Three-dimensional simulation and inversion of borehole temperatures for reconstructing past climate in complex settings

Peter O. Hopcroft,[1,2] Kerry Gallagher,[3] Christopher C. Pain,[1] and Fangxin Fang[1]

Received 30 September 2008; revised 23 February 2009; accepted 20 March 2009; published 3 June 2009.

[1] The majority of inversion methods used for inferring past ground surface temperatures (GST) from borehole temperature-depth profiles rely on the assumption that heat flow is in the vertical direction only. This means that accounting for certain effects caused by the local terrain of a borehole is not possible and consequently, many borehole profiles cannot be used with confidence. Here, we describe a methodology to avoid this problem by solving the heat conduction forward problem in 3-D using finite elements (FE). In order to make the inversion approach computationally tractable, we reduce the dimensions of this FE model using proper orthogonal decomposition. The inverse problem is cast in a probabilistic Bayesian framework for which the posterior probability distribution of the past GSTs is sampled using a reversible jump Markov chain Monte Carlo algorithm. This allows the resolution of the GST history over time to be explored by varying the parameterization of the GST model. Synthetic examples calculated with moderate topographies demonstrate the efficacy of the Bayesian 3-D inversion method, and the results are compared with those using a 1-D approach. For moderate topography, the latter can lead to spurious GST reconstructions. A further synthetic example demonstrates that the effect of incorrectly assuming lateral geological homogeneity is negligible. The inversion method is also compared with a more standard inversion method. A significant advantage of the Bayesian approach is that uncertainties in all of the model parameters can be accounted for, leading to a more realistic interpretation of the range of GST histories supported by the data. The methods presented here should allow a broader range of geothermal data to be used for paleoclimate reconstruction purposes in the future.

Citation: Hopcroft, P. O., K. Gallagher, C. C. Pain, and F. Fang (2009), Three-dimensional simulation and inversion of borehole temperatures for reconstructing past climate in complex settings, *J. Geophys. Res.*, *114*, F02019, doi:10.1029/2008JF001165.

1. Introduction

[2] Past ground surface temperatures (GST) can be estimated by inverting borehole temperature-depth profiles [e.g., *Lachenbruch and Marshall*, 1986; *Huang et al.*, 2000; *Pollack and Huang*, 2000; *Harris and Chapman*, 2001; *Beltrami*, 2002; *Pollack and Smerdon*, 2004]. Deviations from the equilibrium thermal regime can, in the absence of other perturbing factors, be related to past changes in ground surface temperatures through the law of heat conduction. As the thermal diffusivity of groundrock is low and of the order $10^{-6}m^2s^{-1}$, important information on the last 500–750 years can be derived from temperatures measured in boreholes of around 500 m depth. However, this low thermal diffusivity also means that the resolution over time of the derived reconstructions is very low and decreases rapidly with time before the present. Nonetheless, the reconstructions from borehole temperatures provide important constraints for past surface temperatures because there is no empirical calibration required (in contrast to other proxy data for example) and because the methods and data used are independent of these proxy data reconstruction methods.

[3] In order to find an appropriate GST history given a measured present day borehole temperature depth profile, an inverse problem is typically solved whereby the fit between the simulated temperatures in the borehole to the measured temperatures is minimized by changing the GST history (which is the upper surface boundary condition of a diffusion forward problem). Typically this is solved using either a Bayesian least squares method [e.g., *Shen and Beck*, 1992] or a singular value decomposition of a matrix of analytical solutions of the heat conduction equation [e.g., *Mareschal and Beltrami*, 1992]. The common assumptions to both methods are (1) convective heat transport is negligible, (2) heat flow at the borehole site is purely vertical and (3) the rock surrounding the borehole is laterally homogeneous; i.e., it consists of a series of horizontal layers of

[1] Applied Modelling and Computation Group, Department of Earth Science and Engineering, Imperial College, London, UK.
[2] Now at Bristol Research Initiative for the Dynamic Global Environment, School of Geographical Sciences, University of Bristol, Bristol, UK.
[3] Géosciences Rennes, Université de Rennes 1, Rennes, France.

0148-0227/09/2008JF001165

possibly differing thermal properties. However, the importance of accounting for three-dimensional effects has been highlighted in a number of geothermal studies [*Šafanda*, 1994; *Shen et al.*, 1995; *Kukkonen and Šafanda*, 1996; *Kohl*, 1999; *Kohl et al.*, 2001; *Bense and Beltrami*, 2007]. Two factors that may be important in many situations are the effects of subsurface thermal conductivity variations and the subsurface thermal signatures of varying surface topography [*Lachenbruch*, 1969; *Blackwell et al.*, 1980; *Šafanda*, 1999]. In many real data case studies, these two factors have been identified as potentially significant influences on subsurface temperatures [e.g., *Kukkonen and Šafanda*, 1996; *Kohl*, 1999; *Şerban et al.*, 2001]. For example, out of 194 temperature profiles analyzed by *Chouinard and Mareschal* [2007], 34 were rejected for climate reconstruction purposes on the grounds of topography alone.

[4] Surface topographic variations cause departures from vertical heat flow through the slope of the ground surface and the temperature distribution on this surface. This can lead to curved isotherms and in extreme cases, nearly horizontal heat flow. The variation of air temperature with altitude leads to lateral heat flow across hills and around valleys and differences in solar insolation amounts across mountains and hills cause further asymmetries in underground temperature profiles [*Blackwell et al.*, 1980]. *Kohl* [1999] compared a variety of methods for evaluating the subsurface effects of topography and found that finite element methods are best suited to simulate heat transfer in 3-D. *Kohl* [1999] also demonstrated that 1-D inversions for past GSTs of synthetic temperature-depth profiles show pronounced cooling and reduced warming when the temperature-depth profiles are located at topographic peaks, and the opposite effects for profiles located in topographic troughs. For a sinusoidal topography of amplitude 100 m and wavelength of 20 km, the topographically induced effect leads to an error in the GST reconstruction (compared with a flat surface) of 1 K at 200 years before present for a box car shaped GST model of +1 K from 200 to 100 years before present. This effect is clearly of importance when inferring GST histories for the last few hundred years in a range of natural settings.

[5] To our knowledge only one study has focused on incorporating these effects into an inverse scheme for reconstructing past GSTs [*Kohl and Gruber*, 2003], and other GST reconstruction methods rely on the assumption of purely conductive 1-D vertical heat flow. As a result many borehole data sets are excluded from paleoclimate investigation, as the assumption of purely 1-D heat flow may lead to spurious GST reconstructions.

[6] In this work we outline an inversion method for dealing with these issues by developing a scheme which utilizes a reduced-order transient 3-D finite element (FE) heat transfer forward model. This numerical approach has the advantage that it allows incorporation of complex subsurface conductivity structures as well as arbitrary surface topographies, by suitable refinement of the FE mesh. Moreover, the discretization in space can be made coarser at depth and more refined near the surface where the GST effects are largest and where accurate representation of the surface is required. Additionally known geological structures or faults can be dealt with by increasing the number of nodes at appropriate locations such as known boundaries between layers.

[7] This 3-D forward model has been integrated into a Bayesian trans-dimensional inverse approach for inferring GST histories introduced by *Hopcroft et al.* [2007]. This statistical method utilizes reversible jump Markov chain Monte Carlo (rj-MCMC) sampling [*Green*, 1995], so that the number of variables in the model can be varied. Here the number of points used to parameterize the GST history is also a variable, so that the resolution of the GST history over time can be inferred. This also avoids the need to apply ad-hoc regularization, relying on the natural parsimony of the Bayesian approach, whereby simpler models which can adequately fit the data are automatically preferred over more complex models. The Bayesian framework also allows for uncertainty in all model parameters and so for example, the thermal conductivity can be treated as unknown in all or part of the volume (e.g., data are more likely to be available at the borehole from cores). This has the advantage that the specified range of plausible conductivity values will be taken into account in the reconstruction of past GSTs rather than relying on a single set of values.

[8] The rj-MCMC method requires many different samples of the GST history to be generated and the associated forward model evaluated. In order to make this computationally tractable as an inversion approach, a reduced-order model is derived from the full FE simulation using the method of proper orthogonal decomposition (POD) (also known as principal component analysis, Karhunen-Loéve decomposition or empirical orthogonal functions). In this method output from a full-scale FE model is subject to an eigenvalue decomposition. The leading eigenvectors are then used to construct a reduced model [*Sirovich and Kirby*, 1987; *Sirovich*, 1987a, 1987b] which preserves the dominant behavior of the original output. Because the dominant modes are generally contained in just a few of the eigenvectors, the reduced-order model is of much lower dimension compared to the original simulation. The degree to which the first few eigenvalues dominates the complete set of eigenvalues determines the accuracy of the reduced-order model and the errors introduced by reducing the forward model can be monitored (a comparison of full and reduced-order model output is given in section 3).

[9] Detailed descriptions of both the full forward and the reduced forward models are given in sections 2 and 3. Subsequently we give examples of reduced models and compare reduced-order calculations with their full model counterparts. This is followed by an introduction to Bayesian inference and the reversible jump Markov chain Monte Carlo algorithm used to sample the posterior probability distribution of the model parameters. In section 6, synthetic examples are shown to demonstrate how the inverse method performs on realistic data in 3-D settings. Comparisons between GST reconstructions obtained with a standard 1-D method are shown to illustrate the strength of the topographic influence on the GST reconstructions. Further examples demonstrate the effects of incorrectly assuming lateral geological homogeneity. Comparisons between results obtained using the reversible jump Markov chain Monte Carlo inversion and those obtained using a more standard inversion method (nonlinear conjugate gradient)

are also given. This paper concludes with suggestions for future research in this field.

2. Finite Element Forward Model

[10] In this section we develop a reduced-order transient 3-D FE heat conduction model to simulate underground temperatures. Linear basis functions are used in the space domain and a θ time stepping method is used. For $\theta = 1/2$ (as used in this work), the method is the familiar Crank-Nicholson times stepping scheme. The boundary conditions are applied at the upper and lower surfaces, with zero heat flux at the side boundaries. For the upper surface a Robin boundary condition is applied which accounts for heat transfer between the air and the ground. The coupling between the two is determined by a transfer coefficient, β. In all simulations presented, β is kept constant and takes the value of $5.0 Wm^{-2}K^{-1}$. The effect of varying this parameter is not explored here, but see *Stieglitz and Smerdon* [2007] for examples in 1-D settings.

2.1. Formulation

[11] In order estimate the climatic signal, the transient signal must be separated from the steady state thermal regime. This is achieved by solving Fourier's law subject to a heat flux and a long-term equilibrium surface temperature. This latter parameter encompasses climatic fluctuations at the site which occurred prior to the chosen reconstruction length. For a steady state calculation, Fourier's law in 3-D is

$$\nabla \cdot [k_c \nabla T] = 0, \tag{1}$$

where k_c is the thermal conductivity tensor (allowing for anisotropy) and T is temperature. The boundary conditions are then defined by two constants: the surface equilibrium temperature T_{eq} and the basal heat flux q_b,

$$q = \beta(T - T_{eq})|_{\Gamma_\infty} \tag{2}$$

$$-k_c \nabla T = q_0|_{\Gamma_b} \tag{3}$$

where Γ_b is the lower surface, Γ_∞ is the upper surface in contact with the air and q_b is the basal heat flux. In the FE method, the temperature field is approximated by the nodal temperature values:

$$T(x, t^n) = \sum_{j=1}^{N} N_j(x) T_j^n \tag{4}$$

where N_j are a set of piecewise linear basis functions. The FE approximation of equation (1) is

$$\mathbf{K}\,\mathbf{T} = \mathbf{q}, \tag{5}$$

where $\mathbf{T}$ and $\mathbf{q}$ are the vectors which contain the nodal temperatures and heat flux values respectively.

