MathematicS

MathS In A.

In Action

Jérôme J. Fehrenbach, Oleg E. Melnik & Anastassia Y. Borisova

Inverse Stefan problem from indirect measurements, application to zircon crystallization

Volume 14 (2025), p. 1-21.

https://doi.org/10.5802/msia.41

MathematicS In Action *est membre du*
Centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2102-5754

MathematicS In Action
Vol. 14, 1-21 (2025)

Inverse Stefan problem from indirect measurements, application to zircon crystallization

Jérôme J. Fehrenbach [*]
Oleg E. Melnik [**]
Anastassia Y. Borisova [***]

[*] Institut de Mathématiques de Toulouse UMR 5219, CNRS, Université de Toulouse, F-31062 Toulouse Cedex 9, France
E-mail address: jerome.fehrenbach@math.univ-toulouse.fr
[**] Institut des Sciences de la Terre, Université Grenoble Alpes, CS 40700 F-38058 Grenoble, Cedex 9, France & Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK
E-mail address: oleg.melnik@earth.ox.ac.uk
[***] Géosciences Environnement Toulouse, GET, Université de Toulouse, CNRS, IRD, 14 Avenue E. Belin, 31400 Toulouse, France
E-mail address: anastassia.borisova@get.omp.eu.

Abstract

The growth of zircon crystals in cooling magmas is modelled by the one-phase Stefan problem, with the growth rate that depends on the magma cooling rate. Some rare elements (like e.g. U, Th, Hf) are incorporated in the crystal at trace concentration. These elements have different temperature-dependent diffusion and partition coefficients. As a consequence their final spatial repartition in the crystal depends on the temperature evolution of the magma during the cooling.

The present work proposes to reconstruct the temperature evolution from the measurements of trace elements concentration in natural zircon. This inverse problem is solved by minimizing the misfit between calculated and measured trace element concentrations. The tangent model to the one-phase Stefan problem provides the sensitivity matrix, and the quadratic cost-function is minimized using Gauss–Newton method. The identifiability and the error on the retrieved parameters are studied in the framework of BLUE (Best Linear Unbiased Linear Estimator). The algorithm is tested on two synthetic datasets, and on real data obtained in a zircon crystal from early Fish Canyon tuff eruption. Reconstructed temperature ranges and cooling duration are in good agreement with available petrological interpretation.

1. Introduction

Motivation

Most of volcano-magmatic processes are hidden from direct observations. For prehistoric eruptions only analysis of their products is suitable for reconstruction of the evolution. Crystals of different minerals provide access to magma storage and ascent conditions. Among these crystals zircons are widely used for dating of different events on timescales ranging from decades to millions of years. This is due to very slow zircon growth, extreme resistance to external factors and active incorporation of trace elements, especially U, Th and Hf, during the growth.

Distribution of trace elements within zircon crystals reflects conditions, especially the temperature evolution in a parcel of magma in which the crystal was growing. At thermodynamic equilibrium the concentration of trace elements in the crystal is proportional to its concentration

OM acknowledges support from College de France PAUSE program (C7H-PUB23A59) and thanks visiting professor position in Observatoire Midi-Pyrénées in Toulouse, France.
Keywords: One phase Stefan problem, inverse problem, accessory minerals, crystal growth.
2020 *Mathematics Subject Classification:* 65M32, 86-08, 86A22.

in the surrounding melt. The proportionality constant is called partition or distribution coefficient. Usually it strongly depends on the temperature and magma composition. If conditions in the magma body change faster than a characteristic time of diffusion equilibration, only local thermodynamic equilibrium can occur on the interface between zircon crystal and surrounding melt. A boundary layer is formed around the growing crystal where the trace element content is different from the bulk composition. For trace elements with slow diffusion and high partition coefficient, such as Hf, the boundary layer will be strongly depleted.

A forward model of zircon growth with account for trace element incorporation was developed in [13]. This model is adapted to the case of a spherically symmetric domain, it consists of a set of 1D diffusion equations for zirconium and trace elements together with appropriate boundary conditions that reflect mass conservation and conditions of local thermodynamic equilibrium on moving interfaces. Partition and diffusion coefficients are temperature dependent. Note that the temperature is considered to be constant in space within the domain since the time scale of thermal diffusion is much faster than the time scale of chemical diffusion.

Results of the forward model indicate that small variations of the temperature with time can lead to significant variations in trace element concentrations inside the crystal due to a strong dependence of trace element contents in the boundary layer on the growth rate of the crystal, which in turn is governed by the diffusive flux of zirconium.

Objectives and methods

The goal of the present paper is to propose and evaluate a reconstruction method for the temperature evolution during zircon crystal growth based on the distribution of different trace elements within the crystal. This inverse problem will be solved within a variational data assimilation framework.

Variational data assimilation is a general approach to estimate the parameters of a model, when observations on the solution are available. It can be viewed as a constrained minimization problem, where the constraint is the model (here it is a partial differential equation) and the objective function $\mathcal{J}$ to be minimized is the error between the observations and the predictions of the model. Relevant mathematical tools are based on the theory of optimal control, see e.g. the reference book [11].

The dependence of the objective function $\mathcal{J}$ on the parameters is composed of two steps: the solution of the model depends on the parameters (parameter-to-state map), and the observations depend on the solution (state-to-observation map). In weather forecasting applications, the number of parameters is large, and the gradient of the objective function $\mathcal{J}$ is estimated using an adjoint approach [10]. The parameters are then retrieved using a gradient descent, or a method of the L-BFGS type [12].

In our case, the state-to-observation map is non-linear, since the state is described over a time interval $[0, T]$ by the evolution of the location of the boundary of the crystal, and the concentration of Zr and trace elements in the liquid phase. The observations consist of the concentration of each trace element at positions that correspond to the locations of the boundary of the domain, but the time is not available. However, despite its non linearity, the derivative of the state-to-observation map is analytically available.

Moreover, we take advantage of having a quadratic cost function of the form $\mathcal{J} = \frac{1}{2}\|F\|^2$ which allows to use Gauss–Newton algorithm, that has efficient convergence properties, at least close to the minimizer [15]. It requires information of first order derivatives, namely the Jacobian matrix DF, and provides a descent direction (with optimal step 1, so no line search is necessary) that yields asymptotically a quadratic convergence.

Another feature of the problem addressed here is that we proceed using a coarse-to-fine approach for the description of the temperature evolution, leading to a small number of unknown parameters. The evolution of the temperature is described by 2 parameters (in the case of

a linear ramp evolution), or a small number of parameters (piecewise linear evolution). In this situation, we propose to compute the Jacobian matrix DF column-wise using the tangent model. This requires the solution of one tangent model for each parameter, which is a reasonable computational load. The sensitivity of the parameter-to-state map is given by the tangent model, and as we already mentioned the sensitivity of the state-to-observation map is analytically calculated. The Jacobian matrix of the resulting parameter-to-observation map is called the sensitivity matrix.

If the measurements are affected by a centered gaussian noise with known covariance matrix, the framework of BLUE (Best Linear Unbiased Estimator) provides the covariance matrix for the error on the retrieved parameters. This assumes that the model is locally replaced by its tangent model, following the traditional practice in geophysical applications.

The approach proposed to retrieve the temperature evolution is illustrated by numerical experiments in order to document its efficiency. In the first two experiments, we generate a solution of the forward problem that is then perturbed by additive Gaussian noise. The last experiment uses experimental profiles measured in a zircon crystal from Fish Canyon Tuff (USA) eruption, where we use the measured concentration profiles of Hf, Y, U and Th. The Python code that was developed is provided on a repository [1], and an online version with graphical interface is available [2].

