Parameter estimation for energy balance models with memory

Lionel Roques¹, Mickaël D. Chekroun², Michel Cristofol³, Samuel Soubeyrand¹ and Michael Ghil²,⁴

¹UR 546 Biostatistique et Processus Spatiaux, INRA, 84000, Avignon, France
²Department of Atmospheric and Oceanic Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095-1565, USA
³Université Aix-Marseille, LATP, 13397 Marseille, France
⁴Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL), Ecole Normale Supérieure, Paris, France

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 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 that 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.
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 EBMs, the temperature $T$ is globally averaged and thus depends on time alone, and not on any space variables. In this simplification, 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 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 one-dimensional 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 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 two-dimensional 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 behaviour 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, Bhattacharya et al. 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 behaviour 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 for further details and [20] for a 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 Intergovernmental Panel on Climate Change (IPCC)-class GCMs, or in the modelling 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 this 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, subjected to reasonable mathematical assumptions about the PDE and its solutions. However, climate data are generally local and noisy, that is, $T(t, x)$ can be measured only at some points or in some (often small) subregions of the domain $(0, 1)$, and the measurements are subjected 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 one-dimensional 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 $D = (0, 1)$ by using local information about the temperature $T$ that is available only on a small subdomain $E$ 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 [21–23], 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.
2. Energy balance models

(a) Classical energy balance models

One of the simplest zero-dimensional EBMs,

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

has been analysed in [3,26]. 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 $\mu$ thus correspond to changes in these conditions and lead to changes in the number and stability of the model’s equilibria [27,28].

The functions $a(T(t))$ and $g(T(t))$ in the incoming and outgoing radiation terms of equation (2.1), respectively, correspond to the albedo function and to the greyness function; the latter models the difference between black-body radiation $\sigma T^4$—with $\sigma$ 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, the presence of clouds [17], and on the wave distribution on the ocean surface [29].

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

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

where $a_0 > a_1 > 0$ and $T_1 < T_2$ are all constants. The values of $T_1$ and $T_2$ are usually between $-20^\circ$C and $0^\circ$C; see figure 3 in [30]. Other albedo functions have been proposed for the study of more local problems, in restricted temperature ranges; for instance, Fraedrich [27] proposed linear or quadratic functions of the form

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

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

The function $g$ is positive and can take several forms as well [1,2,6,17,29]. 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 equation (2.1) [3].

The one-dimensional reaction–diffusion model presented in [6,29] has the form

$$c(x) \frac{dT}{dt} = \frac{\partial}{\partial x} \left( k(x) \frac{\partial T}{\partial x} \right) + \mu Q(x)[1 - a(x, T)] - g(x, T),$$  (2.3)

for $t > 0$ and $x \in (0, 1)$. On the RHS of (2.3), 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 $\phi$ 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 (2.3), 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 \([6,17,29]\).

(b) Energy balance models with memory terms

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

$$H(t, x, T) = \int_{-\tau}^{0} \beta(s, x) T(t + s, x) \, ds,$$

for all $t > 0$ and $x \in (0, 1)$, \hspace{1cm} (2.4)

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

$$\int_{-\tau}^{0} \beta(s, \cdot) \, ds \equiv 1 \quad \text{and} \quad \beta(s, \cdot) \geq 0 \quad \text{for all } s \in [-\tau, 0].$$

(2.5)

At this level of generality, one-dimensional EBMMs can be described by the equation

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

for $t > 0$ and $x \in (0, 1)$. As in (2.3), the reaction term $f$ in the RHS of (2.6) describes the balance between incoming and outgoing radiation. In the general setting of equation (2.6), the reaction term depends not only on the space position $x$ and on the current temperature $T$, but also on time $t$ and on the history function $H(t, x, T)$. The dependence on $H(t, x, T)$ corresponds to the delayed response of the incoming and outgoing radiation functions, such as the albedo function $a$ [11,17].

(c) Initial and boundary conditions

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

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

(2.7)

for some function $T_0$ defined on $[-\tau, 0] \times [0, 1]$, which has Lipschitz-continuous derivatives up to order 2 with respect to $x$ and up to order 1 with respect to $t$.

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

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

(2.8)

Note that the assumption (2.8) also applies at $t = 0$. This compatibility condition is required to obtain the time continuity of the solution $T$ of (2.6) up to $t = 0$.