[12] The matrix $\mathbf{K}$ has elements:

$$K_{ij} = \int_\Omega k_c \nabla N_j \cdot \nabla N_i \, dV \tag{6}$$

where the integration is over the volume of each element. The steady state solution of equation (5), $\mathbf{T}_{st}$ is then used as the initial condition for the transient finite element simulation.

2.2. Transient Finite Element Formulation

[13] The diffusion equation is given by

$$\nabla \cdot [k_c \nabla T] = \rho c \frac{\partial T}{\partial t} \tag{7}$$

where ρ and c are the density and specific heat capacity respectively. This equation is solved subject to two boundary conditions:

$$q = \beta(T - T_\infty(t))|_{\Gamma_\infty}, \tag{8}$$

$$-k_c \nabla T = q_0|_{\Gamma_b} \tag{9}$$

and the initial conditions $T = T_{st}$, where $T_\infty(t)$ is the surface air temperature (varying over time). Applying the FE method to equation (7) subject to these two boundary conditions leads to the time stepping FE model [e.g., *Zienkiewicz and Morgan*, 1983]:

$$\frac{\mathbf{M}}{\triangle t}\left[\theta\left(\mathbf{T}^{m+1} - \mathbf{T}^m\right) + (1-\theta)\left(\mathbf{T}^{m+1} - \mathbf{T}^m\right)\right] + \mathbf{K}(1-\theta)\mathbf{T}^m + \mathbf{K}\theta\mathbf{T}^{m+1} = (1-\theta)\mathbf{f}^m + \theta\mathbf{f}^{m+1} \tag{10}$$

where the superscript m refers to the time level, θ ($0 \le \theta \le 1$) controls the combination of forward and backward finite differencing schemes used, $\mathbf{K}$ is given in equation (6) and the entries of $\mathbf{M}$ are given by:

$$M_{ij} = \rho c \int_\Omega N_j N_i dV \tag{11}$$

where ρ and c are assumed constant within Ω. The forcing vector $\mathbf{f}$ is given by:

$$\mathbf{f} = \int_{\Gamma_\infty} \beta\,(T - T_\infty) N_i \, d\Gamma_\infty + \int_{\Gamma_b} N_i \hat{q}_b d\Gamma_b \tag{12}$$

and for which the βT term is subsequently moved over to the left hand side and incorporated into the matrix $\mathbf{K}$. The time stepping equation can be rearranged to give:

$$\mathbf{A}\mathbf{T}^{m+1} + \mathbf{B}\mathbf{T}^m = (1-\theta)\mathbf{f}^m + \theta\mathbf{f}^{m+1} \tag{13}$$

where $\mathbf{A}$ and $\mathbf{B}$ are given by:

$$\mathbf{A} = \left(\frac{\mathbf{M}}{\triangle t} + \theta\mathbf{K}\right) \tag{14}$$

$$\mathbf{B} = \left(-\frac{\mathbf{M}}{\triangle t} + (1-\theta)\mathbf{K}\right) \tag{15}$$

Equations (5) and (13) then constitute the full 3-D FE model.

2.3. Computational Issues

[14] The FE mesh is generated over a horizontal spatial domain covering 1000 m $\times$ 1000 m and a vertical extent that depends on the topography used. For the typical scenario the model depth is 500 m with 90 divisions in the vertical direction and 12 each horizontal direction. The total number of FE nodes produced by the mesh generation software is around 20,000. The time step is 1 year and in the synthetic examples there are 700 steps. The matrix equations in the full FE model are solved by the method of conjugate gradients whereas in the reduced model a partial pivoting Gaussian elimination scheme is used because of the smaller size of the matrices in the reduced-order model.

2.4. Validation of the Forward Model

[15] Solutions from the transient 3-D FE model have been compared with those obtained analytically for a model with a flat surface. The analytical solution for a step increase in surface temperature applied to a semi-infinite half-space is given by *Carslaw and Jaeger* [1959]. For a 1°C step increase applied 150 years before present and with a time step of 1 year and a spatial discretization of 50 m in the x-y plane and 10 m in the z (vertical) direction the maximum difference between the FE and analytical solutions was found to be 3.7×10^{-3} K.

3. Model Reduction Using Proper Orthogonal Decomposition

3.1. Proper Orthogonal Decomposition

[16] Proper orthogonal decomposition (POD) was independently discovered by *Kosambi* [1943], *Loéve* [1945], and *Karhunen* [1946]. It provides a method for describing large data sets using a small number of dominant eigenvectors or principal components. The method has the advantage that it is completely data-dependent. An important innovation in use of POD for large problems in fluid dynamics involves using a series of snapshots, which consist of a set of solutions of the governing equation(s) evaluated at different time instants and determined from the evolution in time of the full model [*Sirovich and Kirby*, 1987; *Sirovich*, 1987a, 1987b]. The snapshots are used to compute the POD basis vectors to yield an optimal representation of the data (here the output solutions of the full model) so that for any given basis vector size, the two norm of the error between the original and reconstructed snapshot is minimized. Output data from a particular computer model are used to derive a reduced description of that simulation. The solutions obtained in this way can be shown to be optimal representations in a least squares sense and the accuracy can be controlled by calculating the energy percentage contained in the reduced representation. A mathematical analysis of POD is given by *Rathinam and Petzold* [2003] and an application of POD model reduction to ocean modeling can be found in the work of *Fang et al.* [2009].

[17] In general the snapshot data $\mathbf{S}_i \; \forall i \in \{1,\ldots m\}$ consisting of solutions to the full model, is an n dimensional vector which has been sampled m times, where n is the size of the state space, for example the number of grid points in a numerical model. The mean vector is found from this data set:

$$\bar{\mathbf{S}}_j = \frac{1}{m}\sum_{i=1}^{m} \mathbf{S}_{i,j}, \;\; \forall j \in \{1,\ldots n\} \tag{16}$$

and $\mathbf{S}$ is then modified to give a zero mean set of numbers. Thus a new ensemble is formed by subtracting this mean from the values in each vector:

$$\Psi_{ij} = \mathbf{S}_{ij} - \bar{\mathbf{S}}_j, \;\; \forall j \in \{1,\ldots n\} \tag{17}$$

An optimal (in a least squares sense) representation Φ_{ij}, of this data set can be found using the following relation:

$$\Phi_j = \sum_{i=1}^{m} a_{ij}\Psi_{ij}, \forall j \in \{1,\ldots n\} \tag{18}$$

where $\mathbf{a}_{ij}$ are the POD coefficients and correspond to the unknown values which are solved for by the forward model. The basis functions, Φ_j must maximize

$$\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m}(\Psi_{ij}\Phi_j)^2, \tag{19}$$

subject to

$$\sum_{j=1}^{n}(\Phi_j)^2 = 1. \tag{20}$$

Singular value decomposition is used to find the optimal basis functions Φ_j, by decomposing the matrix Ψ_{ij},

$$\Psi_{ij} = \mathbf{U}\Delta\mathbf{V}^T, \tag{21}$$

where $\mathbf{U}$ are the singular vectors and the diagonal entries of Δ are the singular values. Singular value decomposition is equivalent to the eigenvalue decomposition in this case. The POD basis functions are then given by the singular vectors corresponding to the k largest singular values of the decomposition:

$$\Phi = \Psi_{ij}\mathbf{U}_i/\sqrt{\lambda_i}, \;\; \forall i \in \{1,\ldots k\} \tag{22}$$

Here k must be large enough that the majority of the information in the original set is captured in the model reduction procedure. The ratio of the sums of the first k singular vectors to the sum of all of the singular values must be kept close to 1, i.e.,

$$\frac{\sum_{i=1}^{k}\Delta_{ii}}{\sum_{i=1}^{n}\Delta_{ii}} \geq \gamma. \tag{23}$$

where the value of γ is usually chosen to be 0.99 or larger and so k an be estimated accordingly. The ratio in equation (23) is referred to as the energy content.

3.2. Model Reduction

[18] Consider the numerical model in section 2, the relation is of the form:

$$\mathbf{A}\mathbf{T}^{m+1} + \mathbf{B}\mathbf{T}^{m} = \mathbf{f}^{m} \tag{24}$$

where $\mathbf{T}^m$ are the nodal temperature values at time step m. This model can be reduced using the POD basis function Φ_j to give

$$\hat{\mathbf{A}}\mathbf{a}' + \hat{\mathbf{B}}\mathbf{a} = \hat{\mathbf{f}} \tag{25}$$

where

$$\hat{\mathbf{A}} = \Phi^T \mathbf{A} \Phi, \tag{26}$$

$$\hat{\mathbf{B}} = \Phi^T \mathbf{B} \Phi, \tag{27}$$

$$\hat{\mathbf{f}} = \Phi^T \mathbf{f}, \tag{28}$$

and $\mathbf{a}$ are the POD coefficients. Applying this to the 3-D FE heat transfer model, $\hat{\mathbf{A}}$, $\hat{\mathbf{B}} \in \mathbb{R}^{k \times k}$ and $\mathbf{a} \in \mathbb{R}^k$. The solution of the reduced-order model, $\mathbf{a}$ can then converted back to the FE nodal temperatures using the relation

$$\mathbf{T} = \Phi \mathbf{a}. \tag{29}$$

[19] Since $k \ll n$ (the number of nodes in the FE grid), the POD model has a much smaller dimension than the FE model. Both equations (24) and (25) are of the same form so that the number of time steps in the reduced model is the same as in the full simulation. In the following section examples of model reduction are given for 1-D and 3-D FE simulations where k is of the order 8. In contrast the number of nodes in the original model n, is typically 10,000–20,000.

3.3. Reduction of the 3-D FE Model

[20] In this example the reduced-order and FE steady state and transient models are compared for 3-D implementations. In both cases, the full-scale 3-D FE model is reduced using POD with the basis vectors found using singular value decomposition method [*Lehoucq et al.*, 1998]. The domain is 1000 × 1000 m with a laterally symmetric surface topography, which ranges from 600 m to 525 m, with a central 200 m wide plateau at 550 m (see Figure 1). The full model has 18,949 nodes and the transient model is run for 330 time steps of 6 years, the reduced model has the same number of time steps, but is constructed using just 8 of the leading singular vectors. The applied boundary condition is a linear increase in surface temperature of 1 K over the model duration of 1980 years.

