuous on $[0,1]$ and verifies

$$|\alpha(x)| < M, \ \forall\, x \in [0,1],$$

for some positive constant M.

Appendix B. Existence and uniqueness of the solution of the general EBMM (2.5) for small times

The assumptions of Appendix A on c, f_1, f_2, α imply that we can define two Lipschitz-continuous functions f^- and f^+, which are independent of the (unknown) function α and such that,

$$f^-(T) \leq \min_{x\in[0,1], t\in[0,\delta]} \frac{f_1(t,x,T,H(t,x,T_0)) - M|f_2(t,T,H(t,x,T_0))|}{c(x,H(t,x,T_0))}, \tag{B 1}$$

and

$$f^+(T) \geq \max_{x\in[0,1], t\in[0,\delta]} \frac{f_1(t,x,T,H(t,x,T_0)) + M|f_2(t,T,H(t,x,T_0))|}{c(x,H(t,x,T_0))}, \tag{B 2}$$

for all $T \in \mathbb{R}$ (δ is defined by eq. (3.3)). From the Cauchy-Lipschitz theorem, there exists $t^* \in (0,\delta)$ (which does not depend on α) and two functions T^- and T^+ in $C^1[0,t^*]$ which satisfy:

$$\begin{cases} (T^-)' = f^-(T^-) \text{ for } t \in (0,t^*) \text{ and } T^-(0) = \min\limits_{x\in[0,1]} T_0(0,x), \\ (T^+)' = f^+(T^+) \text{ for } t \in (0,t^*) \text{ and } T^+(0) = \max\limits_{x\in[0,1]} T_0(0,x). \end{cases} \tag{B 3}$$

We have the following Lemma:

Lemma 5.1. *The problem* (2.5)-(2.7) *with f defined by* (3.1) *admits a unique classical solution $T \in C_1^2([0,t^*) \times [0,1])$ (i.e. the derivatives up to order two in x and order one in t are continuous) such that $T^-(t) \leq T(t,x) \leq T^+(t)$ for all $(t,x) \in [0,t^*) \times [0,1]$.*

Proof of Lemma 5.1: Let us set $\mathcal{T}^-(t,x) = T^-(t)$ and $\mathcal{T}^+(t,x) = T^+(t)$ for all $(t,x) \in [0,t^*) \times [0,1]$. Since $t^* < \delta$, the term $H(t,x,T)$ in (2.5) is equal to $H(t,x,T_0)$ for all $(t,x) \in [0,t^*) \times [0,1]$. Moreover, $\mathcal{T}^-$ verifies, for all $(t,x) \in [0,t^*) \times [0,1]$:

$$c(x,H(t,x,T_0))\,\frac{\partial \mathcal{T}^-}{\partial t} - \frac{\partial}{\partial x}\left(k(x)\,\frac{\partial \mathcal{T}^-}{\partial x}\right) - f_\alpha(t,x,\mathcal{T}^-,H(t,x,T_0)) \leq 0, \tag{B 4}$$

together with the boundary conditions $\frac{\partial \mathcal{T}^-}{\partial x}(t,0) = \frac{\partial \mathcal{T}^-}{\partial x}(t,1) = 0$ for $t \in [0,t^*)$ and the initial condition $\mathcal{T}^-(0,x) = T^-(0) \leq T_0(0,x)$ for all $x \in [0,1]$. Thus, $\mathcal{T}^-$ is a subsolution of the problem (2.5)-(2.7). Similarly, $\mathcal{T}^+$ is a supersolution of (2.5)-(2.7) and $\mathcal{T}^- \leq \mathcal{T}^+$ for all $(t,x) \in [0,t^*) \times [0,1]$. This implies (see e.g. [34]) that the problem (2.5)-(2.7) admits a unique classical solution $T \in C_1^2([0,t^*(T_0,f_1,f_2,M)) \times [0,1])$ such that $\mathcal{T}^-(t) \leq T(t,x) \leq \mathcal{T}^+(t)$ for all $(t,x) \in [0,t^*) \times [0,1]$. □

Appendix C. Proof of Theorem 3.1

Let us set $V = T - \tilde{T}$. Then the function V verifies:

$$\begin{array}{ll} c(x,H(t,x,T)) & \frac{\partial T}{\partial t} - c(x,H(t,x,\tilde{T}))\,\frac{\partial \tilde{T}}{\partial t} = \\ & \frac{\partial}{\partial x}\left(k(x)\,\frac{\partial V}{\partial x}\right) + f_\alpha(t,x,T,H(t,x,T)) - f_{\tilde{\alpha}}(t,x,\tilde{T},H(t,x,\tilde{T})), \end{array}$$

for all $t \in (0,t^*)$ and $x \in (0,1)$. Note that $V(0,x) = 0$ for all $x \in [0,1]$.