When $c$ does not depend on $H$, the use of the theory of semigroups typically allows a recasting of equation (2.6), together with the initial and boundary conditions (2.7) and (2.8), 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 (2.6)-(2.8) 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].
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 [39,40]. (b) For EBMMs, the region that is used in this study.

When \( c \) is a genuine function of \( H \), the problem (2.6)–(2.8) becomes a \textit{quasi-linear 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 quasi-linear 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 (2.6)–(2.8), 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 \((0,1)\) [36–38] (figure 1a). For polynomial reaction terms, more recent approaches [39,40] 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 \( T(0,x) \) are known over the entire interval \((0,1)\). Thus, one of the main differences between these two approaches is that, in the method presented in [39,40], 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 1a. Contrary 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 show that comparable results can be obtained for EBMMs with memory terms and nonlinear reaction terms (figure 1b).

We assume here that the function \( f \) in equation (2.6) has the form

\[
f[\alpha](t,x,T,H(t,x,T)) = f_1(t,x,T,H(t,x,T)) + \alpha(x)f_2(t,T,H(t,x,T)),
\]

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 \( \alpha \).

More precisely, consider a solution \( \tilde{T} \) of equation (2.6), where \( f[\alpha] = f_1 + \alpha f_2 \) 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. (2.7) and (2.8), respectively—as \( T \). Our goal is to prove that for any subset of positive measure \( \mathcal{E} \subset (0,1) \),
and any $\theta > 0$,
\[ \{ \tilde{T}(t, x) = T(t, x) \text{ in } (0, \theta) \times \mathcal{E} \} \Rightarrow \{ \tilde{\alpha}(x) \equiv \alpha(x) \text{ over } [0, 1] \}. \]  
(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 $\beta$ of the history function $H$ in equations (2.4) and (2.5), namely that
\[ \exists \delta > 0 \text{ such that } \beta(s, \cdot) \equiv 0 \quad \text{for all } s \in [-\delta, 0]. \]  
(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 $\delta$, 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 (2.6).

Under assumption (3.3) and the regularity assumptions of appendix A, we prove in appendix B that the problem (2.6) with initial condition (2.7) and boundary condition (2.8) admits a unique classical solution, at least for small times $t < t^*$, and $t^*$ does not depend on $\alpha$. Global existence in time of solutions to equations (2.6)–(2.8) 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 need the solution $T$ to be defined only for small times. Further existence results for parabolic equations with delay terms are given in references [31,34].

(a) Uniqueness result

We assume that the unknown coefficient $\alpha(x)$ of (3.1) lies in the function space
\[ \mathcal{M} := \{ \psi \text{ is Lipschitz-continuous and piecewise analytic on } [0, 1] \}. \]  
(3.4)

**Remark 3.1.** A continuous function $\psi$ is called piecewise analytic if there exist $n \geq 1$ and an increasing sequence $(k_j)_{1 \leq j \leq n}$ such that $k_1 = 0$, $k_n = 1$, and
\[ \psi(x) = \sum_{j=1}^{n-1} \chi_{[k_j, k_{j+1})} \varphi_j(x), \quad \text{for all } x \in (0, 1); \]
here $\varphi_j$ are some analytic functions defined on the intervals $[k_j, k_{j+1}]$, and $\chi_{[k_j, k_{j+1})}$ are the characteristic functions of the intervals $[k_j, k_{j+1})$ for $j = 1, \ldots, n - 1$.

**Remark 3.2.** 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 3.3.** Let $\alpha$ and $\tilde{\alpha}$ satisfy assumption (3.4). Let $T$ and $\tilde{T}$ be the solutions of equations (2.6)–(2.8) on $[0, t^*) \times [0, 1]$, with $f = f(\alpha)$ and $f = f(\tilde{\alpha})$, respectively, cf. equation (3.1) and $t^* \in (0, \delta)$. Assume that there exist $x_0 \in [0, 1]$ and $\theta \in (0, t^*)$ such that
\[ \begin{cases} T(t, x_0) = \tilde{T}(t, x_0), \\ \frac{\partial T}{\partial x}(t, x_0) = \frac{\partial \tilde{T}}{\partial x}(t, x_0), \end{cases} \quad \text{for all } t \in (0, \theta). \]  
(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 3.3 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 §3a, we showed that space-dependent coefficients of general EBMMs can be uniquely determined by local measurements of the temperature $T$. Here, 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 [24]. 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 [21–23]. 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.

