The Cryosphere Discuss., 8, 5511–5537, 2014
www.the-cryosphere-discuss.net/8/5511/2014/
doi:10.5194/tcd-8-5511-2014

The Cryosphere Discussions
Open Access

Model calibration for ice sheets and glaciers dynamics: a general theory of inverse problems in glaciology

M. Giudici[1,2,3], F. Baratelli[1,4], A. Comunian[1], C. Vassena[1], and L. Cattaneo[1,2]

[1]Università degli Studi di Milano, Dipartimento di Scienze della Terra "A. Desio", via Cicognara 7, 20129 Milano, Italy
[2]CNR-IDPA (Consiglio Nazionale delle Ricerche, Istituto per la Dinamica dei Processi Ambientali), via Mario Bianco 9, 20131 Milano, Italy
[3]CINFAI (Consorzio Interuniversitario Nazionale per la Fisica delle Atmosfere e delle Idrosfere), Piazza Niccolò Mauruzi 17, 62029 Tolentino (MC), Italy
[4] MINES ParisTech, Centre de Géosciences, 35 Rue Saint-Honoré, 77305 Fontainebleau, France

Received: 18 September 2014 – Accepted: 2 October 2014 – Published: 28 October 2014

Correspondence to: M. Giudici (mauro.giudici@unimi.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Numerical modelling of the dynamic evolution of ice sheets and glaciers requires the solution of discrete equations which are based on physical principles (e.g. conservation of mass, linear momentum and energy) and phenomenological constitutive laws (e.g. Glen's and Fourier's laws). These equations must be accompanied by information on the forcing term and by initial and boundary conditions (IBCs) on ice velocity, stress and temperature; on the other hand the constitutive laws involve many physical parameters, some of which depend on the ice thermodynamical state. The proper forecast of the dynamics of ice sheets and glaciers requires a precise knowledge of several quantities which appear in the IBCs, in the forcing terms and in the phenomenological laws. As these quantities cannot be easily measured at the study scale in the field, they are often obtained through model calibration by solving an inverse problem (IP). The objective of this paper is to provide a thorough and rigorous conceptual framework for IPs in cryospheric studies and in particular: to clarify the role of experimental and monitoring data to determine the calibration targets and the values of the parameters that can be considered to be fixed; to define and characterise identifiability, a property related to the solution to the forward problem; to study well-posedness in a correct way, without confusing instability with ill-conditioning or with the properties of the method applied to compute a solution; to cast sensitivity analysis in a general framework and to differentiate between the computation of local sensitivity indicators with a one-at-a-time approach and first-order sensitivity indicators that consider the whole possible variability of the model parameters. The conceptual framework and the relevant properties are illustrated by means of a simple numerical example of isothermal ice flow, based on the shallow-ice approximation.

1 Introduction

The physics of the dynamic evolution of ice sheets and glaciers is based on physical principles (e.g. conservation of mass, linear momentum and energy) and phenomenological constitutive laws (e.g. Glen's and Fourier's laws) which are used to derive partial differential equations. These equations must be accompanied by information on the forcing terms and by initial and boundary conditions (IBCs) on ice velocity, stress and temperature. On the other hand, the constitutive laws involve many physical parameters, some of which depend on the ice thermodynamical state. The basic equations of ice sheet models can be found, e.g. in Hutter (1983), van der Veen (1999), Hooke (2005), Greve and Blatter (2009) and Cuffey and Paterson (2010).

Difficulties associated with heterogeneity and anisotropy of the physical parameters and of the climatic forcing, non-linearity of the physical processes, complex geometries, etc. do not permit to compute analytical solutions, which can be obtained only if strong approximations are introduced. Therefore, numerical methods of solution of partial differential equations (e.g. finite differences, finite elements, etc.) are used and discrete numerical models are developed and applied, such as, e.g. SICOPOLIS (Greve, 1995), GLIMMER (Rutt et al., 2009), PISM (the PISM authors, 2014; Winkelmann et al., 2011), and many others, some of which were also tested in intercomparison experiments (Huybrechts et al., 1996; Payne et al., 2000; Pattyn et al., 2008).

The proper forecast of the dynamics of ice sheets and glaciers (forward problem, FP) does not depend only on the goodness of the approximations introduced by the discretization of the domain and of the equations. It requires also a precise knowledge of several quantities which appear in the IBCs, in the forcing terms and in the phenomenological laws. Unfortunately, field measurements are often affected by large uncertainties and poor space and time sampling, whereas laboratory measurements are relevant to scales which are very different from those involved in the FP. Therefore, one must rely on model calibration to infer the input model parameters, i.e. the solution of inverse problems (IPs) is necessary.

Roughly speaking, an IP aims at finding the optimal values of the model parameters that yield the best agreement of the model output with the field observations and data. Several applications of inverse modelling were proposed in glaciology (MacAyeal, 1992, 1993; Arthern and Hindmarsh, 2003; Joughin et al., 2004; Truffer, 2004; Gudmundsson and Raymond, 2008; Raymond and Gudmundsson, 2009; Avdonin et al., 2009; Morlighem et al., 2010; Arthern and Gudmundsson, 2010; Gillet-Chaulet et al., 2012; Habermann et al., 2012; Petra et al., 2012; Pollard and DeConto, 2012; Bonan et al., 2014, and others). However, IP theory, which is well developed in several areas of science and geophysics (see, e.g. Parker, 1994; Tarantola, 2004; Menke, 2012), has not yet become popular in glaciological sciences (see Gudmundsson, 2014, for a review). In fact, very excellent methods of IP solution have been tested in several examples (see the papers cited above, among many others), but some basic mathematical properties are not fully considered in the applications and therefore a somehow formal and abstract review can be useful.

IPs are often claimed to be ill-posed. However, this is rigorously true only for continuous domain models. For discrete numerical models, the properties of the IP must be analysed with more care and when this is done, it appears that difficulties sometimes arise from ill-conditioning or non-uniqueness (Giudici, 2002). Moreover, it is necessary to clarify the role of experimental and monitoring data to determine the calibration targets and the values of the parameters that can be considered to be fixed, whereas only the model output should depend on the subset of the parameters that can be identified with the calibration procedure and the solution to the IP. It is actually difficult to guarantee the existence and uniqueness of a solution to the IP for complex non-linear models. Also identifiability (Giudici, 1989, 1991), a property related to the solution to the FP, and resolution should be carefully considered. Moreover, instability of the IP should not be confused with ill-conditioning and with the properties of the method applied to compute a solution. Finally, sensitivity analysis is of paramount importance to assess the reliability of the estimated parameters and of the model output. It is often based on the one-at-a-time approach, through the application of the adjoint-state method (see, e.g.