Since $t^* < \delta$, we observe that for all $x \in [0,1]$ and all $t \in (0,t^*)$

$$H(t,x,T) = H(t,x,\tilde{T}) = H(t,x,T_0).$$

As a consequence, we get that the function V verifies:

$$\begin{array}{ll} c(x,H(t,x,T_0))\,\frac{\partial V}{\partial t} = & \frac{\partial}{\partial x}\left(k(x)\,\frac{\partial V}{\partial x}\right) \\ & +f_\alpha(t,x,T,H(t,x,T_0)) - f_{\tilde{\alpha}}(t,x,\tilde{T},H(t,x,T_0)), \end{array} \tag{C 1}$$

for all $t \in (0,t^*)$ and $x \in (0,1)$. Using (3.1), (C 1) leads to

$$\begin{array}{ll} c(x,H(t,x,T_0))\,\frac{\partial V}{\partial t} = & \frac{\partial}{\partial x}\left(k(x)\,\frac{\partial V}{\partial x}\right) \\ & +f_1(t,x,T,H(t,x,T_0)) - f_1(t,x,\tilde{T},H(t,x,T_0)) \\ & +\alpha(x)\,f_2(t,T,H(t,x,T_0)) - \tilde{\alpha}(x)\,f_2(t,\tilde{T},H(t,x,T_0)), \end{array} \tag{C 2}$$

for all $t \in (0,t^*)$ and $x \in (0,1)$.

Let us set

$$\mu_1(t,x) = \begin{cases} \frac{f_1(t,x,T,H(t,x,T_0)) - f_1(t,x,\tilde{T},H(t,x,T_0))}{T-\tilde{T}} & \text{if } T(t,x) \neq \tilde{T}(t,x), \\ \frac{\partial f_1}{\partial T}(t,x,T,H(t,x,T_0)) & \text{otherwise,} \end{cases} \tag{C 3}$$

and

$$\mu_2(t,x) = \begin{cases} \frac{f_2(t,T,H(t,x,T_0)) - f_2(t,\tilde{T},H(t,x,T_0))}{T-\tilde{T}} & \text{if } T(t,x) \neq \tilde{T}(t,x), \\ \frac{\partial f_2}{\partial T}(t,T,H(t,x,T_0)) & \text{otherwise.} \end{cases} \tag{C 4}$$

Since f_1 and f_2 are C^1 with respect to T and continuous with respect to the other variables, the functions $\mu_1(t,x)$ and $\mu_2(t,x)$ are continuous and bounded on $[0,t^*)\times[0,1]$. Using the definitions (C 3) and (C 4) together with the equation (C 2), we obtain:

$$\begin{array}{ll} c(x,H(t,x,T_0))\,\frac{\partial V}{\partial t} = & \dfrac{\partial}{\partial x}\left(k(x)\,\dfrac{\partial V}{\partial x}\right) + \mu_1\,V + \tilde{\alpha}(x)\,\mu_2\,V \\ & +(\alpha-\tilde{\alpha})(x)\,f_2(t,T,H(t,x,T_0)), \end{array} \tag{C 5}$$

for all $t \in (0,t^*)$ and $x \in (0,1)$.

Let x_0 be defined as in the assumptions of Theorem 3.1, and let us set:

$$\mathcal{A}_+ = \Big\{ x \geq x_0 \text{ s.t. } (\alpha-\tilde{\alpha})(y) \equiv 0 \text{ for all } y \in [x_0,x] \Big\},$$

and

$$x_1 := \begin{cases} \sup(\mathcal{A}_+) & \text{if } \mathcal{A}_+ \text{ is not empty,} \\ x_0 & \text{if } \mathcal{A}_+ \text{ is empty.} \end{cases}$$

Assumption: We assume that $x_1 < 1$.