(a) An energy balance models with memory with unknown coefficient

We assume that the temperatures are given by an EBMM corresponding to a particular case of equation (2.6), with

\[
\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \alpha(x)(1 - a(T)) - q_0 - q_1 T - \left( \frac{1}{\tau} \int_{-\tau}^{0} T(t + s) \, ds \right)^3,
\]

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, $\tau = 0.2$ and $\tau = 0.7$ kyr, during the time interval $0 \leq t \leq 5$ kyr, where 1 kyr = 1000 years. We refer to §4d for the precise numerical setting of equation (4.1).

(b) 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 $\theta = 5 \geq t_1 > \cdots > t_I \geq -\tau$, 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 owing 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 [41]. Furthermore, the uncertainty in measuring the temperature value implies that $T(s(t_i), S_k)$ also contains noise.

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 $\mathcal{N}$:

\[
Y_k(t_i) \mid s(t_i) \sim \text{indep. } \mathcal{N}\{T(s(t_i), S_k), \sigma^2\},
\]

where $\sigma^2$ is the variance of the noise in the temperature measurements.
Second, we construct a model for \( s(t_i) \)
\[
s(t_i) = \theta - \sum_{j=1}^{i} \eta_j \text{ with } \eta_j \sim \text{indep. } \Gamma \left( \frac{t_{j-1} - t_j}{k^2}, k^2 \right),
\]
where \( \Gamma \) denotes the gamma distribution, \( k^2 \) is a positive parameter that controls the shape of the distribution and \( t_0 = \theta \). The expectation of \( s(t) \) in equation (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 \( t_j > t_i \), then \( s(t_j) > s(t_i) \). 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 equations (4.2) and (4.3). Using this model, we are able to compute the joint distribution of the observations \( Y := \{ Y_k(t_i), i = 1, \ldots, I, k = 1, 2, 3 \} \), conditioned on the ‘real’ temperatures \( T := \{ T(s(t_i), S_k), i = 1, \ldots, I, k = 1, 2, 3 \} \). This joint distribution, denoted by \( F(Y) \), is computed explicitly in appendix D. If the temperatures \( T = \{ T(s(t_i), S_k) \} \) are governed by the EBMM (4.1), then they depend deterministically on the unknown coefficient \( \alpha \) and the unknown initial data \( T_0 \) of equation (2.7). Thus, the conditional distribution \( F(Y) \) of the observation process is equal to the likelihood \( L_Y(\alpha, T_0) \) of the model parametrized by \( \alpha \) and \( T_0 \).

(c) Bayesian inference

Let \( T_0 \) correspond to the temperatures during the time window \([-\tau, 0]\) with prior distribution \( \pi_1 \). Let \( \alpha \) denote the coefficient of our model with prior distribution \( \pi_2 \); we assume that the other coefficients and parameters are known. The posterior distribution \( p(T_0, \alpha \mid Y) \) of \( T_0 \) and \( \alpha \) is proportional to
\[
p(T_0, \alpha \mid Y) \propto L_Y(\alpha, T_0)\pi_1(T_0)\pi_2(\alpha).
\]
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
\[
\pi(T_0) \sim \delta_K,
\]
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.

(d) Numerical results

For a given coefficient \( \alpha(x) \), we computed the solution \( T(t, x) \) of the EBMM (4.1) for \( \tau = 0.2 \) kyr and \( \tau = 0.7 \) kyr, emanating from the same constant initial history \( T_0 = 10^\circ C \) given for \([-\tau, 0]\) and \( x \in [0, 1] \). The parameter values used in these computations for solving equation (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 equation (2.2).

When \( \tau = 0.2 \) kyr, the solution of (4.1) exhibits small temporal variations that are quickly damped and a stable steady state is quickly reached (figure 2a). For \( \tau = 0.7 \), a stable periodic orbit \( T_p(t, x) = T_p(t + p, x) \) of period \( p \approx 1.33 \) kyr is reached asymptotically, but the transient is considerably longer and larger amplitudes persist for quite a while (figures 2b and 3). Such large-amplitude variations that precede the setting in of the periodic solution \( T_p \) point to the presence of non-normal modes associated with the linearization of equation (4.1) about the stable periodic solution \( T_p \). Such non-normal modes are associated with polynomial rather than with exponential behaviour in time and they are responsible for large transient energy growth in the linearized
Figure 2. The solution $T(t,x)$ of the EBMM model (4.1). (a) $\tau = 0.2$ kyr; (b) $\tau = 0.7$ kyr. (Online version in colour.)