Related work

To our knowledge, no previous work addresses the specific inverse problem of retrieving the temperature evolution $\theta(t)$ from the measurements of trace elements.

However numerous contributions address inverse Stefan problems in different frameworks, where the position of the interface between phases and/or boundary conditions and/or coefficients are measured or unknown. We cite only some of them since they involve only one Stefan problem, and not the nested advection-diffusion problem describing the trace elements. The monograph [7] provides a detailed study of conditions that ensure existence, stability and regularity of the solution of one-phase and two-phase Stefan problems. In [16] the flux at the boundary is retrieved from the knowledge of the interface location by a minimization in nested finite dimensional spaces. In [5] the temperature at the fixed boundary is retrieved from the position of the interface between phases, using an adjoint approach. In [19] an approach of physics-informed neural networks is used to solve two types of Stefan inverse problems: type I problems consist in determining the Dirichlet and Neumann condition at the boundary that achieve a desired evolution of the boundary (free boundary design), and type II problems aim at reconstructing the boundary condition and the evolution of the boundary from the measurement of a small number of values of the temperature at different times and locations.

Contents

The manuscript is organized as follows: in Section 2 we describe the forward problem of zircon crystallization (it is a one-phase Stefan problem) together with the incorporation of trace elements (one advection-diffusion problem for each element). The continuous formulation in spherical symmetry is provided, together with the discretization that was implemented. The result depends on the temperature evolution $\theta(t)$, for $t \in [0, T]$ where T is the total cooling time that is considered. In Section 3 we discuss the ill posedness of the inverse problem, and illustrate this ill posedness by a numerical experiment. We also describe our choice for the regularization strategy, namely by dimension reduction, and the choice of the minimization method, namely Gauss–Newton using the tangent model. In Section 4 we derive the tangent model to the Stefan problem for Zircon and to the advection-diffusion problems for the trace elements. We compute the derivative with respect to any perturbation of the temperature evolution η. We also obtain

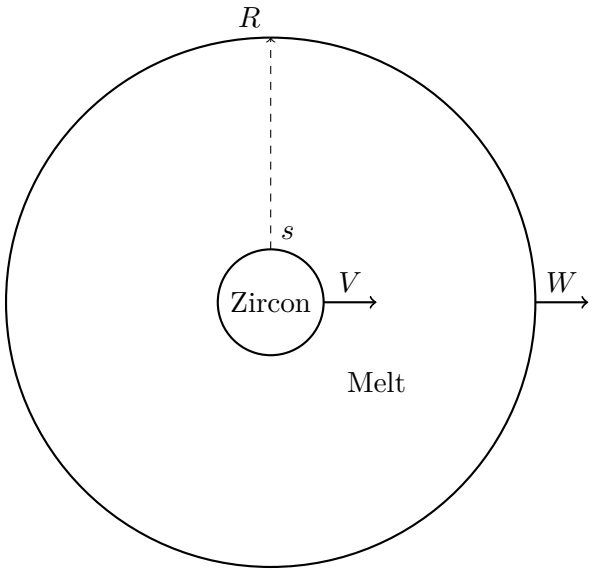

Figure 2.1. Radial model setup with two moving boundaries

the derivative with respect to the total cooling time T. In Section 5 we detail the variational approach that was used to estimate the temperature history, together with the algorithms that were implemented to obtain respectively the total cooling time (assuming linear decrease of the temperature), and a more detailed temperature evolution is a small dimensional subspace. Section 6 shows results obtained on synthetic and real data, and the manuscript closes by summary and a discussion on the geological implications of this work.

2. The forward problem

The crystal growth is described by a one-phase Stefan problem, which is a diffusion problem in a moving domain, where the thermodynamic equilibrium condition at the boundary leads to a fixed point problem that admits as solution both the concentration in the domain and the boundary velocity. Following [3] we assume spherical symmetry of the domain, which leads [13] to a one-dimensional Stefan problem in a moving interval $[s(t), R(t)]$, where $s(t)$ is the position of the interface between the crystal and the melt, and $R(t)$ is the outer boundary of the melt cell, see Figure 2.1. After spatial rescaling the space variable is $x \in [0, 1]$.

2.1. Notation

We follow the notation from [13], excepted that we have a different convention for the orientation of the velocities of the crystal boundaries, see Figure 2.1. The outgoing velocities of the inner, resp. outer, boundaries of the domain are denoted $V(t)$, resp. $W(t)$.

We introduce the following notation: $u_{\mathrm{Zr}}(t, x)$ is the zirconium concentration in the melt phase, C_{Sat} is its value at the interface between the growing zircon crystal and the melt, $C^{\mathrm{S}}_{\mathrm{Zr}} = 490\,000\,\mathrm{ppm}$ is the concentration of zirconium in zircon crystal, Kpl is the partition coefficient for zirconium in the crystals of major magmatic phases that are growing on the outer boundary, Kpl_{i} and Ktr_{i} are partition coefficients for trace element i between the zircon and major phases, respectively. We use experimentally determined values from [17] for these 3 quantities.

2.2. **Continuous forward model**

In the rescaled geometry, the space variable is $x \in [0,1]$ and the concentrations in the melt are denoted respectively $u_{\mathrm{Zr}}(t,x)$ for zircon, and $u_{\mathrm{i}}(t,x)$, $i = 1 \dots p$ for trace elements. The temperature evolution is described by its value $\theta(t)$ at time t. The evolution of the geometry is governed by the cristallization/dissolution of zircon. The quantities $u_{\mathrm{Zr}}(t,x), V(t)$ solve the following :

$$\begin{cases} \dfrac{\partial u_{\mathrm{Zr}}}{\partial t} + \dfrac{(x-1)V + xW}{R-s}\dfrac{\partial u_{\mathrm{Zr}}}{\partial x} = \dfrac{D_{\mathrm{Zr}}}{(R-s)^2(s+x(R-s))^2}\dfrac{\partial}{\partial x}\left((s+x(R-s))^2\dfrac{\partial u_{\mathrm{Zr}}}{\partial x}\right), \\ -\dfrac{D_{\mathrm{Zr}}}{R-s}\dfrac{\partial u_{\mathrm{Zr}}}{\partial x}(x=0) = V(C_{\mathrm{Sat}} - C^{\mathrm{S}}_{\mathrm{Zr}}), \\ u_{\mathrm{Zr}}(x=0) = C_{\mathrm{Sat}}, \\ -\dfrac{D_{\mathrm{Zr}}}{R-s}\dfrac{\partial u_{\mathrm{Zr}}}{\partial x}(x=1) = W(1-Kpl), \end{cases} \tag{2.1}$$

where the diffusion coefficient and equilibrium concentration change with temperature θ: $D_{\mathrm{Zr}} = D_{\mathrm{Zr}}(\theta(t))$ and $C_{\mathrm{Sat}} = C_{\mathrm{Sat}}(\theta(t))$. The velocity of the outer boundary is given by

$$W(t) = -\frac{1}{3R^2}\frac{\mathrm{d}X}{\mathrm{d}t}, \tag{2.2}$$

where $X = X(\theta)$ is the volume fraction of the major phases that simultaneously crystallise in magma and reduce the volume of the melt cell ($W(t) \leq 0$). The volume fraction depends on the temperature θ, and is given as an explicit analytic function of θ. The system (2.1) is thus a one-phase Stefan problem with unknowns u_{Zr}, V. The one-dimensional problem (without the terms accounting for spherical geometry) are known to have solutions that possess Hölder regularity together with their derivatives of order 2 in space and 1 in time, as long as the parameters $D(t), C_{\mathrm{Sat}}(t), W(t)$ are regular enough, see e.g. Theorem 4.4.4 in [7] for a precise statement.