(i) *We prove that there exist* $\rho \in (0, t^*)$ *and* $x_2 \in (x_1, 1)$ *such that:*

$$(\alpha - \tilde{\alpha})(x)\, f_2(t, T, H(t, x, T_0)) \neq 0 \text{ for all } t \in [0, \rho] \text{ and } x \in (x_1, x_2].$$

From assumption (3.4), the functions α and $\tilde{\alpha}$ belong to $\mathcal{M}$. Thus $\alpha - \tilde{\alpha}$ also belongs to $\mathcal{M}$. Therefore, there exists $x_1' > x_1$ such that $\alpha - \tilde{\alpha}$ is analytic on $[x_1, x_1']$. This implies that either:

- (a) $\alpha - \tilde{\alpha} \equiv 0$ on $[x_1, x_1']$;
  or
- (b) there exists $x_2 \in (x_1, x_1')$ such that $(\alpha - \tilde{\alpha})(x) \neq 0$ for all $x \in (x_1, x_2]$.

Case (a) is in contradiction with the definition of x_1. Thus, $\alpha - \tilde{\alpha} \neq 0$ in $(x_1, x_2]$. By continuity, our assumption (A 5) on f_2 implies that $f_2(t, T, H(t, x, T_0)) \neq 0$ for $t > 0$ small enough and for all $x \in [0, 1]$. Finally, there exists $\rho \in (0, t^*)$ such that

$$(\alpha - \tilde{\alpha})(x)\, f_2(t, T, H(t, x, T_0)) \neq 0 \text{ for all } t \in [0, \rho] \text{ and } x \in (x_1, x_2].$$

(ii) *We show that the assumption* $x_1 < 1$ *of step (i) leads to a contradiction.*

From step (i), we know that the quantity $(\alpha - \tilde{\alpha})(x)\, f_2(t, T, H(t, x, T_0))$ has a constant strict sign in $[0, \rho] \times (x_1, x_2]$. Assume that $(\alpha - \tilde{\alpha})(x)\, f_2(t, T, H(t, x, T_0)) > 0$ in $[0, \rho] \times (x_1, x_2]$. Then, computing (C 5) at $t = 0$ and $x = x_2$, we obtain

$$c(x_2, H(0, x_2, T_0))\, \frac{\partial V}{\partial t} = (\alpha - \tilde{\alpha})(x_2)\, f_2(0, T_0(0, x_2), H(0, x_2, T_0)) > 0.$$

Thus, from the assumption (A 1) on c, we have $\frac{\partial V}{\partial t}(0, x_2) > 0$. This implies the existence of some time $\rho' \in (0, \rho)$ such that

$$V(t, x_2) > 0 \text{ for all } t \in (0, \rho'). \tag{C 6}$$

Moreover, the assumption (3.5) of Theorem 3.1 implies that $V(t, x_0) = 0$ for all $t \in (0, \theta)$. Setting

$$\theta' = \min(\rho', \theta),\ K = \max_{t \in [0, \theta'],\ x \in [x_0, x_2]} \frac{\mu_1 + \tilde{\alpha}\, \mu_2}{c(x, H(t, x, T_0))},\ r(t, x) = \mu_1 + \tilde{\alpha}\, \mu_2 - K\, c \leq 0,$$

and

$$W(t, x) = V(t, x)\, e^{-K\, t},$$

we observe that W verifies:

$$c(x, H(t, x, T_0))\, \frac{\partial W}{\partial t} - \frac{\partial}{\partial x}\left(k(x)\, \frac{\partial W}{\partial x}\right) - r(t, x)\, W \geq 0, \tag{C 7}$$

for all $t \in (0, \theta')$ and $x \in (x_0, x_2)$ together with the initial and boundary conditions:

$$\begin{cases} W(0, x) = 0, \text{ for all } x \in [x_0, x_2], \\ W(t, x_0) = 0 \text{ and } W(t, x_2) > 0, \text{ for all } t \in (0, \theta'). \end{cases} \tag{C 8}$$

Finally, the function W satisfies all the assumptions of the strong maximum principle (Theorems 5 and 7 in Chap. 3 of [58]), which implies that either $W > 0$ in $(0, \theta') \times (x_0, x_2)$ or there exists some time $t_0 \in (0, \theta')$ such that $W \equiv 0$ in $(0, t_0) \times (x_0, x_2)$. From the boundary condition satisfied by W at x_2, we necessarily have $W > 0$ in $(0, \theta') \times (x_0, x_2)$.