Figure 3. Time series of the global average temperature. The case $\tau = 0.7$ kyr exhibits a transient with large variations that suggest the presence of non-normal modes [44]; see supplementary information in [45] for a similar signature of such modes in a simpler model. (Online version in colour.)

problem. This growth can be further amplified by nonlinear effects or by noise; see [44,46–48] for many manifestations of such modes and the supporting information of [45] 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. equation (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 $\tau = 0.2$ kyr and $\tau = 0.7$ kyr. 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 made only 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 $\tau = 0.2$ kyr.
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, whereas the lower row corresponds to the measured temperatures at estimated times. (a) $\tau = 0.2$ kyr; (b) $\tau = 0.7$ kyr. (Online version in colour.)

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 +. (a) $\tau = 0.2$ kyr; (b) $\tau = 0.7$ kyr.

Overall, the main difference between figure 5a,b is that the marginal distribution is more variable in the case $\tau = 0.2$ kyr, 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 $\tau = 0.7$ kyr 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 $\tau = 0.7$ kyr.

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

The results are presented in figure 6 for $\tau = 0.2$ kyr (blue curve) and for $\tau = 0.7$ kyr (red curve). In both cases, the $L^2$-response increases linearly with $\varepsilon$, but the slope is more than twice as large in the latter case. As noted above, the non-normal 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 [44,47]). 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 EBMMs can be determined or estimated based on localized measurements of 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, whereas 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 $E$ in space and during a short time window $0 < t < \theta$ (figure 1b). 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, $\alpha$ 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 modelled statistically by taking into account the two main problems that arise in the interpretation of such past-temperature data:
(i) age-dating errors [24,41], and (ii) converting the isotopic, microfaunal, dendrochronological or other measurements into past temperature changes [24,50]. The two kinds of errors are 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 non-normal modes [44,48].

This empirical fact has been noted in other parameter estimation problems. Considerable experience with sequential estimation has accumulated over the past three decades in meteorology, oceanography and climate dynamics [52,53]. 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 [54]. We have shown in this study that the model’s response to random parameter variations, as illustrated in [45], may help quantify these heuristic considerations and provide further insights 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 non-normal modes, might lead to more accurate parameter estimation when the observational data are subjected 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 [45], and compare it with the results for strongly mixing chaotic systems, such as the Lorenz [55] 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 [56].

Acknowledgements. The authors are grateful to D. Kondrashov for helpful discussions.

Funding statement. This study was supported by the US 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 \( \alpha \)

The heat capacity \( c \) and the diffusion coefficient \( k \) satisfy

\[
(x,u) \mapsto c(x,u) \in C^1([0,1] \times \mathbb{R}) \quad \text{and} \quad c \geq \gamma > 0 \quad \text{in} \quad [0,1] \times \mathbb{R},
\]

for some positive constant \( \gamma \), and

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

The functions \( f_1 \) and \( f_2 \) in (3.1) verify

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

and

\[
f_2(t,u,v) \in C^1(\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}).
\]
Furthermore, we assume that
\[
f_2(0, T_0(0, x), H(0, x, T_0(0, x))) \neq 0 \quad \text{for all } x \in [0, 1]. \tag{A 5}
\]
The parameter \(\alpha\) is assumed to be Lipschitz-continuous on \([0, 1]\) and verifies
\[
|\alpha(x)| < M, \quad \forall x \in [0, 1],
\]
for some positive constant \(M\).

### Appendix B. Existence and uniqueness of the solution of the general energy balance models with memory for small times