Heimbach and Bugnion, 2009; Petra et al., 2012; Heimbach and Losch, 2012), to compute local sensitivity, i.e. the uncertainty on the model output due to small variations of the input parameters. However, first-order approaches that consider the whole possible variability of the model parameters should be considered (see, e.g. Hill and Tiedeman, 2006; Saltelli et al., 2008; Baratelli et al., 2012).

Therefore, the objective of this paper is to provide a further step towards a thorough and rigorous theoretical conceptual framework for IPs in cryospheric studies, in order to improve the definition and the comprehension of the properties of IPs with a formal approach which might help to close the gap between mathematical abstraction and applied simulation modelling. The conceptual framework and the relevant properties of IPs are illustrated by means of a simple numerical model of isothermal ice flow, based on the shallow-ice approximation (SIA, see, e.g. Hutter, 1983; Baratelli et al., 2011).

2 The paradigmatic example and definition of the forward problem

The conceptual framework is introduced with a paradigmatic example based on the application of the model developed by Bueler (2014) at the University of Alaska (Fairbanks AK, USA) and implemented in the Matlab code `siageneral.m`. The model is based on a finite-difference discretization of the ice-sheet equation under the SIA and the hypothesis of constant and uniform temperature, that is (Bueler et al., 2005):

$$\partial_t H = M + \nabla \cdot (D \nabla h), \tag{1}$$

where H is the ice thickness, t is time, h is the ice surface elevation, M is the surface mass balance and D is the diffusivity. The model does not consider ice-shelves, so that $h = b + H$, where b is the bed elevation. D is given by

$$D = 2EA(\rho g)^n \frac{H^{n+2}}{n+2} |\nabla h|^{n-1}, \tag{2}$$

where A and n are the flow parameters and E the enhancement factor in Glen's law, which is the constitutive relation assumed between strain rate $\dot{\varepsilon}$ and deviatoric stress τ:

$$\dot{\varepsilon}_{ij} = EA\tau_e^{n-1}\tau_{ij}, \tag{3}$$

where τ_e is the effective deviatoric stress.

The model is applied to the Antarctic ice-sheet: the bed elevation and the initial ice thickness are taken from the dataset ALBMAP v1 (Le Brocq et al., 2010) and it is assumed that $n = 3$ and $A = 10^{-16}\,\mathrm{Pa}^{-3}\,\mathrm{a}^{-1}$. The latter value corresponds to the reference value used for some experiments of model intercomparison: EISMINT I (Huybrechts et al., 1996) and ISMIP-HOM (Pattyn et al., 2008). In the paradigmatic example, it is assumed that M is constant in space and time. Of course these are very strong approximations, but they are useful to have a test to introduce some concepts and to show some results in a very simple way. A reference solution $h^{(\mathrm{ref})}$ is generated by running the model from the present-day geometry for a time period of 20 ka with the reference parameters $M^{(\mathrm{ref})} = 0.3\,\mathrm{m\,a}^{-1}$ and $E^{(\mathrm{ref})} = 3$.

In real-world applications, acquired data should include information about the geometry of the domain, the positions of the measurement points, the measured values of physical quantities (e.g. ice-sheet-surface height, temperature at the surface and in few boreholes, ice-sheet velocity at the surface, ice accumulation rate at some monitoring stations on the surface, etc.). All these data are collected in an array $\boldsymbol{d}$. For the paradigmatic example, $\boldsymbol{d}$ contains the synthetic data corresponding to the nodal values of $h^{(\mathrm{ref})}$.

The input model parameters, included those that describe the geometry of the discretization grid (e.g. the spacing of the grid), are contained in the array $\boldsymbol{p}$. Some of these parameters are fixed before the application of the model and can be grouped in a "sub-array" $\boldsymbol{p}^{(\mathrm{fix})}$, which depends on the data, i.e. $\boldsymbol{p}^{(\mathrm{fix})}(\boldsymbol{d})$. Instead, the model parameters whose values are obtained from calibration, via the solution of an IP, are collected in the array $\boldsymbol{p}^{(\mathrm{cal})}$. Therefore, $\boldsymbol{p} = (\boldsymbol{p}^{(\mathrm{fix})^t}, \boldsymbol{p}^{(\mathrm{cal})^t})^t$. For the paradigmatic example

$\boldsymbol{p}^{(\mathrm{cal})} = (e, M)$, whereas $\boldsymbol{p}^{(\mathrm{fix})}$ includes, among the others, the length of the simulated time period (20 ka) and the prescribed initial condition.

The state of the system is the ice-sheet surface under stationary conditions at the end of the simulation: the values computed at all the nodes of the discretization grid are collected in an array $\boldsymbol{s}$. With this notation, the model can be written as

$$\mathbf{A}(\boldsymbol{p}, \boldsymbol{s})\boldsymbol{s} = \boldsymbol{b}(\boldsymbol{p}, \boldsymbol{s}), \tag{4}$$

where the matrix **A** is used to discretize not only convective and diffusive flow terms, but also capacity terms related to time-variations, whereas the array $\boldsymbol{b}$ is related to accumulation, energy production and to boundary conditions. Therefore, Eq. (4) is a prototype for the linear system of equations arising from the discretization of the partial differential equations which translate mass, momentum and energy conservation principles in mathematical form, even for transient conditions, for which the array $\boldsymbol{s}$ is usually split in the sub-arrays corresponding to different time steps. Also, Eq. (4) is a prototype of different methods of solution of the partial differential equations: for instance, for finite elements or spectral methods, the array $\boldsymbol{s}$ could include the coefficients of the basis functions. Both **A** and $\boldsymbol{b}$ depend on the system state because the relevant equation are usually non linear. **A** is usually a sparse matrix and for approaches based on the discretization of integral balance equations, it is also symmetric and definite positive.

Roughly speaking, the FP aims at solving Eq. (4) with respect to $\boldsymbol{s}$, given the model parameters $\boldsymbol{p}$, whereas the IP aims at identifying the values of some of the model parameters, $\boldsymbol{p}^{(\mathrm{cal})}$, given data ($\boldsymbol{d}$) that can be used to estimate $\boldsymbol{p}^{(\mathrm{fix})}$ and to predict the state $\boldsymbol{s}$ of the system. The solution of the FP can be expressed in explicit form as

$$\boldsymbol{s} = \boldsymbol{g}(\boldsymbol{p}), \tag{5}$$

which is the forward mapping $\boldsymbol{p} \rightarrow \boldsymbol{s}$.