Since $W > 0$ in $(0, \theta') \times (x_0, x_2)$ and $W(t, x_0) = 0$, the Hopf's Lemma (Theorem 6 and 7 in Chap. 3 of [58]) implies that

$$\frac{\partial W}{\partial x}(t, x_0) > 0 \text{ for all } t \in (0, \theta'),$$

which contradicts the assumption (3.5) of Theorem 3.1.

The case $(\alpha - \tilde{\alpha})(x)\, f_2(t, T, H(t, x, T_0)) < 0$ in $[0, \rho] \times (x_1, x_2]$ can be treated similarly and also leads to a contradiction.

As a consequence, the assumption $x_1 < 1$ is false, and therefore $\alpha \equiv \tilde{\alpha}$ on $[x_0, 1]$.

(iii) *We prove that* $\alpha \equiv \tilde{\alpha}$ *on* $[0, x_0]$.

Setting:

$$\mathcal{A}_- = \Big\{ x \leq x_0 \text{ s.t. } (\alpha - \tilde{\alpha})(y) \equiv 0 \text{ for all } y \in [x, x_0] \Big\},$$

and

$$y_1 := \begin{cases} \inf(\mathcal{A}_-) & \text{if } \mathcal{A}_- \text{ is not empty,} \\ x_0 & \text{if } \mathcal{A}_- \text{ is empty,} \end{cases}$$

we can prove, by applying the same arguments as above, that $y_1 = 0$ and consequently $\alpha \equiv \tilde{\alpha}$ on $[0, 1]$. The uniqueness result of Lemma 5.1 then implies that $T \equiv \tilde{T}$ on $[0, t^*) \times [0, 1]$. This concludes the proof of Theorem 3.1. □

Appendix D. Computation of the distribution function for the observations

From the assumptions of our statistical model (4.2)-(4.3), at each site k the joint distribution F_k of $Y_k(t_1), \ldots, Y_k(t_I)$ is

$$F_k(y_1, \ldots, y_i) = \int_{\mathbb{R}_+^I} \prod_{i=1}^{I} \phi\{y_i \mid s(t_i)\} h(s_1, \ldots, s_I) ds_1 \ldots ds_I, \tag{D 1}$$

where $\phi(\cdot \mid s(t_i))$ is the probability density function of the conditional Gaussian distribution $\mathcal{N}\left\{T(s(t_i), S_k), \sigma^2\right\}$ and h is the joint density function of $s(t_1), \ldots, s(t_I)$. Equation (D 1) can also be written:

$$\begin{aligned} F_k(y_1, \ldots, y_i) &= \int_{\mathbb{R}_+^I} \left[\prod_{i=1}^{I} \phi \left\{ y_i \mid \theta - \sum_{j=1}^{i} \eta_j \right\} \right] \left[\prod_{i=1}^{I} g_{i,k}(\eta_i) \right] d\eta_1 \ldots d\eta_I \\ &= \int_{\mathbb{R}_+^I} \prod_{i=1}^{I} \left[\phi \left\{ y_i \mid \theta - \sum_{j=1}^{i} \eta_j \right\} g_{i,k}(\eta_i) \right] d\eta_1 \ldots d\eta_I, \end{aligned}$$

where $g_{i,k}$ is the probability density function of the gamma distribution $\Gamma\{(t_{i-1} - t_i)/\kappa^2, \kappa^2\}$.

If we consider a finite set of locations S_k, $k = 1, 2, 3$, the joint distribution of $\{Y(t_i, S_k),\ k = 1, \dots, 3\}$ is

$$F(\mathcal{Y}) = \prod_{k=1}^{3} \int_{\mathbb{R}_+^I} \prod_{i=1}^{I} \left[\phi_k \left\{ y_{i,k} \mid \theta - \sum_{j=1}^{i} \eta_j \right\} g_{i,k}(\eta_i) \right] d\eta_1 \dots d\eta_I, \tag{D 2}$$

where $\mathcal{Y} = \{y_{i,k},\ i = 1, \dots, I,\ k = 1, 2, 3\}$, $\phi_k(\cdot \mid \theta - \sum_{j=1}^{i} \eta_j)$ is the probability density function of the conditional Gaussian distribution $\mathcal{N}\left\{T(\theta - \sum_{j=1}^{i} \eta_j, S_k), \sigma^2\right\}$, and $g_{i,k}$ is the probability density function of the gamma distribution $\Gamma\{(t_{i-1} - t_i)/\kappa^2, \kappa^2\}$.