The assumptions of appendix A on \(c, f_1, f_2, \alpha\) imply that we can define two Lipschitz-continuous functions \(f^-\) and \(f^+\), which are independent of the (unknown) function \(\alpha\) and such that
\[
f^-(T) \leq \min_{x \in [0,1], \xi \in [0, \delta]} \frac{f_1(t, x, T, H(t, x, T_0)) - M[f_2(t, T, H(t, x, T_0))]}{c(x, H(t, x, T_0))} \tag{B1}
\]
and
\[
f^+(T) \geq \max_{x \in [0,1], \xi \in [0, \delta]} \frac{f_1(t, x, T, H(t, x, T_0)) + M[f_2(t, T, H(t, x, T_0))] - M}{c(x, H(t, x, T_0))}, \tag{B2}
\]
for all \(T \in \mathbb{R}\) (\(\delta\) is defined by equation (3.3)). From the Cauchy–Lipschitz theorem, there exists \(t^* \in (0, \delta)\) (which does not depend on \(\alpha\)) and two functions \(T^-\) and \(T^+\) in \(C^1[0, t^*)\) which satisfy
\[
\begin{aligned}
T^- &= f^- (T^-) \quad \text{for } t \in (0, t^*) \quad \text{and} \quad T^- (0) = \min_{x \in [0,1]} T_0(0, x) \\
T^+ &= f^+ (T^+) \quad \text{for } t \in (0, t^*) \quad \text{and} \quad T^+ (0) = \max_{x \in [0,1]} T_0(0, x).
\end{aligned}
\]
We have the following lemma B.1.

**Lemma B.1.** The problem (2.6)–(2.8) with \(f\) defined by (3.1) admits a unique classical solution \(T \in C^2([0, t^*) \times [0, 1])\) (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 B.1.** Let us set \(T^-(t, x) = T(t, x)\) and \(T^+(t, x) = T^+(t)\) for all \((t, x) \in [0, t^*) \times [0, 1]\). Because \(t^* < \delta\), the term \(H(t, x, T)\) in (2.6) is equal to \(H(t, x, T_0)\) for all \((t, x) \in [0, t^*) \times [0, 1]\). Moreover, \(T^-\) verifies, for all \((t, x) \in [0, t^*) \times [0, 1]\):
\[
c(x, H(t, x, T_0)) \frac{\partial T^-}{\partial t} - \frac{\partial}{\partial x} \left( k(x) \frac{\partial T^-}{\partial x} \right) - f[\alpha](t, x, T^-, H(t, x, T_0)) \leq 0, \tag{B 4}
\]
with the boundary conditions \((\partial T^- / \partial x)(t, 0) = (\partial T^- / \partial x)(t, 1) = 0\) for \(t \in [0, t^*)\) and the initial condition \(T^-(0, x) = T^- (0) = T_0(0, x)\) for all \(x \in [0, 1]\). Thus, \(T^-\) is a subsolution of the problem (2.6)–(2.8). Similarly, \(T^+\) is a supersolution of (2.6)–(2.8) and \(T^- \leq T^+\) for all \((t, x) \in [0, t^*) \times [0, 1]\). This implies [34] that the problem (2.6)–(2.8) admits a unique classical solution \(T \in C^2([0, t^*) \times [0, 1])\) such that \(T^-(t) \leq T(t, x) \leq T^+(t)\) for all \((t, x) \in [0, t^*) \times [0, 1]\). ■

### Appendix C. Proof of theorem 3.3

**Step 1:** we derive a linear equation for \(V = \Theta - \tilde{T}\).

The function \(V\) verifies
\[
c(x, H(t, x, T)) \frac{\partial \Theta}{\partial t} - c(x, H(t, x, \tilde{T})) \frac{\partial \tilde{T}}{\partial t} = \frac{\partial}{\partial x} \left( k(x) \frac{\partial V}{\partial x} \right) + f[\alpha](t, x, T, H(t, x, T)) - f[\tilde{\alpha}](t, x, \tilde{T}, H(t, x, \tilde{T})),
\]
for all \(t \in (0, t^*)\) and \(x \in (0, 1)\). Note that \(V(0, x) = 0\) for all \(x \in [0, 1]\).
Because \( t^* < \delta \), we observe that for all \( x \in [0, 1] \) and all \( t \in (0, t^*) \)
\[
H(t, x, T) = H(t, x, \bar{T}) = H(t, x, T_0).
\]
As a consequence, we obtain that the function \( V \) verifies
\[
c(x, H(t, x, T_0)) \frac{\partial V}{\partial t} = \frac{\partial}{\partial x} \left( k(x) \frac{\partial V}{\partial x} \right) + f(\alpha)(t, x, T, H(t, x, T_0)) - f(\bar{\alpha})(t, x, \bar{T}, H(t, x, T_0)),
\]
for all \( t \in (0, t^*) \) and \( x \in (0, 1) \). Using (3.1), (C 1) leads to
\[
c(x, H(t, x, T_0)) \frac{\partial V}{\partial t} = \frac{\partial}{\partial x} \left( k(x) \frac{\partial V}{\partial x} \right) + f_1(t, x, T, H(t, x, T_0)) - f_1(t, x, \bar{T}, H(t, x, T_0))
\]
\[
+ \alpha(x)f_2(t, T, H(t, x, T_0)) - \bar{\alpha}(x)f_2(t, \bar{T}, H(t, x, T_0)),
\]
for all \( t \in (0, t^*) \) and \( x \in (0, 1) \).