3 Definition of the inverse problem

The model outcome, i.e. the state of the system, can be used to forecast other quantities that depend also on the model parameters and possibly on some of the data. Therefore, the model forecast are expressed as an array $\boldsymbol{y}$, which is function of $\boldsymbol{s}$, $\boldsymbol{p}$ and $\boldsymbol{d}$: $\boldsymbol{y}(\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{p})$.

For instance, with reference to the paradigmatic example, the basic option is that the model forecast coincides with the ice cap surface, i.e. $\boldsymbol{y} \subset \boldsymbol{s}$. An alternative option is that the model forecast used for calibration is the ice cap volume. In this case, it is required to include other parameters in $\boldsymbol{p}^{(\mathrm{fix})}$ to compute the model forecast, e.g. parameters related to the geometry of the system, namely the bedrock topography.

Roughly speaking, the IP consists in the determination of the optimal values of $\boldsymbol{p}^{(\mathrm{cal})}$, i.e. those values that reduce the misfit between the model forecasts and some target values $\boldsymbol{t}$. In the simplest case of the basic option, when measurements of the system state are available, some elements of $\boldsymbol{s}$ can be directly compared with the corresponding elements of $\boldsymbol{d}$, i.e. $\boldsymbol{t} \subset \boldsymbol{d}$. Instead, for the aforementioned alternative option, the calibration target is obtained from the processing of field data, in order to obtain an estimate of the total ice cap volume and therefore it requires additional data and some processing parameters. Then it is necessary to express the calibration targets as an array depending on $\boldsymbol{d}$ and $\boldsymbol{p}^{(\mathrm{fix})}$: $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{d}, \boldsymbol{p}^{(\mathrm{fix})})$.

The IP is therefore related to the determination of $\boldsymbol{p}^{(\mathrm{cal})}$ through the inverse mapping $\{\boldsymbol{d}, \boldsymbol{p}^{(\mathrm{fix})}\} \rightarrow \boldsymbol{p}^{(\mathrm{cal})}$.

The most common approach to IP is the search for the minimum of an objective function $\mathcal{O}$, given by

$$\mathcal{O}(\boldsymbol{p}^{(\mathrm{cal})}) = \|\boldsymbol{y}(\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{p}) - \boldsymbol{t}(\boldsymbol{d}, \boldsymbol{p}^{(\mathrm{fix})})\|. \tag{6}$$

The classical choice is the least-squares approach, when the norm appearing in the right hand side of Eq. (6) is the sum of squared differences between $\boldsymbol{t}$ and $\boldsymbol{y}$ components (l_2 norm). This is the objective function applied in the paradigmatic example

of this paper. Of course, many other choices are possible, among which the sum of absolute differences (l_1 norm) and the maximum absolute difference (l_∞ norm).

The above described formalism includes also more complex objective functions, such as that based on the the logarithm of the misfit between modelled and observed glacier-surface velocity proposed by Morlighem et al. (2010) or that developed by Arthern and Gudmundsson (2010) who introduced the Dirichlet-to-Neumann map approach (Calderòn, 2006; Kohn and Vogelius, 1984) in glaciological modelling. These objective functions were tested to model the Greenland ice sheet dynamics by Gillet-Chaulet et al. (2012).

The misfit between $\boldsymbol{y}$ and $\boldsymbol{t}$ depends on several factors: measurement errors; relevance of the measurement support volumes with respect to the spatial and temporal scales of the model; model approximations; data processing; etc. This motivated several researchers (Berliner et al., 2008; Raymond and Gudmundsson, 2009, among the others) to consider the data and the model parameters, and therefore also the model forecasts and the calibration targets as stochastic processes. Then the Bayes' theorem can be invoked; with the formalism that has been previously introduced, it can be cast as:

$$(\boldsymbol{p}^{(\mathrm{cal})}|\boldsymbol{y}-\boldsymbol{t}) = \frac{f(\boldsymbol{y}-\boldsymbol{t}|\boldsymbol{p}^{(\mathrm{cal})})\cdot f(\boldsymbol{p}^{(\mathrm{cal})})}{f(\boldsymbol{y}-\boldsymbol{t})}, \tag{7}$$

where f functions are (possibly conditioned) probability density functions (pdfs) of the respective arguments. In particular, $f(\boldsymbol{p}^{(\mathrm{cal})})$ is the prior pdf of the model parameters to be calibrated, i.e. independent from the measurements of state variables or other independent quantities; instead, $f(\boldsymbol{p}^{(\mathrm{cal})}|\boldsymbol{y}-\boldsymbol{t})$ represents the posterior pdf, which is conditioned on the measured data.

Notice that Eq. (7) is slightly different from the standard formulation proposed by other authors (see, e.g. the textbook by Menke, 2012), because the framework introduced in this paper is more general, as it accounts for different types of model outputs and calibration targets.

Most of the applications of the Bayesian approach compute the optimal parameters by means of the maximum likelihood method (Edwards, 1972), which searches for the array $\boldsymbol{p}^{(\mathrm{cal})}$ which maximizes the posterior probability given by Eq. (7). Among the most commonly introduced approximations it is worth recalling the assumption that both $f(\boldsymbol{y}-\boldsymbol{t}|\boldsymbol{p}^{(\mathrm{cal})})$ and $f(\boldsymbol{p}^{(\mathrm{cal})})$ can be expressed as multigaussian distributions. Moreover, it is often implicitly assumed that $f(\boldsymbol{y}-\boldsymbol{t})$ is independent from $\boldsymbol{p}^{(\mathrm{cal})}$. Under these hypotheses, the maximum-likelihood method reduces to the least-squares approach (Menke, 2012).

At the authors' knowledge, no test has been conducted in glaciological sciences with different hypotheses of pdfs. However, an exponential pdf, above all for $f(\boldsymbol{y}-\boldsymbol{t}|\boldsymbol{p}^{(\mathrm{cal})})$, might be a better guess in presence of outliers, i.e. of model predictions which are very far from the expected target values. In practice, exponential pdfs yield the minimization of l_1-like norms, which are expected to provide robust estimates, i.e. to be able to yield reasonable results even in presence of high errors.

4 Properties of the IP

The material is now ready to introduce and discuss some properties of the IP.

4.1 Ill-posedness and ill-conditioning

The first question is whether model parameters are identifiable, i.e. if different values of the parameters always yield different predictions of the state of the system with the FP. In other words, the model parameters are said to be identifiable, if for every couple of arrays $\boldsymbol{p}$ and $\boldsymbol{p}'$, $\boldsymbol{p} \neq \boldsymbol{p}'$, the corresponding solutions to Eq. (4), $\boldsymbol{s} = \boldsymbol{g}(\boldsymbol{p})$ and $\boldsymbol{s}' = \boldsymbol{g}(\boldsymbol{p}')$ are such that $\boldsymbol{s} \neq \boldsymbol{s}'$.