Appendix E. Metropolis-Hastings algorithm

For the implementation of the MCMC algorithm, the likelihood function of the model $\mathcal{L}_{\mathcal{Y}}(\alpha, T_0)$ is computed as follows: As mentioned above, given a value for the parameter α and given the initial condition T_0, the equation (4.1) admits a solution T, which is defined for all $t \in [-\tau, \theta]$ and all $x \in [0, 1]$. Thus, we are able to compute the integrals in (D 2) and the likelihood $\mathcal{L}_{\mathcal{Y}}(\alpha, T_0)$ follows.

For our computations, we discretized the space into 31 positions $x_j = j/30$, for $j = 0, \dots, 30$ and we assumed that $\alpha(x)$ was constant equal to α_j on each interval (x_j, x_{j+1}). The Metropolis-Hastings algorithm is an iterative rejection-sampling algorithm with steps that are detailed below:

Start at $k = 0$: initialize $\alpha^0 = \{\alpha_0^0, \dots, \alpha_{30}^0\}$.
while $k \leq N$

- Draw an index j with a uniform law in $\{0, \dots, 30\}$.
- Draw $\hat{\alpha_j}$ from a proposal distribution $Q(\hat{\alpha_j} | \alpha_j^k)$.
- Choose randomly with an uniform law $\zeta \in (0, 1)$.
- Compute $\delta = \dfrac{\mathcal{L}_{\mathcal{Y}}(\hat{\alpha}_j, T_0)\pi_1(T_0)\pi_2(\hat{\alpha}_j)Q(\alpha_j^k | \hat{\alpha}_j)}{\mathcal{L}_{\mathcal{Y}}(\alpha_j, T_0)\pi_1(T_0)\pi_2(\alpha_j)Q(\hat{\alpha}_j | \alpha_j^k)}$.
- **If** $\zeta < \delta$, $\alpha_j^{k+1} = \hat{\alpha}_j$ **else** $\alpha_j^{k+1} = \alpha_j^k$.
- $\alpha_i^{k+1} = \alpha_i^k$, for $i \neq j$.
- $k \leftarrow k + 1$

endwhile

We used normal distributions for the proposals. Namely, at each step k, $\hat{\alpha}_j$ was drawn from a normal distribution with mean α_j and variance 20. For our computations, the number of step was set to $N = 2 \cdot 10^5$.

Appendix F. Computation of $\overline{R}_\varepsilon$

We define the L^2-response of the model (4.1) over the time interval $[0, \theta]$ to a perturbation $\xi \in C([0, 1])$, by

$$R_{\alpha;\alpha+\xi}(\theta) := \frac{1}{\sqrt{\theta}} \|T_\alpha - T_{\alpha+\xi}\|_{L^2(0,\theta)}, \tag{F 1}$$

where $T_\alpha = T_\alpha(t)$ corresponds to the spatially averaged solution of (4.1) and $T_{\alpha+\xi} = T_{\alpha+\xi}(t)$ is the spatially averaged solution of (4.1) where α has been replaced by $\alpha + \xi$.

In order to have a general character, the response of a system cannot rely on a particular perturbation. In that respect it is natural to define an average response of the system to an arbitrary perturbation. Let $\mathcal{Z}$ be a stationary Gaussian random field defined over $[0, 1]$ with mean 0 and autocovariance function $C(x, x') = \exp(-100\|x - x'\|)$, for all x, x' in $[0, 1]$, where $\|\cdot\|$ corresponds to the Euclidian distance.

We define the average L^2-response of our model over $[0, \theta]$ to perturbations of magnitude (standard deviation) ε as the mean value $\overline{R}_\varepsilon$ of $R_{\alpha;\alpha+\varepsilon\,\mathcal{Z}}(\theta)$. The numerical computation of $\overline{R}_\varepsilon$ has been carried out using 10^4 realizations z of the random field $\mathcal{Z}$.