Let us set
\[
\mu_1(t, x) = \begin{cases} f_1(t, x, T, H(t, x, T_0)) - f_1(t, x, \bar{T}, H(t, x, T_0)) & \text{if } T(t, x) \neq \bar{T}(t, x), \\ \frac{\partial f_1}{\partial T}(t, x, T, H(t, x, T_0)) & \text{otherwise,}
\end{cases}
\]
and
\[
\mu_2(t, x) = \begin{cases} f_2(t, T, H(t, x, T_0)) - f_2(t, \bar{T}, H(t, x, T_0)) & \text{if } T(t, x) \neq \bar{T}(t, x), \\ \frac{\partial f_2}{\partial T}(t, T, H(t, x, T_0)) & \text{otherwise.}
\end{cases}
\]
Because \( f_1 \) 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
\[
c(x, H(t, x, T_0)) \frac{\partial V}{\partial t} = \frac{\partial}{\partial x} \left( k(x) \frac{\partial V}{\partial x} \right) + \mu_1(t, x) + \alpha(x)\mu_2(t, x) + (\alpha - \bar{\alpha})(x)f_2(t, T, H(t, x, T_0)),
\]
for all \( t \in (0, t^*) \) and \( x \in (0, 1) \).

**Step 2:** we define the largest interval \([x_0, x_1]\) where \( \alpha \equiv \bar{\alpha} \).

Let \( x_0 \) be defined as in the assumptions of theorem 3.3, and let us set
\[
A_+ = \{ x \geq x_0 \text{ such that } (\alpha - \bar{\alpha})(y) \equiv 0 \text{ for all } y \in [x_0, x]\},
\]
and
\[
x_1 := \begin{cases} \sup(A_+) & \text{if } A_+ \text{ is not empty,} \\ x_0 & \text{if } A_+ \text{ is empty.}
\end{cases}
\]

**Step 3:** we show that \( x_1 = 1 \).

Arguing by contradiction, we assume that
\[
x_1 < 1. \tag{C 6}
\]

**Step 3.1:** we prove that there exist \( \rho \in (0, t^*) \) and \( x_2 \in (x_1, 1) \) such that
\[
(\alpha - \bar{\alpha})(x)f_2(t, T, H(t, x, T_0)) \neq 0 \quad \text{for all } t \in [0, \rho] \text{ and } x \in (x_1, x_2).
From assumption (3.4), the functions $\alpha$ 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

(i) $\alpha - \tilde{\alpha} \equiv 0$ on $[x_1, x'_1]$; or

(ii) 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 (i) 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

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

\textbf{Step 3.2:} we show that the assumption (C.6) leads to a contradiction.

From step 3.1, 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 $(\partial V/\partial t)(0, x) > 0$. This implies the existence of some time $\rho' \in (0, \rho)$ such that

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

Moreover, the assumption (3.5) of theorem 3.3 implies that $V(t, x_0) = 0$ for all $t \in (0, \theta)$. Setting

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

and

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

we observe that $W$ verifies

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

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

$$W(0, x) = 0, \quad \text{for all } x \in [x_0, x_2]$$

and

$$W(t, x_0) = 0 \text{ and } W(t, x_2) > 0, \quad \text{for all } t \in (0, \theta').$$

Finally, the function $W$ satisfies all the assumptions of the strong maximum principle (theorems 5 and 7 in chap. 3 of [57]), 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 $(0, \theta') \times (x_0, x_2)$.

Because $W > 0$ in $(0, \theta') \times (x_0, x_2)$ and $W(t, x_0) = 0$, the Hopf’s lemma (theorems 6 and 7 in chap. 3 of [57]) implies that

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

which contradicts the assumption (3.5) of theorem 3.3.

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 (C.6) is false, and therefore $\alpha \equiv \tilde{\alpha}$ on $[x_0, 1]$.