Any mathematical problem that is applied to model physical processes is required to be well-posed, i.e. it is required that a solution exists, is unique and is stable with respect to the data. For IPs, in principle, it is very easy to state that the uniqueness of

the solution corresponds to the property that a unique array $\boldsymbol{p}^{(\mathrm{cal})}$ yields an absolute minimum of Eq. (6). In principle, it is also very simple to state that an IP admits a unique solution if $\mathcal{O}$ is a convex function of its arguments; unfortunately, it is not easy to prove this for complex models, with a great number of parameters.

The paradigmatic example is so simple that it is possible to draw the graph of $\mathcal{O}$, which is expressed for this example as the root-mean-square error (RMSe) between the modelled ice-sheet surface and $h^{(\mathrm{ref})}$: it is shown in Fig. 1 for the basic option of $\boldsymbol{y}$ and $\boldsymbol{t}$. This plot shows that the IP has a unique solution, obviously corresponding to the reference values used to generate the synthetic data $h^{(\mathrm{ref})}$. This is confirmed by Fig. 2: the objective function is convex in a neighbourhood of the reference parameters and no local minima are present.

Notice that if one of the model parameters were not identifiable, it would be possible to find another couple of values that yield the same solution to the FP. Since in this example $\boldsymbol{y}$ does not explicitly depend on $\boldsymbol{p}$, as it is often the case, then the solution to the IP would not be unique. This remark is fundamental to show the strict link between identifiability, which is a property of the FP, and uniqueness of the IP.

IPs are usually claimed to be unstable and this is true for the continuous case, when a continuous domain is considered and the model is built with partial differential equations (Giudici and Vassena, 2008). Instead, for discrete models, algebraic equations have to be considered and this means that the issue of stability could be of minor relevance. Moreover, the definition of stability presumes that the error on the data can be reduced at will, so that the results of the IP can be as close as possible to the reference value or to the solution for ideal, error-free data. However, this is not the case with real-world applications, when measurement, modelling and approximation errors cannot be reduced below a practical limit.

Figure 3 shows the graph of $\mathcal{O}$ when an uncorrelated error with a gaussian distribution with zero average and SD of 1 m is added to the data, i.e. to the reference elevation $h^{(\mathrm{ref})}$. It is difficult to appreciate differences with the graph of Fig. 1. The differences are

even smaller for smaller values of the SD. In other words, this is a graphical proof that the considered IP is stable.

However, it has been stressed that the concept of stability assumes that the error on the data can be reduced at will, but this will never be the case in practical applications, due to the great number of causes of errors that have been recalled before. In several applications some typical characteristics that denote numerical instability (oscillatory behaviour, alternating high and low values) can be observed: however, they might be effects of ill-conditioning or non-identifiability or non-uniqueness, rather than due to instability of the IP.

Of course conditioning is a qualitative concept, which is related to the Lipschitz condition, defined as follows. If $\boldsymbol{p}^{\dagger}$ and $\boldsymbol{p}^{\ddagger}$ are the solutions to the IP corresponding, respectively, to data $\boldsymbol{d}^{\dagger}$ and $\boldsymbol{d}^{\ddagger}$, then a Lipschitz condition holds if

$$\|\boldsymbol{p}^{\dagger} - \boldsymbol{p}^{\ddagger}\| \leq C\|\boldsymbol{d}^{\dagger} - \boldsymbol{d}^{\ddagger}\|, \tag{8}$$

where $C > 0$ is called a Lipschitz constant. If C is big, then the IP is ill-conditioned, because a very small (and often practically unrealistic) error on the input data is necessary to guarantee a small, physically acceptable error on the calibrated parameters. In other words, ill-conditioning means that a small error on the data could be sufficient to yield big differences in the estimated parameters; on the other hand well-conditioned IPs are such that even if the errors on the input data are quite big, the solution to the IP does not change dramatically.

It is also very important to stress that a clear distinction must be done among problems with the intrinsic properties of the IP and the effects of the solution method. Figure 3 shows a very simple example. On top of the graph of $\mathcal{O}$ the minimization paths followed by the application of the Matlab function `fminunc`, based on the quasi-Newton algorithm, with two different initializations are shown. Information listed in Table 1 show that starting from different couples of parameters values might yield different "solutions". In fact, Fig. 2 shows that the reference values lie in a "valley" characterised by a weak slope, which makes difficult the identification of the absolute minimum with

a gradient-based approach. However, notice that the difference between the values of the $\mathcal{O}$ functions at those points are very small and even about one third of the SD of the error added to the data.

Problems with non-uniqueness, instability or ill-conditioning might appear as alternating (high and low) values of the elements of $\boldsymbol{p}^{(\mathrm{cal})}$. They can be handled by introducing regularization terms in the objective function (see, e.g. Habermann et al., 2012; Petra et al., 2012), which is often obtained by including a term like $\|\boldsymbol{p}^{(\mathrm{cal})}\|$ in $\mathcal{O}$. This regularizing term has the effect of cutting the high values and therefore the minima of $\mathcal{O}$, which are characterised by the most accentuated oscillatory behaviour. With the formalism proposed in this paper, this is equivalent to directly insert the values of $\boldsymbol{p}^{(\mathrm{cal})}$ into the model outputs $\boldsymbol{y}$ and by assuming that the corresponding target elements are set to zero. For the Bayesian approach, the regularizing effect is introduced through the prior pdf.

4.2 Sensitivity analysis

4.2.1 Some definitions

Several authors addressed the problem of sensitivity of model outcomes with respect to input parameters (Ritz et al., 2001; Heimbach and Bugnion, 2009; Baratelli et al., 2012; Giudici et al., 2012; Larour et al., 2012; Thoma et al., 2012; Heimbach and Losch, 2012; Schäfer et al., 2012). This is fundamental in order to estimate the errors on model predictions due to the uncertainties on the values of $\boldsymbol{p}$, but also to assess the physical relevance of some parameters and of some physical processes to determine, for instance, glacier thickness, ice velocity, ice temperature, etc.