Other characterizations of the changes in model output, due to variations in the parameters, were considered in [59], where the authors investigated an inverse heat conduction problem. See also [60] for related questions.

References

[1] Budyko, M. I. 1969 The effect of solar radiation variations on the climate of the Earth. *Tellus*, **21**, 611–619; doi:10.1111/j.2153-3490.1969.tb00466.x.

[2] Sellers, W. D. 1969 A global climatic model based on the energy balance of the Earth atmosphere system. *J. Appl. Meteorol.*, **21**, 391–400.

[3] Ghil, M. & Childress, S. 1987 *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics*. Springer, New York.

[4] Held, I. & Suarez, M. 1974 Simple albedo feedback models of the icecaps. *Tellus*, **36**, 613–629; doi:10.1111/j.2153-3490.1974.tb01641.x.

[5] North, G. R., Cahalan, R. F. & Coakley, J. A. 1981 Energy balance climate models. *Rev. Geophys. Space Phys.*, **19**, 91–121; doi:10.1029/RG019i001p00091.

[6] Ghil, M. 1976 Climate stability for a Sellers-type model. *J. Atmos. Sci.*, **33**, 3–20.

[7] Arcoya, D., Diaz, J. I. & Tello, L. 1998 S-shaped bifurcation branch in a quasilinear multivalued model arising in climatology. *J. Diff. Equations*, **150**, 215–225; doi:10.1006/jdeq.1998.3502.

[8] Hetzer, G. 1996 Global existence, uniqueness, and continuous dependence for a reaction-diffusion equation with memory. *Electronic J. Diff. Equations*, **5**, 1–16.

[9] Diaz, J. I., Hetzer, G. & Tello, L. 2006 An energy balance climate model with hysteresis. *Nonlinear Analysis*, **64**, 2053–2074; doi:10.1016/j.na.2005.07.038.

[10] Benzi, R., Sutera, A., & Vulpiani, A. 1981 The mechanism of stochastic resonance. *J. Phys. A*, **14**, 453–457; doi:10.1088/0305-4470/14/11/006.

[11] Imkeller, P. 2001 Energy balance models – viewed from stochastic dynamics in: *Stochastic Climate Models*, P. Imkeller, J-S Von Storch (Eds.), Prog. Prob. **49**, Birkhäuser, 213–240.

[12] North, G. R., Mengel, J. G., & Short, D. A. 1983 Simple energy balance model resolving the seasons and the continents: Application to the astronomical theory of the ice ages. *J. Geophys. Res.*, **88**, 6576–6586; doi:10.1029/JC088iC11p06576.

[13] Stone, D. A., Allen, M. R., Selten, F., Kliphuis, M. & Stott, P. A. 2007 The detection and attribution of climate change using an ensemble of opportunity. *J. Climate*, **20**, 504–516; doi:10.1175/JCLI3966.1.

[14] Ghil, M., 1994 Cryothermodynamics: The chaotic dynamics of paleoclimate, *Physica D*, **77**, 130–159; doi:10.1016/0167-2789(94)90131-7.

[15] Bermejo, R., Carpio, J., Diaz, J. I. & Tello, L. 2008 Mathematical and numerical analysis of a nonlinear diffusive climate energy balance model, *Math. Comput. Modelling*, **49**, 1180–1210; doi:10.1016/j.mcm.2008.04.010.

[16] Ghil, M., Chekroun, M. D. & Simonnet, E. 2008 Climate dynamics and fluid mechanics: Natural variability and related uncertainties. *Physica D,* **237**, 2111–2126; doi:10.1016/j.physd.2008.03.036.

[17] Bhattacharya, K., Ghil, M. & Vulis, I. L. 1982 Internal variability of an energy-balance model with delayed albedo effects. *Journal of the Atmospheric Sciences*, **39**, 1747–1773.

[18] Chekroun, M.D., Simonnet, E. & Ghil, M. 2011 Stochastic climate dynamics: Random attractors and time-dependent invariant measures. *Physica D*, in press; doi:10.1016/j.physd.2011.06.005.

[19] Hetzer, G. 1995 A functional reaction-diffusion equation from climate modeling: S-shapedness of the principal branch of fixed points of the time-1-map. *Differential and Integral Equations*, **8**, 1047–1059; doi:10.1016/S0362-546X(97)00119-3.