The evolution of the boundaries of the melt domain is given by

$$\begin{cases} \dfrac{\partial s}{\partial t} = V, \\ \dfrac{\partial R}{\partial t} = W. \end{cases} \tag{2.3}$$

The system (2.1)–(2.3) is complemented with the following initial condition:

$$u_{\mathrm{Zr}}(t=0,\cdot) = C_{\mathrm{Sat}}(\theta(t=0)), \quad s(t=0) = s_0, \quad R(t=0) = R_0. \tag{2.4}$$

The velocities $V(t), W(t)$ are, thus, imposed by the evolution of the Zr (major phase) crystallisation. For the concentration of trace elements the following advection-diffusion problem are then solved, for $\mathrm{i} = 1 \dots p$:

$$\begin{cases} \dfrac{\partial u_{\mathrm{i}}}{\partial t} + \dfrac{(x-1)V + xW}{R-s}\dfrac{\partial u_{\mathrm{i}}}{\partial x} = \dfrac{D_{\mathrm{i}}}{(R-s)^2(s+x(R-s))^2}\dfrac{\partial}{\partial x}\left((s+x(R-s))^2\dfrac{\partial u_{\mathrm{i}}}{\partial x}\right), \\ -\dfrac{D_{\mathrm{i}}}{R-s}\dfrac{\partial u_{\mathrm{i}}}{\partial x}(x=0) = V(1-Ktr_{\mathrm{i}}), \\ -\dfrac{D_{\mathrm{i}}}{R-s}\dfrac{\partial u_{\mathrm{i}}}{\partial x}(x=1) = W(1-Kpl_{\mathrm{i}}). \end{cases} \tag{2.5}$$

2.3. **Observation operator**

Profiles of concentrations of the trace elements as a function of distance from the core of the crystal can be measured in laboratory (see example on Figure 1 in [13] or Section 6.5 below). These profiles are formed during a growing-only period, when no dissolution occurred, in other words $V(t) \geq 0$ for $t \in [0,T]$. These profiles will be used as observations for the inverse problem.

Each point in the profile corresponds to an instant t of crystallization where the concentration of each trace element at the boundary is

$$u_{\mathrm{i}}(t, x = 0),$$

and the boundary of the crystal is located at $s(t)$. In other words, the observations consist of the values

$$(s(t), u_{\mathrm{i}}(t, x = 0))$$

for some observation instants $t = \tau_0, \tau_1, \dots \tau_{Nobs}$. Let us emphasize that only the values of (s, u_{i}) are observed, but not the values of the time t.

2.4. **Discrete forward model**

The system (2.1)–(2.2)–(2.3)–(2.4)–(2.5) is approximated using a finite volume spatial discretization, and an implicit time discretization, as described in [3]. We recall here the method and define additional notation that will be used in the sequel.

Let us describe first the semi-discretization in space. The space discretization points are fixed and denoted $x_0 = 0 < x_1 < \cdots < x_{N_x} = 1$. We use the notation $\mathbf{u}_{\mathrm{Zr}}(t)$ for the semi-discrete version in space of $u_{\mathrm{Zr}}(t, \cdot)$. We consider in system (2.1) the Dirichlet condition at $x = 0$. For a given value of V the finite volume matrices arising after multiplying by $(R-s)^2(s+x(R-s))^2$ and integrating between $x_{k-1/2}$ and $x_{k+1/2}$ (see [3]) are denoted $L(V), F(V)$. Note that the matrix F depends on D_{Zr} and C_{Sat} and therefore on $\theta(t)$, this dependence will be emphasized by denoting $F(V, \theta)$. The matrices L and F also depend on s and R but for the sake of concision we will not use a notation for this dependence for the moment. The semi-discrete in space version of (2.1) reads

$$\begin{cases} L(V)\dfrac{\partial \mathbf{u}_{\mathrm{Zr}}}{\partial t} = F(V, \theta(t))\mathbf{u}_{\mathrm{Zr}}, \\ \qquad V = -\dfrac{D_{\mathrm{Zr}}}{(R-s)(C_{\mathrm{Sat}} - C^{\mathrm{S}}_{\mathrm{Zr}})}\dfrac{u_{\mathrm{Zr}}(t, x_1) - u_{\mathrm{Zr}}(t, x_0)}{x_1 - x_0}. \end{cases} \tag{2.6}$$

Let us denote $\mathbf{u}^j_{\mathrm{Zr}}$, resp. $V^j, R^j, s^j, j = 0 \dots N$ the value of the fully discrete approximation of u_{Zr}, resp. of the time-approximation of V, R, s at time t_j. Denote $\theta^j = \theta(t_j)$. Then the time-implicit discretization of (2.6)–(2.2)–(2.3)–(2.4) yields

$$\begin{cases} \mathbf{u}^0_{\mathrm{Zr}} = C_{\mathrm{Sat}}(\theta(t = 0)), & \\ L(V^j)\mathbf{u}^j_{\mathrm{Zr}} - \Delta t\, F(V^j, \theta^j)\mathbf{u}^j_{\mathrm{Zr}} = L(V^j)\mathbf{u}^{j-1}_{\mathrm{Zr}} & j \geq 1, \\ V^j = -\dfrac{D_{\mathrm{Zr}}}{(R^j - s^j)(C_{\mathrm{Sat}} - C^{\mathrm{S}}_{\mathrm{Zr}})}\dfrac{u^j_{\mathrm{Zr}}(x_1) - u^j_{\mathrm{Zr}}(x_0)}{x_1 - x_0} & j \geq 1, \\ W^j = \dfrac{1}{3{R^j}^2}X'(\theta^j)\dfrac{\theta^j - \theta^{j-1}}{\Delta t}, & \\ s^j = s^{j-1} + \Delta t V^j & j \geq 1, \\ R^j = R^{j-1} + \Delta t W^j & j \geq 1. \end{cases} \tag{2.7}$$

For each value of $j \geq 1$ this fixed point problem for $\mathbf{u}^j_{\mathrm{Zr}}, V^j$ is solved iteratively by alternating V^j and $\mathbf{u}^j_{\mathrm{Zr}}$ updates. The convergence of the iterations is ensured by the following

Proposition 2.1. *The spatial discretization is fixed. We assume that there exists $h > 0$ such that $R^j - s^j > h$ (in other words the melt domain remains larger than h). Then for Δt small*

enough, the mapping $V \mapsto \Phi_1(V, \Phi_2(V))$ *is contracting, where*

$$\begin{aligned} \Phi_2(V) &= (L(V) - \Delta t\, F(V, \theta^j))^{-1} L(V) \mathbf{u}_{\mathrm{Zr}}^{j-1} \\ \Phi_1(V, \mathbf{u}'_{\mathrm{Zr}}) &= \frac{-D_{\mathrm{Zr}}}{(R^j - s^{j-1} - \Delta t V)(C_{\mathrm{Sat}} - C_{\mathrm{Zr}}^{\mathrm{S}})} \frac{\mathbf{u}'_{\mathrm{Zr}}(x_1) - \mathbf{u}'_{\mathrm{Zr}}(x_0)}{x_1 - x_0}. \end{aligned} \tag{2.8}$$