A simple approach is the computation of quantities related to the ratio between variations of $\boldsymbol{s}$ or $\boldsymbol{y}$ as a response to variations of $\boldsymbol{p}$ around a reference value. From the mathematical point of view, this is nothing but a derivative. In particular, the state sensitivity,

$\mathbf{S}^{(s)}_{mn}$, of a state variable $\boldsymbol{s}_m$ with respect to a single parameter $\boldsymbol{p}_n$, under a linear approximation, i.e. for small local variations of the parameter, is given by:

$$\mathbf{S}^{(s)}_{mn} = \frac{\partial \boldsymbol{s}_m}{\partial \boldsymbol{p}_n} = \frac{\partial \boldsymbol{g}_m}{\partial \boldsymbol{p}_n}(\boldsymbol{p}). \tag{9}$$

Analogously, the prediction sensitivity, $\mathbf{S}^{(y)}_{mn}$, of a model prediction $\boldsymbol{y}_m$ with respect to $\boldsymbol{p}_n$ is given by:

$$\mathbf{S}^{(y)}_{mn} = \frac{d\boldsymbol{y}_m}{d\boldsymbol{p}_n} = \sum_{k=1}^{N_\mathrm{S}} \frac{\partial \boldsymbol{y}_m}{\partial \boldsymbol{s}_k} \cdot \frac{\partial \boldsymbol{s}_k}{\partial \boldsymbol{p}_n} + \frac{\partial \boldsymbol{y}_m}{\partial \boldsymbol{p}_n} = \sum_{k=1}^{N_\mathrm{S}} \frac{\partial \boldsymbol{y}_m}{\partial \boldsymbol{s}_k} \cdot \mathbf{S}^{(s)}_{kn} + \frac{\partial \boldsymbol{y}_m}{\partial \boldsymbol{p}_n}, \tag{10}$$

where N_S is the dimension of the array $\boldsymbol{s}$ and the "total" dependence of $\boldsymbol{y}$ on $\boldsymbol{p}$ is considered explicitly, i.e. both through the direct functional dependence and the indirect dependence through the solution of the FP.

However, these definitions have two weaknesses. First, since parameters and system states are represented by physical quantities, with given measurement units, it is impossible to identify the most sensitive parameters from a straightforward comparison among the elements of $\mathbf{S}^{(s)}$ and $\mathbf{S}^{(y)}$. Therefore, it is necessary to scale or normalize these quantities. Second, both $\mathbf{S}^{(s)}$ and $\mathbf{S}^{(y)}$ are based on a linearized, one-at-a-time approach, so that they take into account only the linear approximation of the model and neglect both non-linear effects and joint effects of the parameters.

The first problem can be overcome by means of the dimensionless normalized sensitivity, which corresponds to the scaling given by the SD of the relevant quantities, or the prediction scaled sensitivity, which is scaled by means of the reference values around which the sensitivity indices are computed (Giudici et al., 2012).

The second issue is overcome by means of the first-order sensitivity (Saltelli et al., 2008), which can be defined as

$$S_n = \frac{\mathrm{var}_{p_n}[E_{p\setminus n}[Y|\boldsymbol{p}_n]]}{\sigma_Y^2}, \tag{11}$$

where Y represents a state variable $\boldsymbol{s}_m$ or a model prediction $\boldsymbol{y}_m$, $E_{p\setminus n}[Y|\boldsymbol{p}_n]$ is the expected value of Y conditioned on the parameter $\boldsymbol{p}_n$ and var_{p_n} is the variance with respect to $\boldsymbol{p}_n$.

4.2.2 Adjoint method for the computation of sensitivity

The computation of $\mathbf{S}^{(s)}$ is often a crucial computational aspect for the application of IPs, because it is required both to compute $\mathbf{S}^{(y)}$ and the gradient of $\mathcal{O}$ for methods of solution which are based on steepest-descent or conjugate-gradient approaches.

The simplest approach is the computation of $\mathbf{S}^{(s)}_{mn}$ with a finite-difference approach: the FP is solved for two different arrays $\boldsymbol{p}^+$ and $\boldsymbol{p}^-$, which differ from each other only for the value of $\boldsymbol{p}_n$ by an amount Δp: if the corresponding solutions to the IP are denoted by, respectively, $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$, then $\mathbf{S}^{(s)}_{mn} \simeq (\boldsymbol{s}^+_m - \boldsymbol{s}^-_m)/\Delta p$.

An alternative is the use of the adjoint-state equation method (Plessix, 2006), which is introduced in the continuous case by making use of variational calculus and by introducing the Frechet's derivative. Here it is shortly revised for its application to discrete models in glaciological sciences (Heimbach and Bugnion, 2009; Heimbach and Losch, 2012; Goldberg and Heimbach, 2013; Martin and Monnier, 2014).

A linearization of Eq. (4) is obtained by imposing that $\mathbf{A} \simeq \mathbf{A}(\boldsymbol{p}, \tilde{\boldsymbol{s}})$ and $\boldsymbol{b} = \boldsymbol{b}(\boldsymbol{p}, \tilde{\boldsymbol{s}})$, where $\tilde{\boldsymbol{s}}$ is fixed as the solution to Eq. (4) corresponding to the parameters $\boldsymbol{p}$ around which the sensitivity is estimated; in other words Eq. (4) is linearized around the values of the reference parameters and system states. If Eq. (4) is multiplied by an arbitrary

array $\boldsymbol{\psi}^{(m)}$, and the derivative of the resulting equation with respect to $\boldsymbol{p}_n$ is taken, one obtains

$$\frac{\partial \mathbf{A}}{\partial \boldsymbol{p}_n}\boldsymbol{s}\cdot\boldsymbol{\psi}^{(m)} + \mathbf{A}\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{p}_n}\cdot\boldsymbol{\psi}^{(m)} - \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{p}_n}\cdot\boldsymbol{\psi}^{(m)} = 0. \tag{12}$$

If $\boldsymbol{\psi}^{(m)}$ is the solution of the so-called "adjoint-state equation"

$$\mathbf{A}^t\boldsymbol{\psi}^{(m)} = \boldsymbol{\delta}_m, \tag{13}$$

where $\boldsymbol{\delta}_m$ is the unit impulse concentrated on the mth element, then

$$\begin{aligned}\frac{\partial \boldsymbol{s}_m}{\partial \boldsymbol{p}_n} &= \frac{\partial \boldsymbol{s}_m}{\partial \boldsymbol{p}_n} - \frac{\partial \mathbf{A}}{\partial \boldsymbol{p}_n}\boldsymbol{s}\cdot\boldsymbol{\psi}^{(m)} - \mathbf{A}\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{p}_n}\cdot\boldsymbol{\psi}^{(m)} + \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{p}_n}\cdot\boldsymbol{\psi}^{(m)} \\ &= -\frac{\partial \mathbf{A}}{\partial \boldsymbol{p}_n}\boldsymbol{s}\cdot\boldsymbol{\psi}^{(m)} + \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{p}_n}\cdot\boldsymbol{\psi}^{(m)}.\end{aligned} \tag{14}$$

Recall that in most cases $\mathbf{A}^t = \mathbf{A}$.