[20] Diaz J. I. & Hetzer, G. 1997 A quasilinear functional reaction-diffusion equation from climate modeling. *Nonlinear Anal.*, **30**, 2547–2556.

[21] Diaz J. I. 2012 *Diaz et al. in this volume.*

[22] Berliner, L. M. 2003 Physical-statistical modeling in geophysics. *J. Geophys. Res.*, **108**, 8776; doi:10.1029/2002JD002865.

[23] Wikle, C. K. 2003 Hierarchical models in environmental science. *International Statistical Review*, **71**, 181–199; doi:10.1111/j.1751-5823.2003.tb00192.x.

[24] Roques, L., Soubeyrand, S. & Rousselet, J. 2011 A statistical-reaction-diffusion approach for analyzing expansion processes. *J. Theor. Biol.*, **274**, 43–51; doi:10.1016/j.jtbi.2011.01.006.

[25] Salamatin, A., Lipenkov, V., Barkov, N., Jouzel, J., Petit, J. & Raynaud, D. 1998 Ice core age dating and paleothermometer calibration based on isotope and temperature profiles from deep boreholes at Vostok Station (East Antarctica). *J. Geophys. Res.*, **103**, 8963–8977; doi:10.1029/97JD02253.

[26] Parrenin, F., Rémy, F., Ritz, C., Siegert, M. J. & Jouzel, J. 2004 New modeling of the Vostok ice flow line and implication for the glaciological chronology of the Vostok ice core. *J. Geophys. Res.*, **109**; doi:10.1029/2004JD004561.

[27] Crafoord, C., & Källén, E. 1978 A note on the condition for existence of more than one steady state solution in Budyko-Sellers type models. *J. Atmos. Sci.*, **35**, 1123–1125.

[28] Fraedrich, K. 1978 Structural and stochastic analysis of a zero-dimensional climate system. *Q. J. R. Meteorol. Soc.*, **104**, 461–474; doi:10.1002/qj.49710444017.

[29] Fraedrich, K. 1979 Catastrophes and resilience of a zero-dimensional climate system with ice-albedo and greenhouse feedback. *Q. J. R. Meteorol. Soc.*, **105**, 147–167; doi:10.1002/qj.49710544310.

[30] Ghil, M. 1984 Climate sensitivity, energy balance models, and oscillatory climate models. *J. Geophys. Res.*, **89**, 1280–1284; doi:10.1029/JD089iD01p01280.

[31] Wu, J. 1996 *Theory and Applications of Partial Functional Differential Equations*, Applied Mathematical Sciences, vol. 119. Springer-Verlag, New York.

[32] Sell, G. R. & You, Y. 2002 *Dynamics of Evolutionary Equations*, Applied Mathematical Sciences, vol. 143, Springer-Verlag, New York.

[33] Amann, H. 2006 Quasilinear parabolic functional evolution equations, In: M. Chipot, H. Ninomiya (editors), *Recent Advances in Elliptic and Parabolic Issues*. Proc. 2004 Swiss–Japanese Seminar, World Scientific, pp. 19–44.

[34] Pao, C. V. 1992 *Nonlinear Parabolic and Elliptic Equations*. Plenum Press, New York.

[35] Klibanov, M. V. & Timonov, A. 2004 *Carleman Estimates for Coefficient Inverse Problems and Numerical Applications*. Inverse and Ill-Posed Series, VSP, Utrecht.

[36] Williams, D. F., Lerche, I. & Full, W. E. 1988 *Isotope Chronostratigraphy: Theory and Methods*. Geology Series, Academic Press, San Diego.

[37] Yamamoto, M. & Zou, J. 2001 Simultaneous reconstruction of the initial temperature and heat radiative coefficient. *Inverse Problems*, **17**, 1181–1202; doi:10.1088/0266-5611/17/4/340.

[38] Belassoued, M. & Yamamoto, M. 2006 Inverse source problem for a transmission problem for a parabolic equation. *J. Inverse Ill-Posed Problems*, **14**(1), 47–56; doi:10.1515/156939406776237456.

[39] Cristofol, M. & Roques, L. 2008 Biological invasions: Deriving the regions at risk from partial measurements. *Math. Biosciences*, **215**(2), 158–166; doi:10.1016/j.mbs.2008.07.004.