[21] For the steady state formulation, a series of full FE models are evaluated over a range of the two controlling model parameters, basal heat flux (q_0) and equilibrium surface temperature (T_{eq}). The ranges used are from 55 to 65 mWm^{-2} in 35 increments and 8.5 to 9.5°C in 10 increments, respectively. The steady state snapshot data are therefore generated by 350 steady state FE model forward evaluations. The transient snapshot data are collected at every time step of a single evaluation of the transient FE model with an applied oscillatory surface temperature boundary condition of time period 700 years and amplitude of ±0.25°C.

[22] The difference in °C, between the full and reduced model outputs for the steady state calculations and for values chosen to lie outside of the range used in the snapshot data (50 mWm^{-2} and 10.0°C respectively) is small. The mean and maximum absolute errors for this simulation are 9.90×10^{-10} K and 9.02×10^{-9} K respectively. This magnitude of error is insignificant for geothermal forward modeling. Clearly, the POD reduced-order steady state model is an extremely good approximation of the full FE steady state model.

[23] The POD-FE discrepancy for the transient component is calculated for the case of 330 snapshots. In the transient case the maximum absolute error is found to be 2.05×10^{-3} K while the average absolute error is 9.40×10^{-4} K. This magnitude of the error will vary with the GST history applied to the reduced model. The size of the error is relatively insensitive when the temperatures are kept approximately within around two or three times the range of the GST history used in the training set, but the error values quoted here can be up to double for larger applied GST values. In both cases, the maximum magnitude of the POD-FE error is either smaller than or comparable to the error introduced by the FE discretization (compared to analytical solutions), and so we are confident that the POD model is suitable for our purposes.

[24] If the number of singular vectors in the transient POD model is reduced from 8 to 4, the mean absolute error is slightly larger, 2.82×10^{-3} K. Approximating the complete eigenvalues series by the largest 15 values only, then equation (23) can be used to calculated the energy content of these model reductions. In this case the POD model calculated using the 4 principal eigenvectors has a value of 99.13%, whereas the model with 8 principal eigenvectors leads to a value of 99.98%. These energy values are unchanged if the complete eigenvalue set is approximated by the 20 largest eigenvalues rather than just 15.

[25] From the theory, the choice of the number of snapshots dictates how well the reduced-order model reproduces the behavior of the full scale FE model. The choice of POD rank will only adversely affect any inverse solutions for GST if it is chosen to be too low, in which case the accuracy of the POD reduced forward model will be poor and large systematic errors in the calculated temperature profile will become evident. In the above case the errors from using the POD approximation have been found to be of the same order of magnitude as introduced by the FE discretization itself. Therefore we can be confident that this choice of POD rank is sufficient for the present work.

[26] When generating suitable snapshot data for a reduced model run, the optimal boundary condition to apply will be the same as the boundary condition applied in the reduced model. However, if the boundary conditions in the reduced model vary (as required in the MCMC inverse method), then this is not possible. The results here show that for differing boundary conditions used for the snapshot data

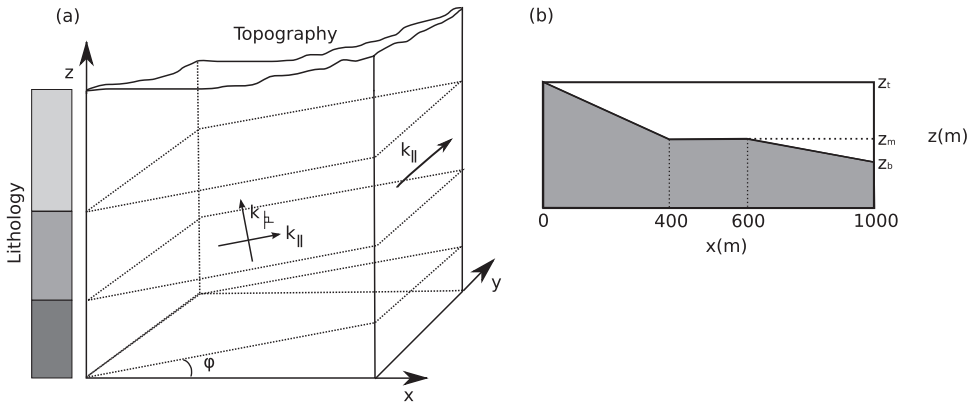

Figure 1. (a) The model setup used in the 3-D simulations. The parameters $k_{\|}$, $k_{\perp}$, c, and ρ are the thermal conductivity (parallel and perpendicular to the formation, respectively), heat capacity, and density, respectively, and these may vary over the volume. The additional parameters are the basal heat flux, q_0, applied to the base of the model, and the surface temperatures (T_{eq} in the steady state and $\mathbf{T}_{\infty}$ in the transient model) applied to the top of the model. (b) The surface topography profile used in the synthetic examples with nonflat upper boundary, where z_t is the height at the highest region of the domain, z_m is the height of the plateau level, and z_b is the height of lowest region.

(oscillatory with time in the transient model) and the reduced-order models (linear with time in the transient test), the accuracy of the reduced-order model is sufficient. This implies that for repeated reduced model evaluations each requiring different boundary conditions (e.g., as required for MCMC inversions) only one full simulation would be needed to generate suitable snapshot data. This can be attributed to the linear nature of the temperature response to heat conduction.

[27] The transient FE forward model takes approximately 4 min to evaluate on a Linux 3.6 GHz Intel Xeon workstation, whereas the POD transient model takes 11 s on the same machine. Evaluating 50,000 forward models would therefore take around 5 months using the full forward model compared to just 6 days using the POD reduced model. In this case, the POD model therefore leads to significant reduction in the required computation time.

4. Inversion, Bayesian Inference, and Markov Chain Monte Carlo

[28] In formulating the inverse problem, the model parameters we need to consider are the temperature history, the equilibrium thermal conditions and the subsurface geological structure (thermal conductivity, specific heat capacity and density). We will treat the geological structure and underground temperatures (down the boreholes) as known to within a given uncertainty level. We then wish to make inference concerning the remaining four quantities. This is a similar setup as in the work of *Hopcroft et al.* [2007] except that heat flow is not now confined to the vertical direction. The model parameter vector is then written as **m**:

$$\mathbf{m} \in \{\mathbf{t}, \mathbf{T}, q_0, T_{eq}\} \tag{30}$$

where **t**, **T** are the time and temperature vectors, given as t_i, T_i, $\forall i \in \{1,\ldots,L\}$, so that there are L nodes in the temperature history, which are used to parameterize the ground surface temperature history with linear interpolation, and $q0$ and T_{eq} are the equilibrium heat flux density and long-term equilibrium surface temperature respectively.

[29] Given these model parameters, the inverse problem is cast in a probabilistic Bayesian framework [e.g., *Mosegaard and Tarantola*, 1995; *Tarantola*, 2005]. This has the advantage that it allows the uncertainty on all model parameters to be assessed. The probability distribution of the model parameter vectors conditioned on the data and prior information is then given by Bayes' law and is termed the posterior [e.g., *Bernardo and Smith*, 1994; *Sivia and Skilling*, 2006]:

$$p(\mathbf{m} \mid \mathbf{d}, \wp) = \frac{p(\mathbf{m} \mid \wp) \times p(\mathbf{d} \mid \mathbf{m}, \wp)}{p(\mathbf{d} \mid \wp)} \tag{31}$$

where p is probability, **m** and **d** are the model and data vectors, and $\wp$ is the theory or hypothesis underlying the model formulation. In words Bayes' law can be expressed as

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} \tag{32}$$

where the terms are in the same order as equation (31), and we return to the specification of these distributions later.

[30] The overall goal of Bayesian inference is to determine the posterior distribution. To do this, we use Markov chain Monte Carlo methods [e.g., *Gilks et al.*, 1996], which allow us to sample the posterior up to a constant of proportionality. This avoids the need to estimate the evidence (or marginal likelihood) term (see equation (32)). In this work we employ reversible jump Markov chain Monte Carlo (rj-MCMC) [*Green*, 1995] as it allows inference on both model parameter values and model dimensionality. Like the more well known Metropolis-Hastings, rj-MCMC constitutes a two stage process of proposing a model probabilistically and then accepting or rejecting this proposed model. The proposal is made by drawing from a

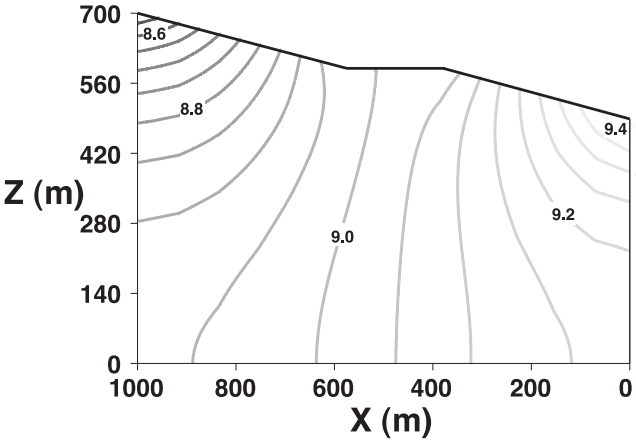

Figure 2. A cross section through the 3-D model with a lapse rate effect of 5 K km^{-1} showing the steady state subsurface isotherms for zero basal heat flow. The same plot with the lapse rate set to zero shows the cross section at a uniform temperature of 9.0°C, the value of T_{eq}.

probability distribution q($\mathbf{m}'|\mathbf{m}$) to give a proposed model m' which is conditional only on the current model **m**. The new model is then accepted with a probability min[1, α($\mathbf{m}'$, **m**)]. If the model is accepted, the current model **m** is replaced by $\mathbf{m}'$ which becomes the current model for the next iteration. If it is not accepted, the current model **m** is retained for another iteration. This process is then iterated many times so that, after a period of initial exploration of the model space (referred to as the burn-in), a series of samples of the model parameters is collected.

[31] In order to ensure convergence of the sampling distribution toward the true distribution, each model proposal needs to be accounted for so that detailed balance is obeyed at each iteration [see *Green*, 1995, 2003]. This is encapsulated by the exact form of the acceptance probability, α, which now includes extra terms to account for possible changes in model dimension.

[32] The 5 proposal types which allow exploration of the model space are as follows:

[33] (1) Create a new GST point (birth). (2) Delete one GST point (death). (3) Perturb one temperature value, T_i. (4) Perturb one time value, t_i. (5) Perturb q_0 and T_{eq}.

[34] At each iteration of the algorithm one proposal type is randomly selected from the above 5. The derivation of the correct values for the acceptance terms for each of these proposal types for this problem is given by *Hopcroft et al.* [2007].

[35] Using the reversible jump form of MCMC allows the construction and comparison of models of differing dimensionality [e.g., *Malinverno*, 2002]. A key feature of Bayesian inference is that it is naturally parsimonious and so simpler models are preferentially sampled in the above methodology, provided that they can adequately explain or fit the data. This means that explicit smoothness constraints on the GST are not required in the prior information. This has the advantage that the amplitude of the reconstructed GST histories can be inferred in a more objective manner, as they should not be artificially smoothed beyond what is implied directly by the data (and the prior).