In other words, computing $\mathbf{S}^{(s)}$ with the adjoint-state approach requires the solution of Eq. (13) for each m and then the application of Eq. (14). This procedure could appear cumbersome, but it is to be recalled that for the application of the model, it is necessary to have an efficient code, function or routine for the solution of the FP: for a single value of m, only one run of the same tool can be used to compute $\boldsymbol{\psi}^{(m)}$ as the solution to Eq. (13), and then $\mathbf{S}^{(s)}_{mn}$, for $n = 1, \ldots, N_\mathrm{p}$ by means of Eq. (14), where N_p is the number of calibrated parameters.

Also the simplest "finite-differences" approach requires the solution of FPs, but it is always approximate. The adjoint-state approach, based on Eqs. (13) and (14), provides a result, which is theoretically perfect and affected only by rounding errors.

5 Conclusions

Inverse modelling is of paramount importance in glaciological sciences to estimate parameters which can hardly be measured (for instance, the parameters of the Glen's law, basal temperature and melt rate, etc.) by taking advantage of the collection of data on more easily accessible physical quantities (for instance, the ice-sheet surface, ice velocity, surface temperature, etc.). Different definitions are given by different authors and there is also a great variety of approaches to the discretization of the equations that translate in mathematical form the basic conservation principles (mass, linear momentum, energy). The conceptual framework proposed in this paper permits to unify the different notations and to facilitate the formal definitions of the IP and its properties.

First, the weight of experimental and monitoring data on IP has been clarified together with the role of the parameters that are kept fixed and are not subjected to the fitting procedure. Model predictions and calibration targets seldom correspond in a straightforward way with, respectively, the results of the balance equations, which can be cast in the form shown by Eq. (4), and the directly measured quantities. Some processing is often required and its role for the mathematical properties of the IP should be considered.

The simple prototype example of inferring the leading coefficient of Glen's law and of accumulation rate for a SIA-based, uncoupled model is a proof that the classical statement that IPs are ill-posed is not always true. In fact, such an example shows an IP which has a unique and stable solution. However, this remark should not be misunderstood as a statement that IP can be easily and efficiently solved. Even in the best cases, when IPs are well-posed, the great number of processes, which introduce discrepancies between model outcomes and calibration targets (measurement errors, modelling errors, wrong estimates of fixed parameters, etc.), do not permit to reduce such discrepancies at will. Moreover, in many cases IPs could be ill-conditioned, so that even a small error on the input data could cause high and physically unacceptable errors on the calibrated parameters.

The paradigmatic example, despite being quite simple, also shows that the methods of solution to the IP can introduce some issues. In particular, they could suffer from some troubles when the objective function has multiple local minima or is very flat around the minimum. Unfortunately, in some cases the methods of solution could also reduce or mask the intrinsic difficulties of the IP because they do not span the whole space of the admissible values of model parameters.

Finally, one of the most important messages of this paper is that before drawing any conclusion on the results of the model calibration, the IP under study should be clearly defined and its properties should be properly and thoroughly analysed.

Acknowledgements. This work was financially supported by CNR through the project "Caratterizzazione biogeochimica dei laghi sub-glaciali antartici" (CaBiLA) (2009/A2.02-PNRA).

References

Arthern, R. J. and Gudmundsson, H.: Initialization of ice-sheet forecasts viewed as an inverse Robin problem, J. Glaciol., 56, 527–533, doi:10.3189/002214310792447699, 2010. 5514, 5519

Arthern, R. J. and Hindmarsh, R. C. A.: Optimal estimation of changes in the mass of ice sheets, J. Geophys. Res.-Earth, 108, 6007, doi:10.1029/2003JF000021, 2003. 5514

Avdonin, S., Kozlov, V., Maxwell, D., and Truffer, M.: Iterative methods for solving a nonlinear boundary inverse problem in glaciology, J. Inverse Ill-Pose. P., 17, 209–310, doi:10.1515/JIIP.2009.018, 2009. 5514

Baratelli, F., Giudici, M., and Vassena, C.: Dynamics of ice sheets: a critical review of the shallow-ice approximation, Bollettino geofisico, XXXIV, 5–18, 2011. 5515

Baratelli, F., Giudici, M., and Vassena, C.: A sensitivity analysis for an evolution model of the Antarctic ice sheet, Reliab. Eng. Syst. Safe., 107, 64–70, doi:10.1016/j.ress.2011.07.003, 2012. 5515, 5523

Berliner, L., Jezek, K., Cressie, N., Kim, Y., Lam, C., and van der Veen, C.: Modeling dynamic controls on ice streams: a Bayesian statistical approach, J. Glaciol., 54, 705–714, doi:10.3189/002214308786570917, 2008. 5519

Bonan, B., Nodet, M., Ritz, C., and Peyaud, V.: An ETKF approach for initial state and parameter estimation in ice sheet modelling, Nonlin. Processes Geophys., 21, 569–582, doi:10.5194/npg-21-569-2014, 2014. 5514

Bueler, E.: Bueler's Karthaus Materials, available at: https://github.com/bueler/karthaus, last access: 24 October 2014. 5515

Bueler, E., Lingle, C. S., Kallen-Brown, J. A., Covey, D., and Bowman, L. N.: Exact solutions and verification of numerical models for isothermal ice sheets, J. Glaciol., 51, 291–306, doi:10.3189/172756505781829449, 2005. 5515

Calderòn, A. P.: On an inverse boundary value problem, Comput. Appl. Math., 25, 133–138, 2006. 5519

Cuffey, K. and Paterson, W.: The Physics of Glaciers, Elsevier Science, Burlington, MA, USA, Oxford, UK, 2010. 5513

Edwards, A.: Likelihood, Cambridge University Press, 1972. 5520

Gillet-Chaulet, F., Gagliardini, O., Seddik, H., Nodet, M., Durand, G., Ritz, C., Zwinger, T., Greve, R., and Vaughan, D. G.: Greenland ice sheet contribution to sea-level rise from a new-generation ice-sheet model, The Cryosphere, 6, 1561–1576, doi:10.5194/tc-6-1561-2012, 2012. 5514, 5519

Giudici, M.: A result concerning identifiability of the inverse problem of groundwater hydrology, Inverse Probl., 5, L31–L36, doi:10.1088/0266-5611/5/3/003, 1989. 5514

Giudici, M.: Identifiability of distributed physical parameters in diffusive-like systems, Inverse Probl., 7, 231–245, doi:10.1088/0266-5611/7/2/007, 1991. 5514