Justification. Since the space discretization is fixed, there is a constant K such that $|D\Phi_1(V, \mathbf{u}_{\mathrm{Zr}}).(\nu, \mu)| \leq K(\Delta t\, \nu + \|\mu\|_\infty)$. Let us denote $A = L - \Delta t F$, so that $\Phi_2(V) = A^{-1}(A + \Delta t F)\mathbf{u}_{\mathrm{Zr}}^{j-1} = \mathbf{u}_{\mathrm{Zr}}^{j-1} + \Delta t A^{-1} F \mathbf{u}_{\mathrm{Zr}}^{j-1}$. The matrices A and F depend on V, s^j, R^j (we do not provide the details here), but under the assumption that $R^j - s^j > h$ the matrix $A^{-1}F$ remains bounded, and therefore there is a constant K' such that $\|\Phi_2(V_1) - \Phi_2(V_2)\| \leq K' \Delta t |V_2 - V_1|$. Gathering the upper bounds that we obtain and denoting $f(V) = \Phi_1(V, \Phi_2(V))$, we have then $|f'(V)| = |D_1\Phi_1 + D_2\Phi_1.\Phi_2'(V)| \leq |K\,\Delta t + KK'\Delta t|$. The choice of a small enough Δt ensures that this map is contracting. □

We now denote for short

$$\begin{cases} \mathbf{u}_{\mathrm{Zr}}^0 = C_{\mathrm{Sat}}(\theta^0), \\ A(V^j, \theta^j)\mathbf{u}_{\mathrm{Zr}}^j = L(V^j)\mathbf{u}_{\mathrm{Zr}}^{j-1} & j \geq 1, \\ V^j = \varphi(\theta^j, \mathbf{u}_{\mathrm{Zr}}^j), \\ s^j = s^{j-1} + \Delta t V^j & j \geq 1, \\ R^j = R^{j-1} + \Delta t W^j & j \geq 1, \end{cases} \tag{2.9}$$

where $A = L - \Delta t F$ and φ is defined by the third equation in (2.7).

For the full discretization for trace elements from (2.5), the velocity V is known from the solution of (2.7), and only one linear solve is necessary at each time step. The finite volume matrix is different since each trace element has its specific diffusion and partition coefficients. The initial condition for each trace element is a constant [13].

Let $\mathbf{u}_{\mathrm{i}}^j$ denote the value of the fully discrete approximation of u_{i} at time t_j. Then we have

$$\begin{cases} \mathbf{u}_{\mathrm{i}}^0 = c_{\mathrm{i}}, \\ A_{\mathrm{i}}(V^j, \theta^j)\mathbf{u}_{\mathrm{i}}^j = L(V^j)\mathbf{u}_{\mathrm{i}}^{j-1} & j \geq 1, \end{cases} \tag{2.10}$$

where $A_{\mathrm{i}} = L - \Delta t F_{\mathrm{i}}$, and F_{i} is the finite volume matrix obtained by discretizing (2.5).

3. Ill posedness and regularization strategy

The general theoretical study of the identifiability for the model (2.1)–(2.2)–(2.3)–(2.4)–(2.5) is beyond the scope of the present work, that aims at proposing an effective numerical method. In [7] different settings of 1D inverse Stefan problems are studied, and numerous illposedness examples are provided. The examples have oscillating coefficients, and the regularization strategies proposed in [7] rely on boundedness assumptions in the space $H^{2+\lambda, 1+\lambda/2}$ of functions that are λ-Hölder w.r.t. space variable up 2 derivatives and $\lambda/2$-Hölder w.r.t. time variable together with their derivative. In other words, the oscillations in the coefficient are ruled out by the proposed regularization strategy.

Here we will not prove the same kind of results as in the reference book [7]. Our framework is more general since in the radial geometry the coefficients of the diffusion equation depend on the variables s and R, which describe the boundary of the domain and are therefore unknowns of the problem. The theoretical study of such a model is not in our scope. We will provide in Section 3.1 numerical examples with an oscillating temperature, where the observations will tend to some limit. This is a practical illustration of ill-posedness of our problem. In Section 3.2

we will propose a regularization strategy and discuss the algorithm to achieve its solution in Section 3.3.

3.1. **Illustration of ill posedness**

We simulated a linear temperature decrease from 1113 K to 993 K, in a time range of $T = 5000$ years. This is considered as the reference temperature evolution. The model (2.1)–(2.2)–(2.3)–(2.4)–(2.5) is then solved with parameters provided in [13] using the finite volume discretization described in Section 2.4.

Then we considered perturbations of the temperature evolution, with oscillations added of fixed amplitude and increasing frequency. Note that the amplitude of the first 2 oscillations is reduced, this is to prevent V to become negative, in other words we want a scenario where the crystal is only growing. We have considered the cases of 50, 150 and 250 oscillations in the total time interval $[0, T]$. The reference temperature evolution and the scenario with 50 oscillations are presented in Figure 3.1.

We have then considered observations of trace elements at 40 equally spaced locations. We do not consider point observations but local average around the observation point. Note that this observation smoothes out the oscillations in the trace elements concentrations. The results are presented in Figure 3.2. We observe that when the frequency of the oscillations increases, the observations issued from the perturbed temperature evolution tend to the observations with the reference temperature evolution. We have thus an example of (first terms of) a sequence of temperature evolution at an increasing $H^{1+\lambda/2}$ distance from the reference (and at a fixed L^∞ distance) that generate observations that "tend to" the reference observations.

This observation on the simulations shows the same behavior as the examples in [7] where solutions with more and more oscillating coefficients tend to provide fixed observations.

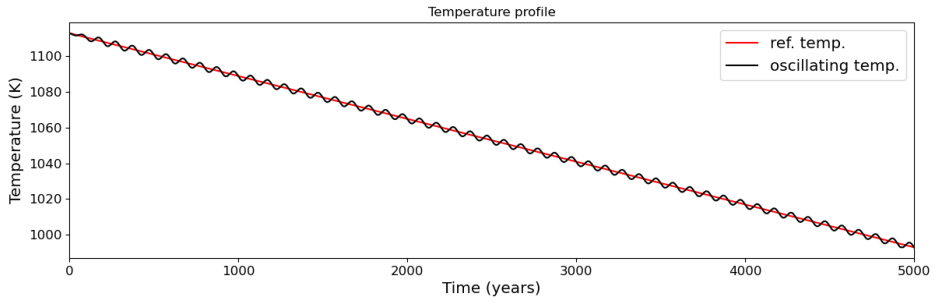

Figure 3.1. Reference and perturbed temperature evolution (case of 50 oscillations).