[36] Assuming the measurement errors on the borehole temperature data are normally distributed and independent between different depths, the likelihood function can be taken as a multivariate Gaussian. That is

$$p(\mathbf{d} \mid \mathbf{m}, \wp, k_p) = \frac{1}{[(2\pi)^{n_d} \det \mathbf{C}_d]^{1/2}} \exp\left[-\frac{1}{2} \cdot (\mathbf{d}_{sim} - \mathbf{d}_{obs})^T \times \mathbf{C}_d^{-1} (\mathbf{d}_{sim} - \mathbf{d}_{obs}) \right], \quad (33)$$

where $\mathbf{d}_{sim}$ and $\mathbf{d}_{obs}$ are the simulated and observed temperatures in the borehole respectively, there are n_d data points and $\mathbf{C}_d$ is the data covariance matrix which is taken as diagonal with entries corresponding to a standard deviation on the data, assuming Gaussian errors, for which a value of 0.1 K has been used for all but the first synthetic examples presented herein.

5. Boundary Conditions

[37] The atmospheric lapse rate, which quantifies the change in air temperature with altitude above sea level, varies from 5 to 10 K km^{-1} depending on the air moisture content [e.g., *Gill*, 1982, Appendix 4]. The standard atmospheric environmental lapse rate value is 6.5 K km^{-1}, so that for a 100 m rise in altitude, the temperature of the air will on average be reduced by 0.65 K. This is significant in terms of climatological investigations because the change in surface temperature over 100 m topographic change is comparable to the strength of atmospheric warming over the twentieth century for example. The effect of a lapse rate is illustrated in Figures 2–4. In the model, the surface topography varies from 500 m to 700 m over a distance of 1000 m with a central 200 m wide plateau at 600 m (Figure 1b). Figure 2 shows the steady state subsurface temperature cross section for the case where by lapse rate is used. Figures 3 and 4 show the transient solutions for the same 3-D volume which

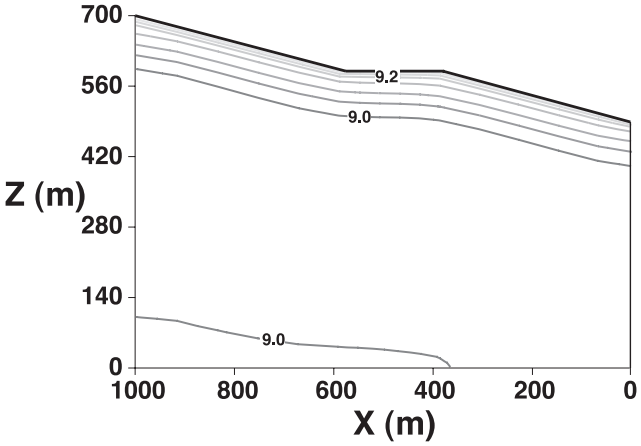

Figure 3. As for Figure 2 but with the lapse rate set to zero and showing the transient response to the last 700 years of the temperature reconstruction from *Moberg et al.* [2005]. Notice that the isotherms are close to parallel to the ground surface.

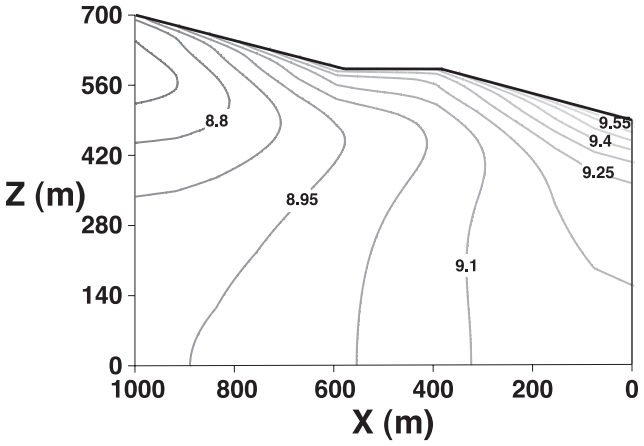

Figure 4. As for Figure 3 but with the lapse rate set to 5 K km^{-1}. Notice now that the isotherms are curved and show induced lateral heat flow.

has now been forced with a ground temperature history based on the last 700 years from *Moberg et al.* [2005]. In the first case (Figure 3) the lapse rate effect is ignored and the isotherms lie close to parallel to the ground surface. In the more realistic example whereby the surface temperature boundary condition is modified by the lapse rate (4), the subsurface isotherms are considerably different and demonstrate lateral heat flow induced by elevation differences. The perturbations due to topography, shown in Figure 5 relative to the 1-D solution, are of a similar magnitude and at similar depths as the perturbations caused by climatic variations [*Lewis and Wang*, 1992]. Therefore a 3-D model is required to interpret underground temperatures located in hilly terrains. In section 6.1 the effect on the GST reconstructions is quantified by comparing inverse results from 1-D and 3-D models.

[38] In reality the orientation of the ground surface relative to the predominant orientation of incoming solar radiation will also have an effect on underground temperatures [*Blackwell et al.*, 1980]. Any surface which is predominantly Sun facing will, on average, absorb a larger amount of solar energy, causing the ground to be warmer than on surfaces which are exposed to more solar radiation. The difference between surface ground temperatures on north and south facing slopes can be up to 3°C, causing heat flow which can be almost horizontal below extreme topography. In the works of *Blackwell et al.* [1980] and *Safanda* [1999] the relationship between total solar irradiation and ground surface temperatures is found to be approximately linear, giving a gradient of the expected average ground surface temperature change with respect to the change in total annual solar insolation of 4×10^{-3} K/(kWhm^{-2} a^{-1}) whereas *Safanda* [1994] and *Gruber et al.* [2004] assume a constant offset in surface temperature between north and south facing slopes for extreme gradients. However, more realistic methods have also been utilized for predicting ground temperatures from meteorological data or regional climate model simulations. For example, *Noetzli et al.* [2007] use output data from a regional climate model to drive an energy balance model coupled with a 3-D FE model in order to simulate underground temperatures below mountainous topographies.

[39] In practice, however, the quantities required by an energy balance model may not be available at suitable locations for a particular borehole. Consequently, it is clear that a simpler method will find wider applicability in practical cases. In this work we follow the approach of *Kohl* [1999] and ignore the effects of irradiation differences. We modify the ground surface temperatures only as a function of altitude using a fixed atmospheric lapse rate. This is a reasonable assumption for sites where topographic variations are small (i.e., nonmountainous). A surface temperature lapse rate of 5 K km^{-1} [*Kohl*, 1999] is used which is close to values calculated by extrapolating borehole temperatures to the surface by *Kubík* [1990] (4.7 K km^{-1}) and *Šafanda* [1999] (4.0 K km^{-1}).

[40] In the majority of borehole climate inversion methods, the subsurface thermal conductivity is assumed to be laterally homogeneous, and in some studies is assumed to be constant with depth. In many real data cases these assumptions are likely to be inappropriate and more realistic models should be used. One issue in geothermics is the possible influence of conductivity anisotropy whereby the thermal conductivity varies as a function of orientation. This has been encountered in real data cases [e.g., *Kukkonen and Šafanda*, 1996; *Clauser et al.*, 1997] and may impair the recovery of past climate changes from borehole data. Here the 3-D FE model is used to assess the magnitude of the effect for typical conductivity variations.

[41] Consider a dipping geological section with anisotropic conductivity so that conductivity parallel to the formation, $k_{||}$ is not equal to the conductivity perpendicular

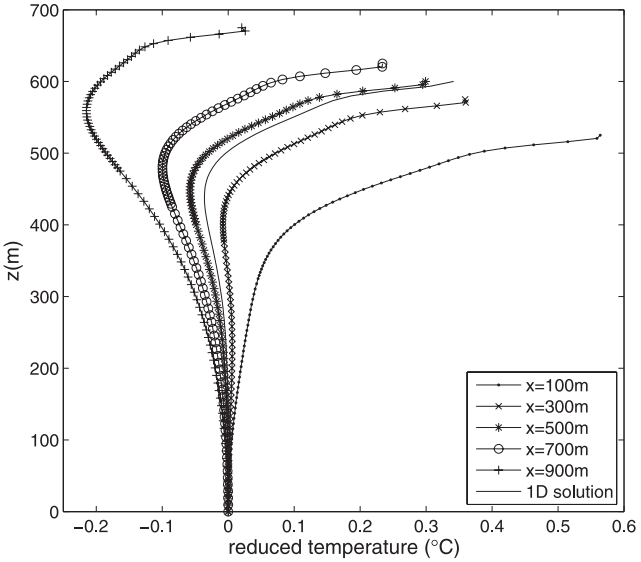

Figure 5. Four reduced temperature depth profiles x = 100, 500, 700, and 900 m in the 3-D volume calculated with 3-D forward model and including the lapse rate effect. The basal heat flow was set to zero. The plots have been shifted to have the same temperature at z = 0 m (the base of the model) for ease of comparison. The applied GST history are the last 700 years of the multiproxy reconstruction of *Moberg et al.* [2005]. For comparison a similar calculation using a 1-D forward model is shown.

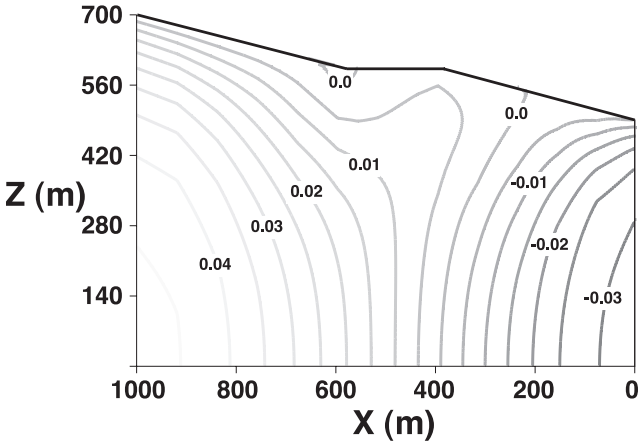

Figure 6. The contours of difference in °C between the anisotropic and isotropic cases for a transient FE simulation. The basal heat flux is set to zero, and the equilibrium surface temperature is 9.0°C. The transient boundary condition is the last 700 years of the proxy reconstruction of *Moberg et al.* [2005].

to the formation $k_\perp$. A typical ratio for these two parameters is 1.5 with the parallel conductivity usually taking the higher value [*Kukkonen and Šafanda*, 1996]. The conductivities for the Cartesian coordinates can be calculated using

$$\mathbf{K}_{cart} = \mathbf{R}^T \begin{pmatrix} k_\parallel & 0 & 0 \\ 0 & k_\perp & 0 \\ 0 & 0 & k_\parallel \end{pmatrix} \mathbf{R} \tag{34}$$

for which **R** is a rotation matrix defining the rigid body rotation of the principal axes of the material to the Cartesian coordinate axes. For the case described here the Cartesian values are given by

$$k_x = \sqrt{k_\perp^2 \sin^2\varphi + k_\parallel^2 \cos^2\varphi}, \tag{35}$$

$$k_y = k_\parallel, \tag{36}$$

$$k_z = \sqrt{k_\perp^2 \cos^2\varphi + k_\parallel^2 \sin^2\varphi}. \tag{37}$$

[42] In order to test the effect of this type of anisotropy on the subsurface temperature distribution, a 3-D transient FE model was constructed with a dipping geological structure. In the first case $k_\parallel$ was set to 2.25 $W(mK)^{-1}$ and $k_\perp$ to 1.5 $W(mK)^{-1}$, i.e., an anisotropy factor ($\frac{k_\parallel}{k_\perp}$) of 1.5. The dip angle was set to $\varphi = 30°$. In the second case, the conductivity is isotropic so that $k_x = k_y = k_z$ with k_z given by equation (37). The difference in °C between the two cases is shown in Figure 6. The maximum absolute difference between the two simulations is 4.7×10^{-2}°C. Reducing the anisotropy factor to 1.1 leads to an approximate halving of this maximum difference. These difference are comparable with the typical error values of borehole temperature measurements. The effect of anisotropic conductivity (where it occurs) can therefore have an affect on paleoclimate inversions but the magnitude is unlikely to be very significant.