Giudici, M.: Some problems for the application of inverse techniques to environmental modelling, Contemporary mathematics 333, American Mathematical Society, Providence, RI, USA, 89–97, 2002. 5514

Giudici, M. and Vassena, C.: Spectral Analysis of the balance equation of ground water hydrology, Transport Porous Med., 72, 171–178, doi:10.1007/s11242-007-9142-3, 2008. 5521

Giudici, M., Baratelli, F., Castellani, G., and Vassena, C.: Modeling the Antarctic ice sheet and ice shelves: assessing the effects of uncertainty on the model parameters by sensitivity analysis, Earth Sciences in the 21st Century Series, Nova Science Publishers, Inc., New York, 121–142, 2012. 5523, 5524

Goldberg, D. N. and Heimbach, P.: Parameter and state estimation with a time-dependent adjoint marine ice sheet model, The Cryosphere, 7, 1659–1678, doi:10.5194/tc-7-1659-2013, 2013. 5525

Greve, R.: Thermomechanisches Verhalten polythermer Eisschilde – Theorie, Analytik, Numerik, Doctoral thesis, Department of Mechanics, Darmstadt University of Technology, Germany, Berichte aus der Geowissenschaft, Shaker Verlag, Aachen, Germany, 1995. 5513

Greve, R. and Blatter, H.: Dynamics of Ice Sheets and Glaciers, Springer, Berlin, Heidelberg, Germany, 2009. 5513

Gudmundsson, G. H.: Inverse methods in glaciology, in: Encyclopedia of Snow, Ice and Glaciers, edited by: Singh, V., Singh, P., and Haritashya, U., Encyclopedia of Earth Sciences Series, Springer Netherlands, Dordrecht, the Netherlands, 653–656, doi:10.1007/978-90-481-2642-2_648, 2014. 5514

Gudmundsson, G. H. and Raymond, M.: On the limit to resolution and information on basal properties obtainable from surface data on ice streams, The Cryosphere, 2, 167–178, doi:10.5194/tc-2-167-2008, 2008. 5514

Habermann, M., Maxwell, D., and Truffer, M.: Reconstruction of basal properties in ice sheets using iterative inverse methods, J. Glaciol., 58, 795–807, doi:10.3189/2012JoG11J168, 2012. 5514, 5523

Heimbach, P. and Bugnion, V.: Greenland ice-sheet volume sensitivity to basal, surface and initial conditions derived from an adjoint model, Ann. Glaciol., 50, 67–80, doi:10.3189/172756409789624256, 2009. 5515, 5523, 5525

Heimbach, P. and Losch, M.: Adjoint sensitivities of sub-ice-shelf melt rates to ocean circulation under the Pine Island Ice Shelf, West Antarctica, Ann. Glaciol., 53, 59–69, doi:10.3189/2012/AoG60A025, 2012. 5515, 5523, 5525

Hill, M. C. and Tiedeman, C. R.: Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions and Uncertainty, J. Wiley & Sons, Hoboken, NJ, USA, 2006. 5515

Hooke, R. L.: Principles of Glacier Mechanics, 2nd Edn., Cambridge University Press, 2005. 5513

Hutter, K.: Theoretical Glaciology, D. Reidel Publishing Co., Dordrecht, Terra Scientific Publishing Co., Tokyo, 1983. 5513, 5515

Huybrechts, P., Payne, T., and Eismint, I. G.: The EISMINT benchmarks for testing ice-sheet models, Ann. Glaciol., 23, 1–12, 1996. 5513, 5516

Joughin, I., MacAyeal, D. R., and Tulaczyk, S.: Basal shear stress of the Ross ice streams from control method inversions, J. Geophys. Res.-Sol. Ea., 109, B09405, doi:10.1029/2003JB002960, 2004. 5514

Kohn, R. and Vogelius, M.: Determining conductivity by boundary measurements, Commun. Pur. Appl. Math., 37, 289–298, doi:10.1002/cpa.3160370302, 1984. 5519

Larour, E., Schiermeier, J., Rignot, E., Seroussi, H., Morlighem, M., and Paden, J.: Sensitivity Analysis of Pine Island Glacier ice flow using ISSM and DAKOTA, J. Geophys. Res.-Earth, 117, F02009, doi:10.1029/2011JF002146, 2012. 5523

Le Brocq, A. M., Payne, A. J., and Vieli, A.: An improved Antarctic dataset for high resolution numerical ice sheet models (ALBMAP v1), Earth Syst. Sci. Data, 2, 247–260, doi:10.5194/essd-2-247-2010, 2010. 5516

MacAyeal, D. R.: The basal stress distribution of Ice Stream E, Antarctica, inferred by control methods, J. Geophys. Res.-Sol. Ea., 97, 595–603, doi:10.1029/91JB02454, 1992. 5514

MacAyeal, D. R.: A tutorial on the use of control methods in ice-sheet modeling, J. Glaciol., 39, 91–98, 1993. 5514

Martin, N. and Monnier, J.: Adjoint accuracy for the full Stokes ice flow model: limits to the transmission of basal friction variability to the surface, The Cryosphere, 8, 721–741, doi:10.5194/tc-8-721-2014, 2014. 5525

Menke, W.: Geophysical Data Analysis: Discrete Inverse Theory, 3rd Edn., Academic Press, Waltham, MA, USA, San Diego, CA, USA, Oxford, UK, Amsterdam, the Netherlands, 2012. 5514, 5519, 5520

Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., and Aubry, D.: Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica, Geophys. Res. Lett., 37, L14502, doi:10.1029/2010GL043853, 2010. 5514, 5519

Parker, R.: Geophysical Inverse Theory, Princeton University Press, Princeton, NJ, USA, Chircherster, West Sussex, UK, 1994. 5514

Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G. H., Hindmarsh, R. C. A., Hubbard, A., Johnson, J. V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A. J., Pollard, D., Price, S., Rückamp, M., Saito, F., Souček, O., Sugiyama, S., and Zwinger, T.: Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP–HOM), The Cryosphere, 2, 95–108, doi:10.5194/tc-2-95-2008, 2008. 5513, 5516

Payne, A. J., Huybrechts, P., Abe-Ouchi, A., Calov, R., Fastook, J. L., Greve, R., Marshall, S. J., Marsiat, I., Ritz, C., Tarasov, L., and Thomassen, M. P. A.: Results from the EISMINT model intercomparison: the effects of thermomechanical coupling, J. Glaciol., 46, 227–238, doi:10.3189/172756500781832891, 2000. 5513