3.2. **Regularization**

It follows from the experiments above that the inverse problem is ill-posed, and that some regularization strategy is required to obtain its solution. Let us recall that given a map (direct model) between two Banach spaces X and Y:

$$K : X \longrightarrow Y,$$

the inverse problem of solving for $x \in X$ the equation

$$Kx = y$$

is ill-posed (in the sense of Hadamard) if it either admits no solution, or several solutions, or if its solution does not depend continously on y. We are interested in the third case of this alternative, which corresponds to inversion of a model where the observations y are corrupted

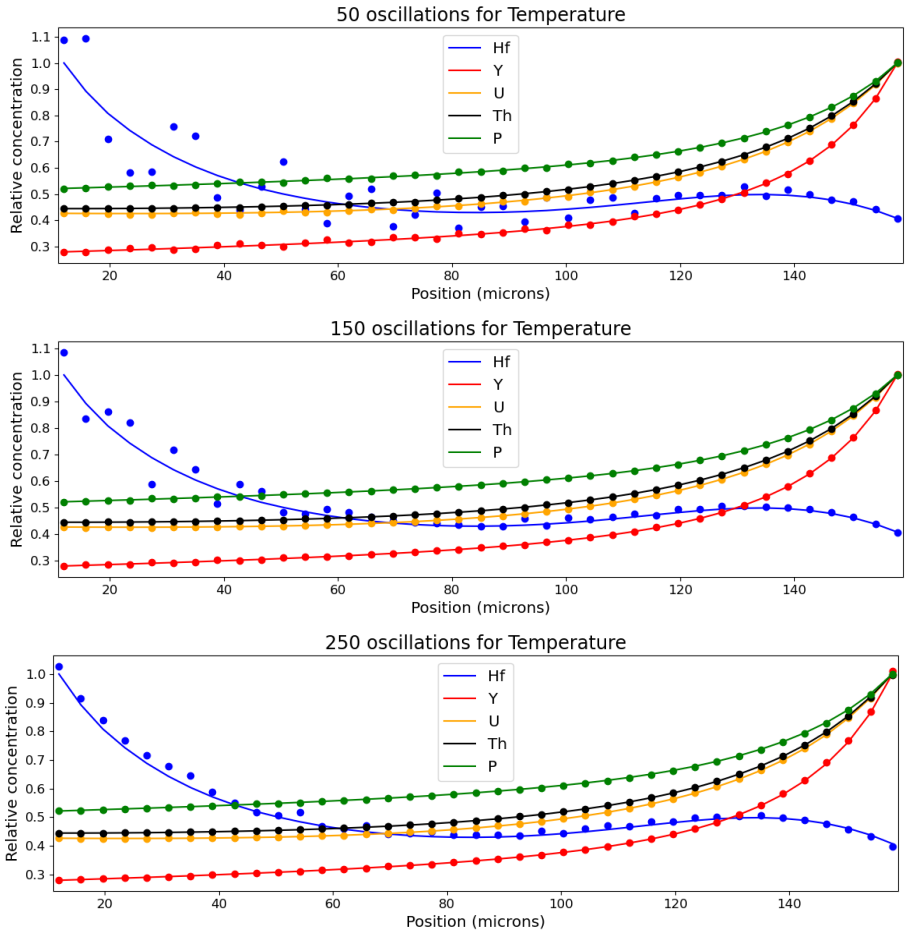

Figure 3.2. Reference (continuous line) and perturbed (dots) trace elements, from top to bottom in the case of 50, 150, 250 oscillations

by some noise, and only some y^δ is available where $\|y - y^\delta\|_Y \leq \delta$. The non-continuity of the inverse operator can be observed when some x_n are away from x but $Kx_n \to Kx$.

The brute force inversion consists of finding $x^* \in X$ that provides observations closest to y^δ:

$$x^* = \text{Argmin}_{x \in X} \|Kx - y^\delta\|_Y.$$

This approach is doomed to fail because K^{-1} is not continuous. Since the inverse problem is ill-posed, some additional information has to be incorporated in order to obtain its solution (in some sense). The strategy that accounts for this additional information is a regularization strategy, and different regularization strategies can be used, see e.g. [6, 9]. Let us cite Tikhonov regularization, that adds to the cost function above the term $\alpha\|x\|_X^2$ for some suitably chosen α depending on the noise level δ. The Landweber regularization method consists of performing some (small) number m of iterations for the steepest descent of the cost function, where m is suitably chosen knowing the noise level δ.

We propose here to implement a very simple regularization method, namely the regularization by discretization. It consists of looking for the minimum of the cost function in a small dimensional subspace X_n. In our case, we will use the subspace of piecewise linear functions on 2^k subintervals, which is of dimension $n = 2^k + 1$. We will not prove theoretical convergence estimates when the noise level tends to zero. Instead, we will experimentally study the singular values of the restriction of the (differential of the) operator K to X_n. These singular values

have to be away from 0 in order to have an efficient regularization strategy, since the continuity modulus of the inverse is the inverse of the smallest singular value of the differential of K.

3.3. **Minimization algorithm**

Let us keep for the moment general notation, and denote the observation $y^{\text{Obs.}}$. The spaces X and Y are Hilbert spaces. We thus aim at minimizing the cost function

$$j(x) = \frac{1}{2}\|Kx - y^{\text{Obs.}}\|_Y^2$$

on the finite dimensional space X_n. We take advantage of the fact that j is a quadratic function to use Gauss–Newton algorithm, which has better convergence properties than the gradient descent [15]. This requires at each iteration the knowledge of the Jacobian matrix of the operator K, which can be assembled column-wise using the tangent model. An alternative is to assemble the Jacobian matrix row-wise using the adjoint model but since the number of unknowns (the dimension of X_n) is much smaller than the number of observations it is more efficient to solve tangent problems.

Note that the classical adjoint method implements the steepest descent approach, that we do not wish to use. In the case of gradient descent, the gradient is provided using the adjoint state (which requires one solution of the adjoint problem), and the step must be adapted (which requires possibly several evaluations of the cost function). This is to be compared with our approach where the number of tangent problems to be solved is the dimension of X_n, which we will take successively to be $2, 3, 5, 9, 17$ and ultimately 33. Let us recall that the optimal step for Gauss–Newton method is 1 [15].

4. **Tangent model**

The Frechet differentiability of the solution of the continuous Stefan problem w.r.t. the parameters (either interface location, or partial measurement of the solution) is established under regularity assumptions on the coefficients in [7]. In the present work, we propose to follow a discretize-then-differentiate approach. This means that we will compute the derivative w.r.t. the temperature evolution θ of the discrete model given by systems (2.9)–(2.10). We also provide in Section 4.3 an approach to estimate the derivative w.r.t. the total time T.

4.1. **Derivative of the zirconium concentration**

We consider a discrete temperature history $\theta = \theta^0, \dots, \theta^N$. The corresponding solution of (2.9) is denoted $\mathbf{u}_{\text{Zr}} = \mathbf{u}_{\text{Zr}}(\theta)$. Now we emphasize that the matrices depend on s and R: they are denoted $A^j = A(s^j, R^j, V^j, \theta^j)$ and $L^j = L(s^j, R^j, V^j)$. Since the solution of (2.9) depends smoothly on each θ^j, the solution $\mathbf{u}_{\text{Zr}}(\theta)$ and the domain boundaries s, R depend smoothly on θ. The derivatives of these quantities in the direction $\eta = \eta^0, \dots, \eta^N$ are denoted $\mathbf{v}_{\text{Zr}}, s', R'$ and

solve the following system obtained by differentiating (2.2) and (2.9):