6. Synthetic Inversion Examples

[43] For the first inversion example, data published by *Beck et al.* [1992] are used in order to provide a comparison of the inverse technique used here [*Hopcroft et al.*, 2007] with the 5 techniques (which employ 1-D models) described by *Beck et al.* [1992]. In order to achieve this, the model is setup with a cuboid 3-D domain with a horizontal upper surface on an areal domain of dimensions 500 m × 500 m and of depth 600 m. The chosen data set labeled DS2 includes temperature, thermal conductivity, density and specific heat capacity and radiogenic heat production values with depth to 600 m. The data have been degraded with random Gaussian noise by *Beck et al.* [1992]. Here we use the noisy values for temperature (degraded with standard deviation of 0.02 K) and thermal conductivity (degraded with standard deviation of 0.25 $Wm^{-1}\ K^{-1}$) but otherwise use the true values, the noise on the density and specific heat capacity have therefore been ignored. We also correct the temperature values for the radiogenic heat production using the steady state formula $T_r = Az^2/k_c$, for which A is the volumetric heat production ($1 \times 10^{-6}\ Jm^{-3}$ is used by *Beck et al.* [1992]). In the rj-MCMC algorithm the model parameters were initialized with a flat GST history with 2 points at 1000 years and 0 years, and with a basal heat flow of $0.45 Wm^{-2}$ and T_{eq} of 7.0°C. The temperature data uncertainty standard deviation was set to 0.02 K, the same value as the standard deviation of the synthetic noise added to the temperature data.

[44] Figure 7 shows the posterior distribution of the GST after 50,000 iterations of the rj-MCMC algorithm alongside the true model (shifted to have the same T_{eq} as inferred by the algorithm). The results from the rj-MCMC algorithm indicate that the cool period between 1500 AD and 1700 AD cannot be fully resolved, and that the expected value (from the posterior mean) for this time period is

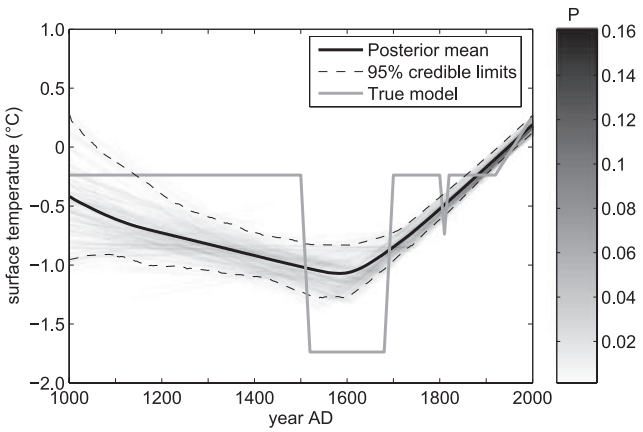

Figure 7. Posterior probability density function on the past GST values as sampled by the rj-MCMC algorithm for the data set DS2 published in the work of *Beck et al.* [1992], and as described in the text.

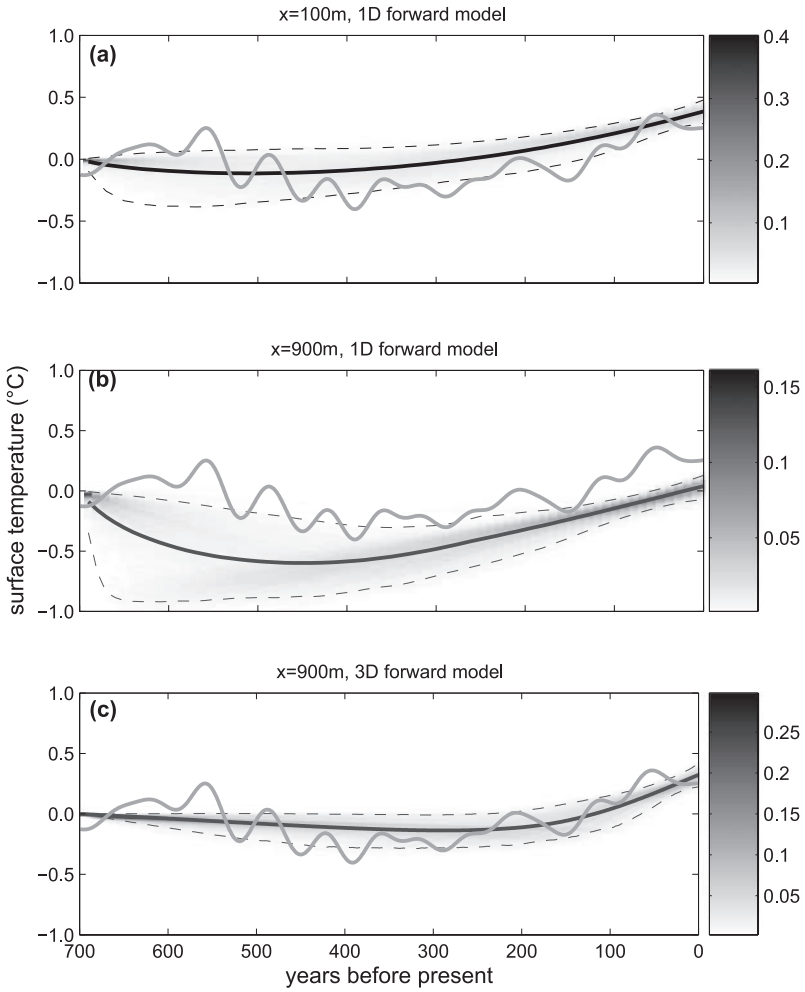

Figure 8. Inversions of the temperature profiles shown in Figure 5. (a) Inversion of x = 100 m profile using a 1-D model, (b) inversion of the x = 900 m profile using a 1-D model, and (c) inversion of the x = 900 m profile using a 3-D forward model. In this latter case the posterior distribution corresponds well with the true model, except for the short excursion at 550 years which is most likely unresolvable due to exponential decay of the surface signal with time and surface perturbation frequency. In Figures 8a and 8b the posterior distributions do not capture the true variability in the synthetic data, and this can be attributed directly to the influence of the surface topography which is not accounted for in the 1-D model. The curves are labeled as in Figure 7.

−1.0°C, just over half the value in the true model (−1.75°C). This underestimation is most likely to be due to the short duration of this temperature excursion in relation to its age (200 years at 300–500 years ago) and the effect of the noise on all of the model parameters. Additionally, the rj-MCMC inversion appears to underestimate the long-term equilibrium temperature (T_{eq}), this may be due to length of the temperature history which means that T_{eq} is unconstrained by the data, since a range of values can be compensated for by a different inferred GST history. More informative prior information would likely counter this effect.

[45] The underestimation of the GST signal in this case demonstrates a useful feature relating to the parsimony of the Bayesian approach whereby the posterior distribution will tend toward the prior where the data provide a poor constraint otherwise. In this case the prior indicates no change in GST over time and so the posterior underestimates the true model somewhat. Direct comparisons of the prior and posterior can be useful in determining the relative influence of the prior and data on the posterior [e.g., *Hopcroft et al.*, 2007]. The 95% credible limits indicate that a cooling to −1.2°C at around 1600 AD should not be ruled out by the data. This shows that the overall posterior distribution contains useful information which may not be conveyed by a single solution.

[46] The equivalent synthetic examples shown in Figure 4 of *Beck et al.* [1992] demonstrate that the amplitude of the cooling at around 1500 AD is unlikely to be well resolved by borehole inversion. Four of the examples in their Figure 4 lead to a similar magnitude of inferred cooling as does the rj-MCMC method here. Additionally, three of the reconstructions show large amplitude variations, of magnitude 1°C, in the recent past at around 1800–1900 AD, which are not consistent with the data. These large-amplitude variations are not inferred by the rj-MCMC method. Addition-

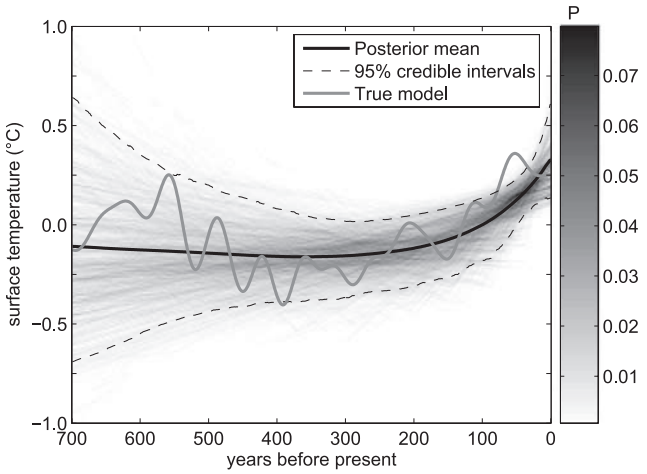

Figure 9. Posterior probability density function on the past GST values as sampled by the rj-MCMC algorithm (50,000 iterations) for noisy synthetic data with 3-D forward model allowing also for the uncertainty on the basal heat flux and equilibrium surface temperature.

ally, because the rj-MCMC is able to deal with the inversion as a nonlinear problem, the inferred credible limits may well be asymmetric, leading to further information about the true GST.

6.1. Idealized Examples 1-D/3-D

[47] In the subsequent synthetic examples a 3-D geometry is used which is designed to be representative of typical moderate topographies as in the work of *Kohl* [1999]. The model has a horizontal extent of 1000 × 1000 m and varies from 700 m to 500 m in the vertical direction. This setup is shown in Figure 1b where z_b = 500 m, z_m = 600 m and z_t = 700 m. In the vertical direction the average nodal spacing is 6.67 m and in the horizontal direction the nodal spacing is 40 m in the central 200 m and 80 m outside of this region. The lapse rate is set to 5 $K\ km^{-1}$. A forward model is run with 700 × 1 year time steps with a surface boundary condition based on the last 700 years of the surface temperature reconstruction by *Moberg et al.* [2005]. The basal heat flux has been set to zero for clarity in interpreting the results. Vertical profiles have been extracted from the resultant temperature field at distances along the slope corresponding to x = 100 m, 500 m, 700 m and 900 m and these are shown in Figure 5. A vertical slice through the resultant temperature field is shown in Figure 4.