[40] Roques, L. & Cristofol, M. 2010 On the determination of the nonlinearity from localized measurements in a reaction-diffusion equation. *Nonlinearity*, **23**, 675–686; doi:10.1088/0951-7715/23/3/014.

[41] Cristofol, M., Garnier, J., Hamel, F. & Roques, L. 2011 Uniqueness from pointwise observations in a multi-parameter inverse problem. *Commun. Pure Appl. Math.*, **in press**.

[42] Metropolis, N., Rosenbluth, A. W., Rosenbluth, N. M., Teller, A. H. & Teller, E. 1953 Equation of state calculations for fast computing machines. *J. Chem. Phys.*, **21**, 1087–1092.

[43] Hastings, W. K. 1970 Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109; doi: 10.1093/biomet/57.1.97.

[44] Farrell, B. F. 1988 Optimal excitation of perturbations in viscous shear flow. *Phys. Fluids*, **31**, 2093–2102; doi:10.1063/1.866609.

[45] Reddy, S. C., Schmid, P. J. & Henningson, D. S. 1993 Pseudospectra of the Orr-Sommerfeld operator. *SIAM J. Appl. Math.*, **53** , 15–47; doi:10.1137/0153002.

[46] Trefethen, L. N., Trefethen, A. E., Reddy, S. C. & Driscoll, T. A. 1993 Hydrodynamic stability without eigenvalues. *Science*, **261**, 578–584; doi:10.1126/science.261.5121.578.

[47] Trefethen, L. N. & Embree, M. 2005 *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton Univ. Press.

[48] Chekroun, M. D., Kondrashov, D. & Ghil, M. 2011 Predicting stochastic systems by noise sampling, and application to the El Niño-Southern Oscillation. *Proc. Natl. Acad. Sci USA*; doi:10.1073/pnas.1015753108.

[49] Cavaterra, C., Lorenzi, A. & Yamamoto, M. 2006 A stability result via Carleman estimates for an inverse source problem related to a hyperbolic integro-differential equation. *Comput. Appl. Math.*, **25**, 229–250; doi:10.1590/S0101-82052006000200007.

[50] National Research Council 2006 *Surface Temperature Reconstructions for the Last 2000 Years*. National Academies Press, Washington, DC, 196 pp.

[51] Solomon, S., et al. (Eds.) 2007 *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* Cambridge University Press, Cambridge & New York, 2007.

[52] Ghil, M., Cohn, S. E., Tavantzis, J., Bube, K. P., & Isaacson, E. 1981 Applications of estimation theory to numerical weather prediction. In *Dynamic Meteorology: Data Assimilation Methods*, L. Bengtsson, M. Ghil & E. Källén (Eds.), Springer Verlag, pp. 139–224.

[53] Ghil, M., & Malanotte-Rizzoli, P. 1991 Data assimilation in meteorology and oceanography. *Adv. Geophys.*, **33**, 141–266.

[54] Talagrand, O. 2011 *In this volume.*

[55] Kondrashov, D., Sun, C.-J. & Ghil, M. 2008 Data assimilation for a coupled ocean-atmosphere model. Part II: Parameter estimation. *Mon. Wea. Rev.*, **136**, 5062–5076; doi:10.1175/2008MWR2544.1.

[56] Lorenz, E. N. (1963) Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.

[57] Chekroun, M. D., Di Plinio, F., Glatt-Holtz, N. E. & Pata, V. 2011 Asymptotics of the Coleman-Gurtin model, *Discrete Contin. Dyn. Syst., Serie S*, **4** , no. 2, 351–369; doi:10.3934/dcdss.2011.4.351.

[58] Protter, M. H. & Weinberger, H. F. 1967 *Maximum Principles in Differential Equations.* Prentice-Hall, Englewood Cliffs, NJ.

[59] Jarny, Y., Ozisik, M. N. & Bardon, J. P. 1991 A general optimization method using adjoint equation for solving multidimensional inverse heat conduction. *Int. J. Heat Mass Transfer*, **34** (11), 2911–2919; doi:10.1016/0017-9310(91)90251-9.

[60] Singler, J.R. 2008 Transition to turbulence, small disturbances, and sensitivity analysis I: A motivating problem. *J. Math. Anal. Appl.*, **337**, 1425–1441; doi:10.1016/j.jmaa.2007.03.094.