$$\begin{cases} \mathbf{v}_{\mathrm{Zr}}^0 = C'_{\mathrm{Sat}}(\theta^0).\eta^0, \\ s'^0 = 0, \\ R'^0 = 0, \\ A^j\mathbf{v}_{\mathrm{Zr}}^j = -D_sA^j.s'^j\mathbf{u}_{\mathrm{Zr}}^j - D_RA^j.R'^j\mathbf{u}_{\mathrm{Zr}}^j - D_VA^j.V'^j\mathbf{u}_{\mathrm{Zr}}^j - D_\theta A^j.\eta^j\mathbf{u}_{\mathrm{Zr}}^j \\ \qquad\qquad + L^j\mathbf{v}_{\mathrm{Zr}}^{j-1} + D_sL^j.s'^j\mathbf{u}_{\mathrm{Zr}}^{j-1} + D_RL^j.R'^j\mathbf{u}_{\mathrm{Zr}}^{j-1} + D_VL^j.V'^j\mathbf{u}_{\mathrm{Zr}}^{j-1} & j \geq 1, \\ V'^j = D_\theta\varphi(\theta^j,\mathbf{u}_{\mathrm{Zr}}^j).\eta^j + D_{\mathbf{u}}\varphi(\theta^j,\mathbf{u}_{\mathrm{Zr}}^j).\mathbf{v}_{\mathrm{Zr}}^j & j \geq 1, \\ s'^j = s'^{j-1} + \Delta t V'^j & j \geq 1, \\ W'^j = -\dfrac{2}{3R^3}\dfrac{\mathrm{d}X^j}{\mathrm{d}t}R'^j - \dfrac{1}{3R^2}\dfrac{\mathrm{d}X'^j}{\mathrm{d}t}, & j \geq 1, \\ R'^j = R'^{j-1} + \Delta t W'^j & j \geq 1, \end{cases} \tag{4.1}$$

where $X'^j = D_\theta X^j.\eta^j$. For every $j \geq 1$, this is a linear fixed point problem w.r.t. $\mathbf{v}_{\mathrm{Zr}}^j, V'^j$. The linear operator that appears is exactly the tangent operator to the mapping that appears in the forward model (2.7). Therefore an iterative solution of this linear fixed point problem will converge under the same condition as the iterations for the nonlinear forward model, see Proposition 2.1.

4.2. **Derivative of the trace elements concentration**

The differentiation of the system (2.10) in the direction $\eta = \eta^0, \ldots, \eta^N$ provides the derivative of $\mathbf{u}_{\mathrm{i}}$, that is denoted $\mathbf{v}_{\mathrm{i}}$. Similarly as for the zirconium case, the matrices depend on s and R and are denoted $A_{\mathrm{i}}^j = A_{\mathrm{i}}(s^j, R^j, V^j, \theta^j)$ and $L^j = L(s^j, R^j)$. The derivative $\mathbf{v}_{\mathrm{i}}$ solves the following:

$$\begin{cases} \mathbf{v}_{\mathrm{i}}^0 = 0, \\ A_{\mathrm{i}}^j\mathbf{v}_{\mathrm{i}}^j = -D_sA_{\mathrm{i}}^j.s'^j\mathbf{u}_{\mathrm{i}}^j - D_RA_{\mathrm{i}}^j.R'^j\mathbf{u}_{\mathrm{i}}^j - D_VA_{\mathrm{i}}^j.V'^j\mathbf{u}_{\mathrm{i}}^j - D_\theta A_{\mathrm{i}}^j.\eta^j\mathbf{u}_{\mathrm{i}}^j \\ \qquad\qquad + L^j\mathbf{v}_{\mathrm{i}}^{j-1} + D_sL^j.s'^j\mathbf{u}_{\mathrm{i}}^{j-1} + D_RL^j.R'^j\mathbf{u}_{\mathrm{i}}^{j-1} + D_VL^j.V'^j\mathbf{u}_{\mathrm{i}}^{j-1} & j \geq 1. \end{cases} \tag{4.2}$$

4.3. **Derivative w.r.t. the total cooling time T**

In practice, the first quantity to estimate is the total cooling time, even before the determination of the temperature profile. This can be on a time scale of 500-50000 years for typical zircon growth in magmatic systems and this time scale has to be determined. For this we use a simple temperature profile, namely a decreasing linear ramp. The aim is to estimate the sensitivity w.r.t. the slope of this ramp. We start from an initial guess where the cooling occurs at a constant slope between the values θ^0 and θ^{final} on a time interval $[0, T]$. When the total cooling time changes from the value T to the value T', the slope is modified. We can obtain an equivalent slope modification by keeping the same total time T and changing the final temperature from θ^{final} to $\theta^{\mathrm{final}'}$. The obtained trace element concentration will be the same if both temperatures θ^{final} and $\theta^{\mathrm{final}'}$ correspond to solid states (there is no further evolution in the remaining time interval).

These equivalent scenarios are presented on Figure 4.1. The relation between the different quantities is given by

$$\frac{\theta^{\mathrm{final}} - \theta^0}{T'} = \frac{\theta^{\mathrm{final}'} - \theta^0}{T}. \tag{4.3}$$

From the practical point of view, the estimation of the total cooling time T is performed by computing the sensitivity w.r.t. θ^{final}, and replacing this sensitivity with a sensitivity w.r.t. T given by Equation (4.3).

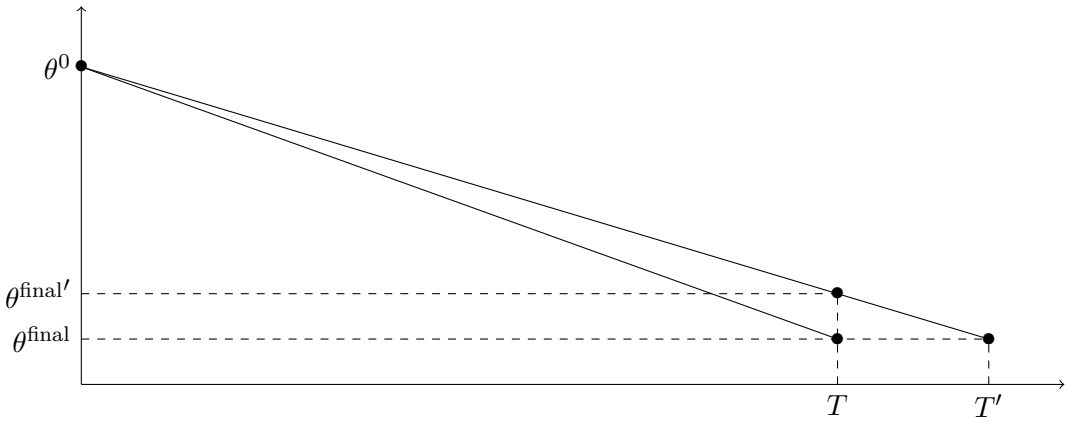

Figure 4.1. The slope is modified by a change either in the final temperature, or in the final time.

5. **Cooling history reconstruction**

We present here the variational framework that was adopted and the algorithms that were implemented to estimate the cooling history.

The cooling history is defined on a time interval $[0, T]$, and discretized at the instants

$$t_j = j\Delta t.$$

The total number of discretization instants is $N + 1$ where $N = T/\Delta t$. In order to reduce the number of unknowns (which is a method to regularize an inverse problem, see [9]) we consider a piecewise linear temperature profile θ that depends on q control points. The temperature profile is thus described by a control variable $\mathbf{x} \in \mathbf{R}^q$. The obtained concentration for trace elements will be denoted $u_{\mathrm{i}}(\mathbf{x})$.

In order to find the optimal $\mathbf{x}$ we use Gauss–Newton algorithm to minimize the quadratic misfit between the predicted and measured trace elements concentrations. We recall that the observations consist of the values of $(s, u_{\mathrm{i}}^{\mathrm{Obs.}})$ and will be presented as $(s, u_{\mathrm{i}}^{\mathrm{Obs.}}(s))$ for s ranging from $s_{\min}$ to $s_{\max}$.