[48] The synthetic data produced by the forward calculation at the borehole sites are degraded with 0.1 K normally distributed noise as in the work of *Hopcroft et al.* [2007]. The rj-MCMC algorithm is then run for 20,000 iterations. The resultant posterior samples are shown for a simplified case in Figures 8a and 8b for which the heat flow and equilibrium surface temperature are treated as known. The data were inverted using a 1-D forward model with the rj-MCMC algorithm. The resultant posterior on the past GSTs shows the effects of topography on the recovery of past surface temperatures, as the posterior on the GST diverges during most of the reconstruction period from the true values, with no inferred cool period centered on 400 years (before present) for the x = 100 m and an overestimate of this cooling in the x = 900 m case. For comparison the x = 900 m data are inverted using a 3-D forward model and the same rj-MCMC algorithm and the resulting posterior on the GST is shown in Figure 8c. Apart from the noise on the data this represents a realistic but ideal inversion case against which to make further comparisons. The posterior mean in this example shows excellent agreement with the true model with the posterior mean GST correctly identifying a cool period centered on 400 years. This example clearly demonstrates the strong effect that moderate topographic variations (here 200 m vertical change over a horizontal distance of 1 km) can have on inferred GST histories when a 1-D forward model is employed.

6.2. Uncertain Basal Heat Flow Example

[49] In this more realistic example, a similar model setup is used to test the more realistic inversion case whereby the true values of heat flow (q_0) and long-term equilibrium temperature (T_{eq}) are not known precisely to start with. This results in a larger model space and so more MCMC iterations are required to properly sample the posterior pdf. In addition the data have been truncated above 20 m depth to mimic real data acquisition and since daily and annual signals are significant above this depth. The model was set up with a more moderate surface topography with a variation of 75 m change over the volume, so that z_b = 500 m, z_m = 550 m and z_t = 575 m in Figure 1b. The rj-MCMC algorithm was run for 50,000 iterations. Figure 9 shows the posterior pdf for the GST history. The data favor a model with 3 or 4 points (not shown), which can be taken as a measure of the model complexity (subject to the discrete representation of what is in reality a continuous function of time). In this case the uncertainty of the GST model increases considerably into the past and reflects this more realistic model setup. The posterior mean GST history again indicates a cool period centered on 300–400 years and subsequent 0.4–0.5°C warming to the present. The 95% credible limits enclose the true model at nearly all times except for a short positive temperature excursion at 50 years. Given the level of noise on the temperature data is relatively large at σ_d = 0.1 K and the overall amplitude of the changes in the true model is small and of the order 0.8 K, this example demonstrates the efficacy of the method, and we can be reasonably confident that similar examples with real data will be feasible.

6.3. Three-Dimensional Thermal Conductivity Structure

[50] A common assumption in 1-D GST inversions is that the underlying geological structure is laterally homogeneous. This assumption is tested here using a realistic synthetic example. The forward model is set up with three sloping geological layers of differing thermal conductivities. In the inversion, the conductivity layers are assumed to be horizontal and take the values which occur at the borehole position in the forward model.

[51] The synthetic data have been recalculated to account for the three parallel layers which are assumed to have a dip angle of 8.53° corresponding to a vertical rise of 150 m over 1000 m horizontally. The topographic profile of the preced-

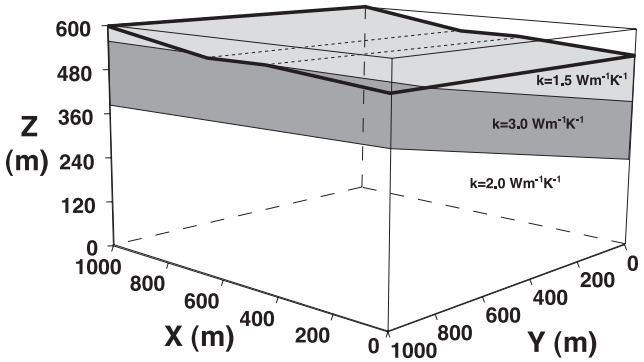

Figure 10. The conductivity profile setup used in the forward model example of section 6.3.

ing example is used (i.e., a change of 75 m is applied at the upper surface from one side of the volume to the other). The thermal conductivity layers from surface downward are isotropic and with values of 1.5, 3.0 and 2.0 $W(mK)^{-1}$. This thermal conductivity setup is depicted in Figure 10. For comparison, the forward model is forced with the same surface temperature reconstruction as in previous examples and is setup with a basal heat flux of 60 mWm^{-2} and an equilibrium surface temperature of 9.0°C. Again the surface temperature is modified according to the ground surface lapse rate of 5 K km^{-1}.

[52] In Figure 11 the posterior pdf of the GST history is depicted after rj-MCMC sampling of 50,000 iterations. It can be seen that the posterior mean GST accurately follows the past temperature variations shown in the true model. Although a general cool period is inferred, the amplitude and timing is modest (around −0.1 K rather than −0.3 K) and the time of the coolest temperatures has been shifted toward the present by around 100 years when compared to Figure 9. The algorithm also shows a tendency toward smoother models (not shown) as the posterior probability of 2 or 3 time-temperature points is higher than for the example in Figure 9. Overall this case demonstrates that, although slightly incorrect thermal conductivity values have been used over much of the volume, the reconstructed GST history is relatively robust to this change. The effect of the lateral thermal conductivity homogeneity assumption (which is not true in this example) is relatively small, other geometries may lead to increased discrepancies in the inferred GST histories.

7. Comparing Reversible Jump-MCMC With a Nonlinear Conjugate Gradient Inversion Method

[53] In order to provide a comparison with the results obtained using rj-MCMC a more standard inverse method is considered (nonlinear conjugate gradient NLCG). For rectangular matrices that the linear conjugate gradient method is equivalent to Lanczos' algorithm [e.g., *Golub and Van Loan*, 1996, chapter 10.2.5], a method used to calculate the singular value decomposition (SVD) of a matrix. Therefore, the NLCG method applied here is a nonlinear analogue of the SVD method widely used for borehole data inversion [e.g., *Mareschal and Beltrami*, 1992], and comparison of this method with the rj-MCMC used here is therefore particularly relevant.

[54] NLCG relies on a first-order gradient calculation and uses a functional term penalizing the roughness of the GST model to constrain the solutions to realistic values. The method is designed such that initial model updates produce generally smooth models, while subsequent iterations can introduce more detailed structure. The gradient calculations are achieved efficiently by using the adjoint method [e.g., *Wang et al.*, 1992]. This means that these calculations are exact for the finite element discretization used here.

[55] The functional quantifies the progress of the inversion method by measuring the fit between the measured and simulated data as in the likelihood used in the Bayesian method. In this work the functional also includes a regularization term which penalizes deviations in the temperature history model from the smoothest (i.e., no change) model. This is formed by the square of the differential of the temperature history with time, for which the continuous and discretized equations are given by:

$$F_r = \frac{1}{2}\lambda \int \nabla^T T \mathbf{k}_{sm} \nabla T \, d\Omega \tag{38}$$

$$F_r = \frac{1}{2}\lambda \mathbf{T}^T \mathbf{K}_\mathbf{r} \mathbf{T} \tag{39}$$

where $\mathbf{k}_{sm}$ is a regularization matrix, $\mathbf{K}_{\mathbf{r}i,j} = \int \nabla^T Q_i \nabla Q_j d\Omega$, where Q_i are the temporal basis functions. The integral is over the time domain of the reconstruction. The data fit contribution to the functional is given by:

$$F_d = \frac{1}{2}(\mathbf{d} - \mathbf{d}_{obs})^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{d}_{obs}) \tag{40}$$

where C_d is the data covariance matrix.

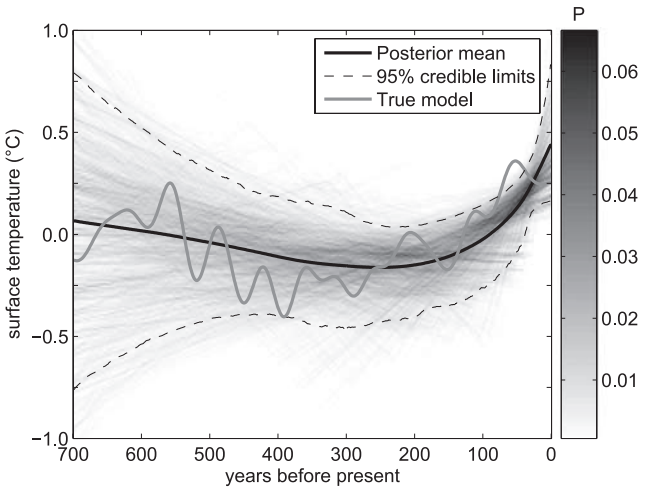

Figure 11. Posterior probability density function of the past GSTs as sampled by the rj-MCMC algorithm (50,000 iterations) for the noisy synthetic case corresponding to the conductivity structure of Figure 10 but with assumed lateral homogeneity so that values at the borehole are assumed for the whole volume.

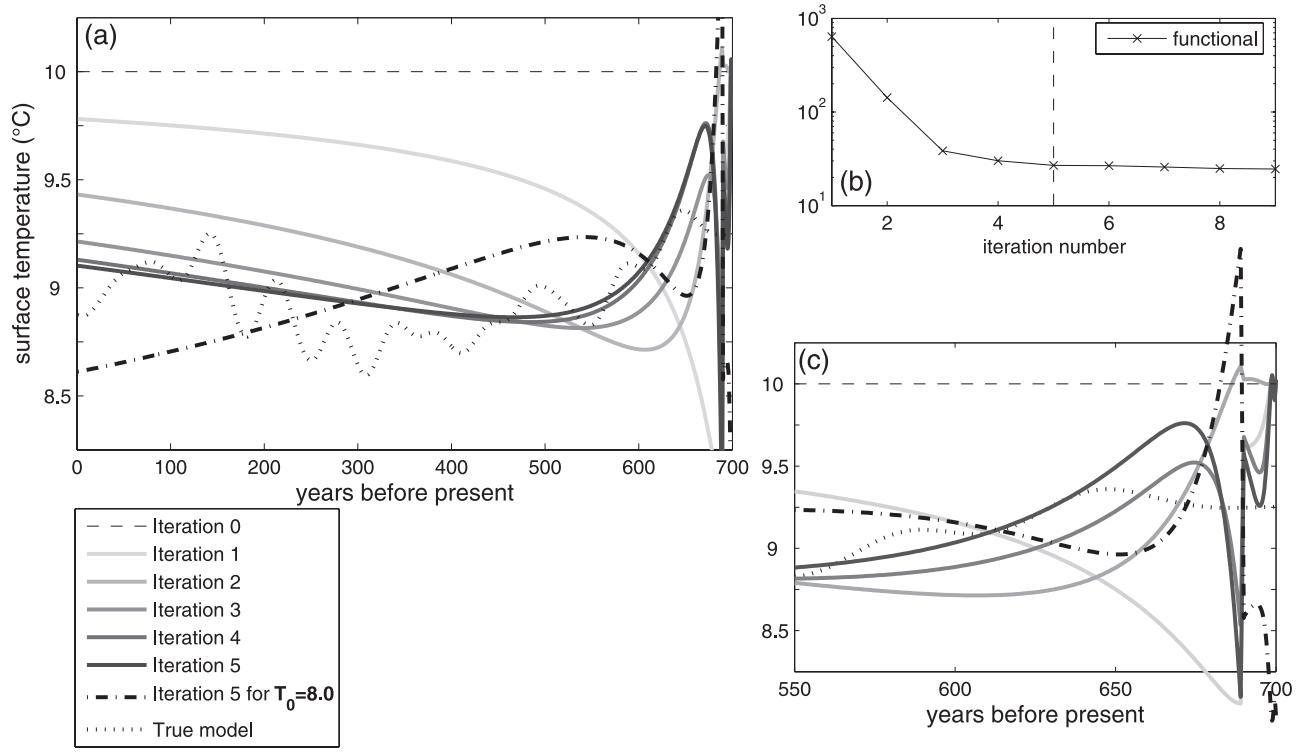

Figure 12. (a) The inverse GST models with iterations for two starting models using the NLCG method. The first five iterations are shown for the model starting at $\mathbf{T}_0 = 10°C$, and the fifth iteration only is shown for the model initialized with $\mathbf{T}_0 = 8°C$. (b) The functional value for the NLCG algorithm starting at $\mathbf{T}_0 = 10°C$. The vertical line denotes the stopping point which is chosen according to the magnitude of the model update as described in the text. (c) As in Figure 12a but showing the last 150 years only.