\textbf{Step 4:} we prove that $\alpha \equiv \tilde{\alpha}$ on $[0, x_0]$.

Setting

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

and

\[
y_1 := \begin{cases} 
\inf (A_-) & \text{if } A_- \text{ is not empty,} \\
x_0 & \text{if } A_- \text{ is empty,}
\end{cases}
\]

we can prove, by applying the same arguments as above, that \(y_1 = 0\) and consequently \(a \equiv \bar{a}\) on \([0, 1]\). The uniqueness result of lemma B.1 then implies that \(T \equiv \bar{T}\) on \([0, t^*] \times [0, 1]\). This concludes the proof of theorem 3.3.

**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

\[
F_k(y_1, \ldots, y_I) = \int_{\mathbb{R}^I_+} \prod_{i=1}^I \phi(y_i \mid s(t_i))h(s_1, \ldots, s_I) \, ds_1 \ldots ds_I, \tag{D1}
\]

where \(\phi(\cdot \mid s(t_i))\) is the probability density function of the conditional Gaussian distribution \(\mathcal{N}(T(s(t_i), S_k), \sigma^2)\) and \(h\) is the joint density function of \(s(t_1), \ldots, s(t_I)\). Equation (D1) can also be written

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

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 \(\{Y(t_0, S_k), k = 1, \ldots, 3\}\) is

\[
F(Y) = \prod_{k=1}^3 \int_{\mathbb{R}^I_+} \prod_{i=1}^I \phi_k \left( y_{i,k} \mid \theta - \sum_{j=1}^{i} \eta_j \right) g_{i,k}(\eta_j) \, d\eta_1 \ldots d\eta_I, \tag{D2}
\]

where \(Y = \{y_{i,k}, i = 1, \ldots, I, k = 1, 2, 3\}, \phi_k(\cdot \mid \theta - \sum_{j=1}^{i} \eta_j)\) is the probability density function of the conditional Gaussian distribution \(\mathcal{N}(T(\theta - \sum_{j=1}^{i} \eta_j, S_k), \sigma^2)\), 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 \(L_Y(a, T_0)\) is computed as follows: as mentioned above, given a value for the parameter \(a\) 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 (D2) and the likelihood \(L_Y(a, T_0)\) follows.

For our computations, we discretized the space into 31 positions \(x_j = j/30\), for \(j = 0, \ldots, 30\) and we assumed that \(a(x)\) is constant equal to \(a_j\) 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 \(a^0 = (a^0_0, \ldots, a^0_{30})\).

**while** \(k \leq N\)

1. Draw an index \(j\) with a uniform law in \([0, \ldots, 30]\).
2. Draw \(\hat{a}_j\) from a proposal distribution \(Q(\hat{a}_j | a^k_j)\).
3. Choose randomly with an uniform law \(\xi \in (0,1)\).
4. Compute \(\delta = \frac{L_Y(\hat{a}_j, T_0)\pi_1(T_0)\pi_2(\hat{a}_j)Q(\hat{a}_j | a^k_j)}{L_Y(a_j, T_0)\pi_1(T_0)\pi_2(a_j)Q(a_j | a^k_j)}\).
5. If $\zeta < \delta$, $\alpha_j^{k+1} = \hat{\alpha}_j$ else $\alpha_j^{k+1} = \alpha_j^k$.
6. $\alpha_j^{k+1} = \alpha_j^k$ for $i \neq j$.
7. $k \leftarrow k + 1$

endwhile

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

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

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

$$R_{\alpha, \alpha + \xi} = \frac{1}{\sqrt{\theta}} \parallel T_\alpha - T_{\alpha + \xi} \parallel_{L^2(0, \theta)}, \quad (F1)$$

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 $\alpha$ 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 $Z$ be a stationary Gaussian random field defined over $[0, 1]$ with mean 0 and autocovariance function $C(x, x') = \exp(-100 \parallel x - x' \parallel)$, for all $x, x'$ in $[0, 1]$, where $\parallel \cdot \parallel$ corresponds to the Euclidian distance.

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

Other characterizations of the changes in model output, owing to variations in the parameters, were considered in [58], where the authors investigated an inverse heat conduction problem. See also [59] for related questions.

References


35. Klibanov MV, Timonov A. 2004 *Carleman estimates for coefficient inverse problems and numerical applications*. Inverse and Ill-Posed Series, VSP, Utrecht.