The continuous version of cost function that is minimized is the following:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\sum_{\mathrm{i}=1}^{p} \omega_{\mathrm{i}} \int_{s_{\min}}^{s_{\max}} (u_{\mathrm{i}}^{\mathrm{Obs.}}(s) - u_{\mathrm{i}}(\mathbf{x})(s))^2,$$

where ω_{i} is the weight that affects the trace element with label i. For the discrete computation, the integral is replaced by a discrete sum over the values of s that are available in the observations. Note that the points where $u_{\mathrm{i}}(\mathbf{x})$ is available are not necessary the same as the points where $u_{\mathrm{i}}^{\mathrm{Obs.}}$ is available and an interpolation is performed. The discrete version reads

$$J(\mathbf{x}) = \frac{1}{2}\sum_{\mathrm{i}=1}^{p} \omega_{\mathrm{i}} \sum_{s\in S} (u_{\mathrm{i}}^{\mathrm{Obs.}}(s) - u_{\mathrm{i}}(\mathbf{x})(s))^2 = \frac{1}{2}\sum_{i=1}^{p} \omega_{\mathrm{i}} \|F_{\mathrm{i}}(\mathbf{x})\|^2, \tag{5.1}$$

where S is the set of values where the observations are available, and $u_{\mathrm{i}}(\mathbf{x})(s)$ denotes the interpolation of $u_{\mathrm{i}}(\mathbf{x})(t, 0)$ between values of t_j such that $s(t_j) \le s < s(t_{j+1})$. More precisely

$$u_{\mathrm{i}}(\mathbf{x})(s) = \alpha u_{\mathrm{i}}(\mathbf{x})(t_j, 0) + (1 - \alpha) u_{\mathrm{i}}(\mathbf{x})(t_{j+1}, 0),$$

where j is such that $s(t_j) \le s < s(t_{j+1})$ and $\alpha = \frac{s(t_{j+1}) - s}{s(t_{j+1}) - s(t_j)} \in [0, 1)$.

In order to obtain the sensitivity of J w.r.t. $\mathbf{x}$ in the direction $\mathbf{h}$ we use the derivative of $u_{\mathrm{i}}(\mathbf{x})(t)$ and the derivative of $s(t)$ in the direction $\mathbf{h}$. The derivatives of the quantities α and $\alpha u_{\mathrm{i}}(\mathbf{x})(t_j, 0) + (1 - \alpha) u_{\mathrm{i}}(\mathbf{x})(t_{j+1}, 0)$ in the direction $\mathbf{h}$ can then be deduced as long as $\alpha \neq 0$.

We commit the abuse of using the same computations when $\alpha = 0$ (even if the piecewise linear interpolant is not smooth).

These provide the derivative of $u_{\mathrm{i}}(\mathbf{x})(s)$ in the direction $\mathbf{h}$, and thus the derivative of F_{i} in the direction $\mathbf{h}$. The direction $\mathbf{h}$ is given the different values of the vectors of a basis of the parameter space, that is each of the q elements of the basis of piecewise linear maps. We obtain this way the columns of the Jacobian matrix DF_{i}.

We now detail the algorithms that implemented the Gauss–Newton minimization with the following objectives:

- Algorithm 1: Find the time scale of the cooling (estimation of the best linear ramp). A linear temperature decrease (time scale of the cooling) is determined by finding optimal θ^0 and T, as described in the following Algorithm 1.
- Algorithm 2: Find a more detailed temperature evolution by finding a piecewise linear history. In practice this algorithm is used on a coarse-to-fine approach. The Gauss–Newton minimization algorithm on a finite dimensional subspace is implemented as described in Algorithm 2 below.

Algorithm 1 Gauss–Newton method to find the optimal linear temperature decrease

Input: measurements $(s, u_{\mathrm{i}}^{\mathrm{Obs.}}(s))$ for $s \in S$ and $i = 1 \dots p$; final temperature θ^{final} ; initial guess for initial temperature θ^0 and time T
Output: initial temperature θ^0 and time T such that the temperature evolution minimizes J.
Let $(\mathbf{e}_1, \mathbf{e}_2)$ be the linear functions on $[0, T]$ s.t. $\mathbf{e}_1(0) = 1, \mathbf{e}_1(T) = 0, \mathbf{e}_2 = 1 - \mathbf{e}_1$
Set $k = 0$
while Not converged **do**
  for $i = 1 \dots p$ **do**
    Compute $F_{\mathrm{i}} = (u_{\mathrm{i}}^{\mathrm{Obs.}}(s) - u_{\mathrm{i}}(\mathbf{x})(s))_{s \in S}$
    Compute $DF_{\mathrm{i}}.\mathbf{e}_1$ and $DF_{\mathrm{i}}.\mathbf{e}_2$ by solving (4.1) and (4.2)
  end for
  Let $\mathbf{d}^k$ be the solution of $(\sum_{i=1}^p DF_{\mathrm{i}}^T DF_{\mathrm{i}})\mathbf{d}^k = -\sum_{i=1}^p DF_{\mathrm{i}}^T F_{\mathrm{i}}$
  update $\theta^0 = \theta^0 + \mathbf{d}^k[0]$, and $T = T/(1 + \mathbf{d}^k[1]/(\theta^{\mathrm{final}} - \theta_0))$
  set $k = k + 1$
end while

6. Numerical results

We first document the relevance of the dimension reduction by providing a study of the sensitivity matrix associated to each trace element, and the influence of the total time scale. Then we present the implementation of the estimation method, and show examples on synthetic and real data.

The direct and tangent problems presented here were implemented in Python 3.8 and run on a laptop equipped with a Intel(R) Core(TM) i7 CPU. The computationally intensive parts of the code (tridiagonal linear solves) were implemented using Cython.

6.1. Study of the sensitivity matrices

Let us recall that we minimize the cost function $J = \frac{1}{2}\sum_{i=1}^p \|F_i\|^2 = \frac{1}{2}\|F\|^2$, with

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_p \end{pmatrix},$$

Algorithm 2 Gauss–Newton method to find the parameter $\mathbf{x}$ that minimizes J in a finite dimensional subspace

Input: measurements $(s, u_{\mathrm{i}}^{\mathrm{Obs.}}(s))$ for $s \in S$ and $i = 1 \dots p$; basis $\mathcal{B} = (\mathbf{h}_1, \dots, \mathbf{h}_q)$ of the search space ; initial guess $\mathbf{x}^0$
Output: $\mathbf{x}$ such that the temperature evolution $\theta(\mathbf{x})$ minimizes J.
set $k = 0$
while Not converged **do**
  The temperature history is $\theta(\mathbf{x}^k) = \sum_{\ell=1}^{q} \mathbf{x}_\ell^k \mathbf{h}_\ell$
  for $i = 1 \dots p$ **do**
    compute $F_{\mathrm{i}} = F_{\mathrm{i}}(\mathbf{x}^k)$
    for $\ell = 1 \dots q$ **do**
      compute $DF_{\mathrm{i}}(\mathbf{x}^k).\mathbf{h}_\ell$ by solving (4.1) and (4.2)
      (This provides the columns of the matrix $DF_{\mathrm{i}} = DF_{\mathrm{i}}(\mathbf{x}^k)$)
    end for
  end for
  Let $\mathbf{d}^k$ be the solution of $(\sum_{i=1}^{p} DF_{\mathrm{i}}^T DF_{\mathrm{i}})\mathbf{d}^k = -\sum_{i=1}^{p} DF_{\mathrm{i}}^T F_{\mathrm{i}}$
  update $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{d}^k$
  set $k = k + 1$
end while

where each F_{i} is the parameter-to-observation operator associated to the trace element i. We consider here 5 trace elements, namely Hf, Y, U, Th, P with their diffusion coefficient and partition coefficient from [13]. The sensitivity matrix DF is a key ingredient in our approach, since each iteration involves the solution of a linear problem with matrix $DF^T DF$. Moreover, since DF measures the influence of each parameter on the observations, the directions associated to null (or very small) singular values of DF indicate directions in which the parameters can be modified without (or with little) effect on the observations. These directions would be source of ill posedness in the inverse problem.