[56] The overall functional used for nonlinear conjugate gradient F, is given by a combination of the data fit and regularization terms:

$$F = F_d + F_r = \frac{1}{2}(\mathbf{d} - \mathbf{d}_{obs})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{d}_{obs}) + \frac{1}{2}\lambda \mathbf{m}^T \mathbf{K}\mathbf{m}. \quad (41)$$

[57] The NLCG algorithm uses Polak-Ribière formula to calculate the search directions and makes use of a parabolic bracketing method in order to find the minimum along each search direction. If the model regularization is not sufficient the algorithm will tend to inverse solutions with large-amplitude variations which are not apparent in the true model. In order to combat this the regularization term is initially set to be large and is then gradually relaxed by 5% per iteration. Additionally, a termination criterion is introduced so that the algorithm is stopped once the functional reduction due to the model update falls below a certain criterion. Here this is determined to be the functional update which corresponds to less than 0.01 K average update at each data point, i.e., when the functional changes by less than 0.4.

[58] To compare the NLCG with the rj-MCMC method, the data from the idealized 3-D example were used, corresponding to the posterior GST history shown in Figure 8 and the results are given in Figure 12. In this example, the NLCG inversion was stopped after 5 iterations. The solution shows reasonable agreement with the true model, depicting a general cool period in the past with more recent warming. However, the model appears to be more sensitive to recent events and so the minimum of the cool period is inferred to have occurred at around 200 years rather than at 400 years.

[59] In Figure 12 we can see that the NLCG method infers a set of rapid variations between 20 years and the present. These variations appear to be an artifact of the gradient calculation and the lack of constraints for very recent GST history as the data are taken from depths greater than 20 m. These variations are therefore only penalized in the functional through the regularization term and are therefore difficult to remove with increased regularization without obscuring part of the true trend at earlier times [e.g., *Chouinard and Mareschal*, 2007]. A plausible solution is to cut off the inverse model at 10–20 years before present, or to average the GST values over this period, here the former option used. Other more sophisticated approaches to this problem involve varying the parameterization of the forward model or using a time-dependent regularization term [*Şerban and Jacobsen*, 2001; *Jacobsen and Rath*, 2007].

[60] Since gradient optimization methods cannot guarantee convergence to a particular functional minimum, it is necessary to use multiple trials which are started from different positions in the model space. Therefore for comparison the inversion of the data used in Figure 12 inverted again with the model initiated as $\mathbf{T} = 8.0°C$. The resultant inverse solutions converge to the same warming signal for the recent part of the history, whereas the earlier part of the

history before 300 years is biased toward lower temperatures. In this example the algorithm is again terminated after 5 iterations and the final model has almost exactly the same functional value. However, here the inverse solution GST history does not show the same variation over time, instead showing a general warming over the whole reconstruction length, this is shown in Figure 12 by the black dash-dot line. It therefore appears that in this example the algorithm has converged to a different functional minimum leading to a different solution.

[61] Realistic synthetic examples presented here indicate that the inverse solutions which are dependent on the gradient of the functional w.r.t. model parameters are most sensitive to the recent climatic events. This means that it is difficult for the method to recover realistic timing of past temperature changes. Evidence from repeated inversions with differing starting models shows that approximately equal data fits can be obtained for solutions with quite different GST histories. This is due to the degeneracy of the inverse solutions caused by the noise in the data. This is a feature with is dealt with robustly in a Bayesian formulation through the quantification of the posterior distribution of the GST history. In our experience this has shown no significant dependence on the starting values used in the rj-MCMC sampling method.

[62] However, the rj-MCMC method used is relatively computationally expensive, typically requiring 10,000 s of forward model evaluations. In contrast the gradient based method requires only around 50–100 equivalent forward model evaluations (the gradient calculation is equivalent to 2 forward runs). For extremely large data sets or 3-D models, it may only be feasible to use an optimization type method unless the forward model is parallelized efficiently, but this is prone to converging on poor solutions. Some hybrid scheme involving both gradient and an rj-MCMC methods could therefore be useful. For example, the faster gradient inversions could be used to approximate the prior pdf of the temperature histories by collecting the results from multiple starting points. This more informative prior could then serve to constrain the plausible model space (and hence computation time) for the rj-MCMC algorithm.

8. Discussion and Conclusions

[63] In this work a new method for performing simulation and inversion for GST histories using a 3-D numerical forward model has been described. A major limitation has been that full 3-D models are typically expensive or slow and the aim here has been to develop a forward model strategy that is efficient enough that a fully nonlinear Bayesian inversion method (MCMC) could be applied. This necessarily implies the use of a further approximation in the finite element method. In order to achieve this, the method of proper orthogonal decomposition has been applied to a conventional 3-D finite element heat transfer model such that a reduced-order model is produced. However, the errors introduced by this approximation have been shown to be small enough for accurate inversions to be considered. This reduced-order model is then taken as the forward model and is linked with a Bayesian approach to the inverse problem in which the trans-dimensional form of Markov chain Monte Carlo sampling is used.

[64] We have demonstrated the application of the new method using synthetic data, which have been degraded with noise in order to replicate realistic data sources. The resultant GST histories then show similar resolution to those obtained in 1-D settings. For comparison, the same data have been treated with an identical inverse method but utilizing instead a 1-D forward model. In examples of moderate topography (200 m over 1 km) the resultant GST histories are shown to be significantly affected as the 1-D inversion results do not recover the true GST trends. The impact of assuming lateral homogeneity of the geological structure has also been assessed. This shows that for large differences between the true and assumed thermal conductivity models, the inferred GSTs are still relatively realistic, but (in the case considered here) leads to slight underestimation of the true trends.

[65] This approach therefore allows the inclusion of a range of processes in the forward model which could not previously be accounted for when inverting borehole data for past surface temperatures. In particular the effects of variable surface topography and heterogeneous subsurface conductivity values can be taken into account. Previously *Kohl* [1999] showed that the effects of topography could be eliminated from temperature-depth profiles by performing a correction, which relies on 3-D calculations. This has the disadvantage that the basal heat flow and equilibrium surface temperature as well as the underground conductivities must be calculated separately from the inversion for the GST histories. In our approach these quantities are inferred jointly from the data along with the GSTs allowing for uncertainty in each parameter and therefore leading to a more robust approach, as the inferred GST signal depends strongly on the values chosen for each of these other quantities.

[66] The trans-dimensional sampling method (rj-MCMC) used here allows for models of differing dimensionality such that the resolution of the GST history with time is addressed directly during the inversion procedure. Comparisons between this method and a more conventional nonlinear conjugate gradient method have shown that the rj-MCMC method is simpler in terms of less tunable parameters and is generally able to reconstruct the applied GST histories more closely.

[67] The synthetic case studies shown in this paper illustrate effects that could be important for particular settings. However, these sorts of terrains are purposefully avoided when selecting data for paleoclimate studies. The effects of topographical settings on previous borehole paleoclimate reconstructions cannot be estimated from the results in this work, since the calculations here relate to a single synthetic model setup. In practice the effects of topography may be expected to cancel upon averaging of multiple reconstructions, since boreholes could be located with equal proportion on positive and negative topographic features. However, this is difficult to justify without resorting to detailed analysis of suitable digital elevation data.

[68] Regional averages of GST histories are particularly important in real data settings in order that robust climate histories can be derived. The method used here could be employed at locations where it is warranted (i.e., where topography is significant), while the 1-D equivalent could be employed for other sites with regular topography.

Regional averages could then be found by averaging the posterior mean GST deviations for these sites. Alternatively a simple statistical model (e.g., Gibbs sampling) could be used to sample jointly from the separate posterior GST history probability distributions for each of the sites considered. For closely spaced boreholes a single 3-D forward model could be used to model all of the sites together. In that case it may be found that the simple parameterization of the surface temperature as a function of altitude (using a lapse rate) is inadequate. In this case the Bayesian methodology used here could be extended to estimate spatial variations of this parameter using for example Bayesian partition modeling [*Denison et al.*, 2002]. This may then allow the influence of historical land cover differences [e.g., *Nitoiu and Beltrami*, 2005; *Bense and Beltrami*, 2007] from one borehole to another to become apparent, since these effects will disrupt the assumed linear relationship between ground surface temperature and ground surface altitude.

[69] A chief limitation for the methods introduced here is the paucity of appropriate geological data required to construct the 3-D forward model. Much of the borehole data currently available for paleoclimate reconstruction are not accompanied by appropriate subsurface geological data. However, topographic data at reasonable horizontal resolution of 90 m is available for the globe, with higher 30 m resolution data available for the U.S. and this could be used to reasonably accurately model the topographic influence at a particular borehole site. Further limitations stem from the simplicity of the model setup considered here, which ignores the effects of vegetation or snow cover which can vary with topography and the influence of topographically induced fluid flow. Nevertheless, a recent study by *Chouinard and Mareschal* [2007] discarded 34 out of 194 borehole temperature profiles for climate studies on the grounds of topography alone. Clearly some of these may have been affected additionally by topographically induced fluid flow, but this result gives an approximate upper limit for the proportion of borehole data that can now be used for climate reconstruction purposes using the methods presented herein.

[70] In future work, data from a digital elevation model (DEM) could be incorporated into the 3-D forward model in order to test the method on real data. Furthermore, the forward model could also be modified to simulate underground advection of ground water. This advection-diffusion model could then also be reduced using the POD method and linked with a suitable inversion scheme. Because of the trade-off between the influence of groundwater flow and paleoclimate on underground temperatures, some method for assessing a priori the fluid velocities would be required. This could be obtained from detailed geological investigations of the site under consideration or by the use of a hydrogeological fluid flow model.

[71] **Acknowledgments.** We thank I. M. Navon for helpful advice on POD and model reduction. We are also grateful to two reviewers and the associate editor for many insightful comments and suggestions. P.O.H. was supported through a studentship grant from the Environmental Mathematics and Statistics program of NERC and EPSRC to K.G.