Petra, N., Zhu, H., Stadler, G., Hughes, T. J., and Ghattas, O.: An inexact Gauss–Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model, J. Glaciol., 58, 889–903, doi:10.3189/2012JoG11J182, 2012. 5514, 5515, 5523

Plessix, R.-E.: A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, Geophys. J. Int., 167, 495–503, doi:10.1111/j.1365-246X.2006.02978.x, 2006. 5525

Pollard, D. and DeConto, R. M.: A simple inverse method for the distribution of basal sliding coefficients under ice sheets, applied to Antarctica, The Cryosphere, 6, 953–971, doi:10.5194/tc-6-953-2012, 2012. 5514

Raymond, M. J. and Gudmundsson, G. H.: Estimating basal properties of ice streams from surface measurements: a non-linear Bayesian inverse approach applied to synthetic data, The Cryosphere, 3, 265–278, doi:10.5194/tc-3-265-2009, 2009. 5514, 5519

Ritz, C., Rommelaere, V., and Dumas, C.: Modeling the evolution of Antarctic ice sheet over the last 420,000 years: implications for altitude changes in the Vostok region, J. Geophys. Res.-Atmos., 106, 31943–31964, doi:10.1029/2001JD900232, 2001. 5523

Rutt, I. C., Hagdorn, M., Hulton, N. R. J., and Payne, A. J.: The Glimmer community ice sheet model, J. Geophys. Res.-Earth, 114, F02004, doi:10.1029/2008JF001015, 2009. 5513

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis – The Primer, John Wiley & Sons, Chichester, UK, 2008. 5515, 5525

Schäfer, M., Zwinger, T., Christoffersen, P., Gillet-Chaulet, F., Laakso, K., Pettersson, R., Pohjola, V. A., Strozzi, T., and Moore, J. C.: Sensitivity of basal conditions in an inverse model: Vestfonna ice cap, Nordaustlandet/Svalbard, The Cryosphere, 6, 771–783, doi:10.5194/tc-6-771-2012, 2012. 5523

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004. 5514

the PISM authors: PISM, a Parallel Ice Sheet Model, available at: http://www.pism-docs.org, last access: 24 October 2014. 5513

Thoma, M., Grosfeld, K., Mayer, C., and Pattyn, F.: Ice-flow sensitivity to boundary processes: a coupled model study in the Vostok Subglacial Lake area, Antarctica, Ann. Glaciol., 53, 173–180, doi:10.3189/2012AoG60A009, 2012. 5523

Truffer, M.: The basal speed of valley glaciers: an inverse approach, J. Glaciol., 50, 236–242, doi:10.3189/172756504781830088, 2004. 5514

van der Veen, C. J.: Fundamentals of Glacier Dynamics, A. A. Balkema, Rotterdam, Brookfield, 1999. 5513

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description, The Cryosphere, 5, 715–726, doi:10.5194/tc-5-715-2011, 2011. 5513

Table 1. Initial points for the search of the minimum of $\mathcal{O}$; final points of the procedure of minimization and the corresponding values of $\mathcal{O}$.

Initial points	Final points	$\mathcal{O}$ at the final point
$(0.38, 2)$	$(0.3002, 3.002)$	0.992 m
$(0.2, 2)$	$(0.2915, 2.914)$	0.992 m

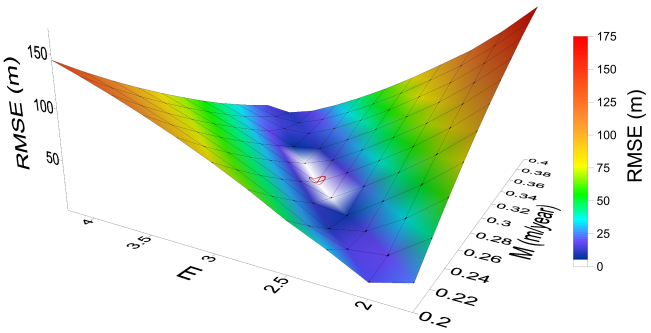

Figure 1. Objective function for the test example: the red symbol corresponds to the reference values.

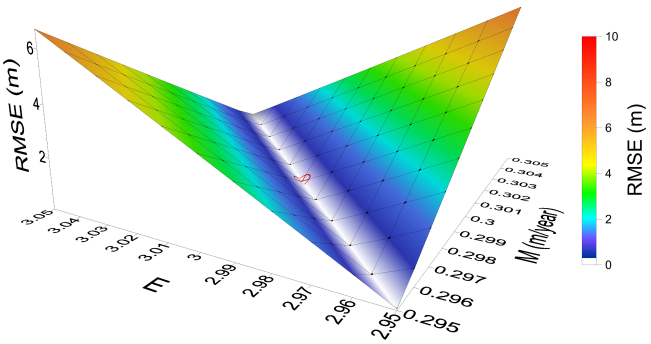

Figure 2. Zoom of the objective function for the test example around the reference values.

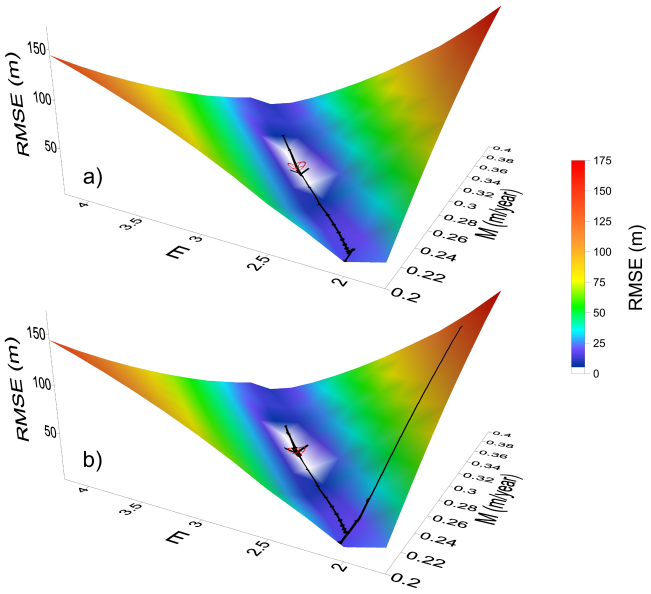

Figure 3. Objective function for the test example when an uncorrelated gaussian error with zero mean and SD of 1 m is added to the input data. The two panels show the minimisation paths obtained with the `fminunc` function starting from two different sets of parameters: **(a)** $M = 0.2$ ma^{-1}, $E = 2$; **(b)** $M = 0.38$ ma^{-1}, $E = 2$. The red symbol corresponds to the reference values $(M^{(\mathrm{ref})}, E^{(\mathrm{ref})})$.