References

Beck, A. E., P. Y. Shen, H. Beltrami, J. C. Mareschal, J. Šafanda, M. N. Sebagenzi, G. Vasseur, and K. Wang (1992), A comparison of five different analyses in the interpretation of five borehole temperature data sets, *Global Planet. Change*, *6*(2–4), 101–112.

Beltrami, H. (2002), Climate from borehole data: Energy fluxes and temperatures since 1500, *Geophys. Res. Lett.*, *29*(23), 2111, doi:10.1029/2002GL015702.

Bense, V., and H. Beltrami (2007), Impact of horizontal groundwater flow and localized deforestation on the development of shallow temperature anomalies, *J. Geophys. Res.*, *112*, F04015, doi:10.1029/2006JF000703.

Bernardo, A. F. M., and J. M. Smith (1994), *Bayesian Theory*, John Wiley, Chichester, U. K.

Blackwell, D. D., J. L. Steele, and C. A. Brott (1980), The terrain effect on terrestrial heat flow, *J. Geophys. Res.*, *85*(B9), 4757–4772.

Carslaw, H. S., and J. C. Jaeger (1959), *Conduction of Heat in Solids*, 2nd ed., Oxford Univ. Press, Oxford, U. K.

Chouinard, C., and J.-C. Mareschal (2007), Selection of borehole temperature depth profiles for regional climate reconstructions, *Clim. Past*, *3*, 297–313.

Clauser, C., P. Giese, E. Huenges, T. Kohl, H. Lehmann, L. Rybach, J. Šafanda, H. Wilhelm, K. Windloff, and G. Zoth (1997), The thermal regime of the crystalline continental crust: Implications from the KTB, *J. Geophys. Res.*, *102*(B8), 18,417–18,441.

Denison, D. G. T., C. C. Holmes, B. K. Mallick, and A. F. M. Smith (2002), *Bayesian Methods for Nonlinear Classification and Regression*, John Wiley, Chichester, U. K.

Fang, F., C. Pain, I. Navon, M. Piggott, G. Gorman, P. Allison, and A. Goddard (2009), Reduced order modelling of an adaptive mesh ocean model, *Int. J. Numer. Methods Fluids*, *59*(8), 827–851, doi:10.1002/fld.1841.

Gilks, W. R., , S. Richardson, and D. J. Spiegelhalter (Eds.) (1996), *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.

Gill, A. E. (1982), *Atmosphere-Ocean Dynamics*, *Int. Geophys. Ser.*, vol. 30, Academic, San Diego, Calif.

Golub, G., and C. Van Loan (1996), *Matrix Computations*, 3rd ed., Johns Hopkins Univ. Press, Baltimore, Md.

Green, P. J. (1995), Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, *82*(4), 711–732.

Green, P. J. (2003), Trans-dimensional Markov chain Monte Carlo, in *Highly Structured Stochastic Systems*, edited by P. J. Green, N. L. Hjort, and S. Richardson, pp. 179–198, Oxford Univ. Press, Oxford, U. K.

Gruber, S., L. King, T. Kohl, T. Herz, W. Haeberli, and M. Hoelzle (2004), Interpretation of geothermal profiles perturbed by topography: The Alpine permafrost boreholes at Stockhorn Plateau, Switzerland, *Permafrost Periglac. Process.*, *15*(4), 349–357, doi:10.1002/ppp.503.

Harris, R. N., and D. S. Chapman (2001), Mid-Latitude (30 degrees–60 degrees N) climatic warming inferred by combining borehole temperatures with surface air temperatures, *Geophys. Res. Lett.*, *28*(5), 747–750.

Hopcroft, P. O., K. Gallagher, and C. C. Pain (2007), Inference of past climate from borehole temperature data using Bayesian Reversible Jump Markov chain Monte Carlo, *Geophys. J. Int.*, *171*(3), 1430–1439, doi:10.1111/j.1365-246X.2007.03596.x.

Huang, S. P., H. N. Pollack, and P. Y. Shen (2000), Temperature trends over the past five centuries reconstructed from borehole temperatures, *Nature*, *403*(6771), 756–758, doi:10.1038/35001556.

Jacobsen, B. H., and V. Rath (2007), Tuning a multiscale prior with generalized cross validation for piecewise constant paleotemperature, *Geophys. Res. Abstracts*, *9*, SRef-ID:1607-7692/gra/EGU2007-A-05557.

Karhunen, K. (1946), Zur spektraltheorie stochastischer prozesse, *Ann. Acad. Sci.*, *37*.

Kohl, T. (1999), Transient thermal effects below complex topographies, *Tectonophysics*, *306*(3–4), 311–324.

Kohl, T., and S. Gruber (2003), Evidence of paleotemperature signals in mountain permafrost areas, in *Proceedings of the 8th International Conference on Permafrost*, edited by L. U. Arenson, M. Phillips, and S. M. Springman, Taylor and Francis, Philadelphia, Pa.

Kohl, T., S. Signorelli, and L. Rybach (2001), Three-dimensional (3-D) thermal investigation below high Alpine topography, *Phys. Earth Planet. Inter.*, *126*(3–4), 195–210, doi:10.1016/S0031-9201(01)00255-2.

Kosambi, D. (1943), Statistics in function space, *J. Indian Math. Soc.*, *7*, 76–88.

Kubík, J. (1990), Subsurface temperature field of the Bohemian Massif, *Stud. Geophys. Geod.*, *34*, 110–128.

Kukkonen, I., and J. Šafanda (1996), Palaeoclimate and structure: The most important factors controlling suburface temperatures in cyrstalline rocks. A case study from Outokumpu, eastern Finland, *Geophys. J. Int.*, *126*, 101–112.

Lachenbruch, A. (1969), The effect of two-dimensional topography on superficial thermal gradients, *U.S. Geol. Surv. Bull.*, *1203-E*, 1–86.

Lachenbruch, A. H., and B. V. Marshall (1986), Changing climate—Geothermal evidence from permafrost in the Alaskan Arctic, *Science*, *234*(4777), 689–696.

Lehoucq, R., D. Sorensen, and C. Yang (1998), *ARPACK Users Guide: Solution of Large-Scale Eigenvalue Problems With Implicitly Restarted Arnoldi Methods*, Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Lewis, T. J., and K. Wang (1992), Influence of terrain on bedrock temperatures, *Global Planet. Change*, *6*(2–4), 87–100.

Loéve, M. (1945), Functions aleatoire de second ordre, *Compte Rend. Acad.*, *220*.

Malinverno, A. (2002), Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, *Geophys. J. Int.*, *151*(3), 675–688, doi:10.1046/j.1365-246X.2002.01847.x.

Mareschal, J. C., and H. Beltrami (1992), Evidence for recent warming from perturbed geothermal gradients: Examples from eastern Canada, *Clim. Dyn.*, *6*, 135–143.

Moberg, A., D. M. Sonechkin, K. Holmgren, N. M. Datsenko, and W. Karlén (2005), Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data, *Nature*, *433*(7026), 613–617, doi:10.1038/nature03265.

Mosegaard, K., and A. Tarantola (1995), Monte-Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*(B7), 12,431–12,447.

Nitoiu, D., and H. Beltrami (2005), Subsurface thermal effects of land use changes, *J. Geophys. Res.*, *110*, F01005, doi:10.1029/2004JF000151.

Noetzli, J., S. Gruber, T. Kohl, N. Salzmann, and W. Haeberli (2007), Three-dimensional distribution and evolution of permafrost temperatures in idealized high-mountain topography, *J. Geophys. Res.*, *112*, F02S13, doi:10.1029/2006JF000545.

Pollack, H. N., and S. P. Huang (2000), Climate reconstruction from subsurface temperatures, *Ann. Rev. Earth Planet. Sci.*, *28*, 339–365.

Pollack, H. N., and J. E. Smerdon (2004), Borehole climate reconstructions: Spatial structure and hemispheric averages, *J. Geophys. Res.*, *109*, D11106, doi:10.1029/2003JD004163.

Rathinam, M., and L. R. Petzold (2003), A new look at proper orthogonal decomposition, *SIAM J. Numer. Anal.*, *41*(5), 1893–1925, doi:10.1137/S0036142901389049.

Šafanda, J. (1994), Effects of topography and climatic changes on the temperature in borehole GFU-1, Prague, *Tectonophysics*, *239*, 187–197.

Šafanda, J. (1999), Ground surface temperature as a function of slope angle and slope orientation and its effect on the subsurface temperature field, *Tectonophysics*, *306*, 367–375, doi:10.1016/S0040-1951(99)00066-9.

Şerban, D., S. Nielsen, and C. Demetrescu (2001), Transylvanian heat flow in the presence of topography, paleoclimate and groundwater flow, *Tectonophysics*, *335*, 331–344, doi:10.1016/S0040-1951(01)00065-8.

Şerban, D.Z., and B. H. Jacobsen (2001), The use of broad-band prior covariance for inverse palaeoclimate estimation, *Geophys. J. Int.*, *147*(1), 29–40, doi:10.1046/j.0956-540x.2001.01509.x.

Shen, P. Y., and A. E. Beck (1992), Paleoclimate change and heat flow density inferred from temperature data in the Superior Province of the Canadian Shield, *Global Planet. Change*, *6*(2–4), 143–165.

Shen, P. Y., H. N. Pollack, S. Huang, and K. Wang (1995), Effects of subsurface heterogeneity on the inference of climate-change from borehole temperature data—Model studies and field examples from Canada, *J. Geophys. Res.*, *100*(B4), 6383–6396.

Sirovich, L. (1987a), Turbulence and the dynamics of coherent structures, part ii: Symmetries and transformations, *Q. Appl. Math.*, *45*, 573–582.

Sirovich, L. (1987b), Turbulence and the dynamics of coherent structures part iii, Dynamics and scaling, *Q. Appl. Math.*, *45*, 583–590.

Sirovich, L., and M. Kirby (1987), Low-dimensional procedure for characterization of human faces, *J. Opt. Soc. Am.*, *4*(3), 519–524.

Sivia, D. S., and J. Skilling (2006), *Data Analysis: A Bayesian Tutorial*, 2nd ed., Oxford Univ. Press, Oxford, U. K.

Stieglitz, M., and J. E. Smerdon (2007), Characterizing land-atmosphere coupling and the implications for subsurface thermodynamics, *J. Clim.*, *20*(1), 21–37, doi:10.1175/JCLI3982.1.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Wstimation*, Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Wang, Z., I. Navon, F. Le Dimet, and X. Zou (1992), The 2nd-order adjoint analysis—Theory and applications, *Meteorol. Atmos. Phys.*, *50*(1–3), 3–20.

Zienkiewicz, O., and K. Morgan (1983), *Finite Elements and Approximations*, Wiley, New York.

F. Fang and C. C. Pain, Applied Modelling and Computation Group, Department of Earth Science and Engineering, Imperial College, London SW7 2AZ, UK. (f.fang@imperial.ac.uk; c.pain@imperial.ac.uk)

K. Gallagher, Géosciences Rennes, Université de Rennes 1, Campus de Beaulieu, F-35042 Rennes, France. (kerry.gallagher@univ-rennes1.fr)

P. O. Hopcroft, Bristol Research Initiative for the Dynamic Global Environment, School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK. (peter.hopcroft@bris.ac.uk)