In Bayesian data assimilation, where measurement errors are assumed to be gaussian with covariance matrix Σ, the matrix

$$(DF^T \Sigma^{-1} DF)^{-1} \tag{6.1}$$

provides the covariance of the error in the estimated parameters, see [8] or [18]. If the noise level for the measurement of each trace element is known, and the noise at every point are gaussian, independent and identically distributed then in the cost function each term can be multiplied by $\omega_{\mathrm{i}} = 1/\sigma_{\mathrm{i}}^2$, where σ_{i}^2 is the variance of the noise that affects the i-th trace element. In that case the cost function J has the weighted form in Equation (5.1), and the minimizer is the Best Linear Unbiased Estimator [18].

In practice the noise level affecting each trace element are of the same order (see Section 6.5) but the absolute magnitude of the concentration for Hf is much larger than for other elements. We investigate numerically the smallest singular value of the sensitivity matrix DF, or equivalently the smallest eigenvalue of $DF^T DF$. This amounts in this theoretical investigation to consider every $\sigma_{\mathrm{i}} = 1$. Since

$$DF^T DF = \sum_{i=1}^{p} DF_{\mathrm{i}}^T DF_{\mathrm{i}},$$

we will also investigate the contribution of each trace element by considering the smallest eigenvalue of $DF_{\mathrm{i}}^T DF_{\mathrm{i}}$. Note that our approach is based on the study of zircon crystals that were formed in conditions where the crystal growth occurs faster than the diffusion time scale in the melt. We will also study the influence of the total cooling time T on the smallest eigenvalue.

In practice, we consider a linear ramp where the temperature decreases from 1113 K to 993 K in a duration T. The initial concentrations for Hf/Y/U/Th are respectively 1400/100/50/20, according to the values in our experimental sample (see Section 6.5). We consider then a space of 9 parameters, corresponding to a piecewise linear temperature evolution where the time interval is subdivided into 8 subintervals. The sensitivity matrices DF_{i} are computed together with their smallest eigenvalues, as well as the smallest eigenvalue of $DF^T DF$. We also consider the case when only Hf and Y are observed. The results are presented in Figure 6.1.

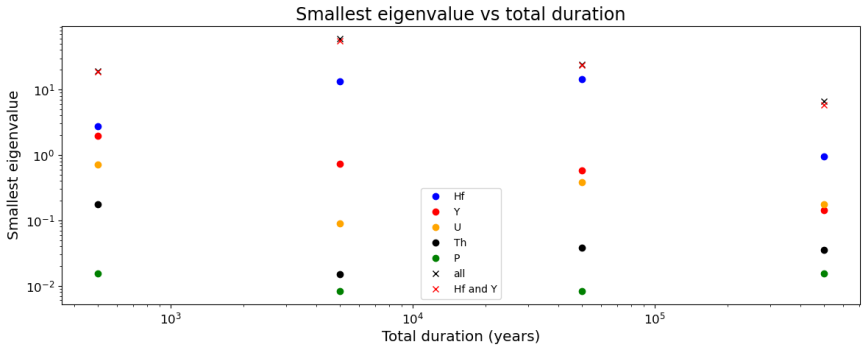

Figure 6.1. Smallest singular value of the sensitivity matrix for different total duration T (log-log scale). The smallest singular value of the observation of each trace element is displayed, as well as for all trace elements, and only for Hf and Y. The different values of T are: 500, 5000, 50000 and 500000 years.

Figure 6.1 shows that when T becomes large, the smallest singular value of the sensitivity matrix decreases. This fact illustrates that the system reaches diffusive equilibrium, and the influence of diffusion on trace element partition to zircon becomes negligible. A second observation is that the two elements Hf and Y provide most of the information.

6.2. **Practical Implementation**

The estimation of a temperature profile on the interval $[0, T]$ described by a few (between 2 and a few 10s) parameters uses a coarse-to-fine method. First the initial temperature and total cooling time are determined using Algorithm 1. Then Algorithm 2 is used to find a more detailed profile. A dyadic refinement of the search space is performed every 4 iterations: the temperature profile is piecewise linear on an increasing number of intervals: 1, then 2, then 4, then 8... The total number of intervals can not be too large, since measurements error impair the estimation. In practice we stopped at 8, 16 or 32 intervals.

The simulations presented here were run with $N_x = 500$ space points and $N = 500$ time points, excepted for Test case 3 where $N_x = 200$ and $N = 200$.

6.3. **Test case 1: temperature profile estimation**

We first present a test case where the temperature profile on a time interval $[0, T]$ is to be determined. We assume that the total cooling time T is known, and only Algorithm 2 is used with dyadic refinement every 4 iterations.

The reference evolution is a linear decrease of the temperature from 1113 K to 993 K, perturbed by 3 oscillations, see Figure 6.2. Knowing this reference temperature evolution, we compute the concentrations of trace elements in the crystal, namely: Hf, Y, U, Th, Sm, Dy, Yb. We then add a gaussian noise to these concentrations to generate the observations. We present in Figure 6.3 the concentration profiles, that are the concentrations of the trace elements, each concentration is normalized by its maximal value.

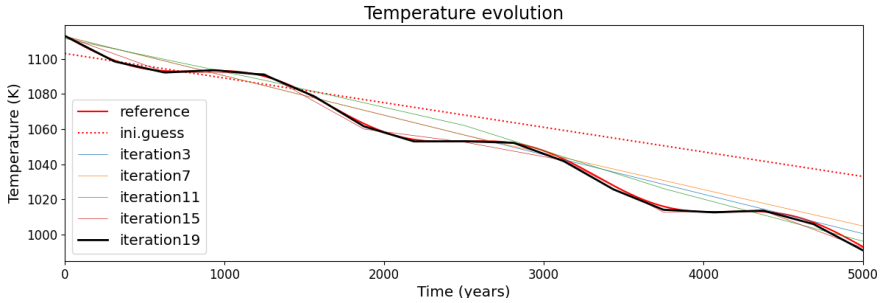

FIGURE 6.2. Test case 1: estimated temperature evolution (fine resolution), along the iterations of the algorithm.

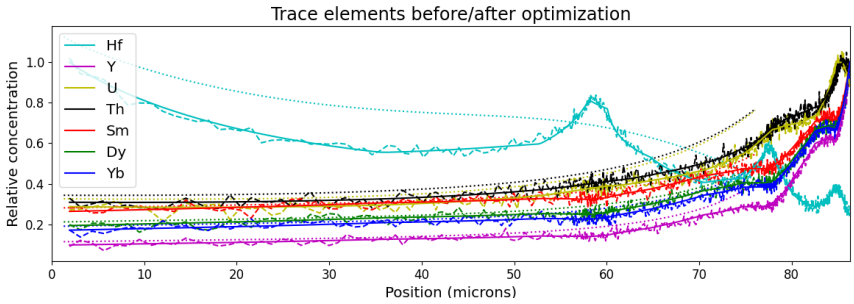

FIGURE 6.3. Test case 1: observations (dashed), initial guess (dotted) and estimations (continuous) for the concentration profiles (fine resolution).

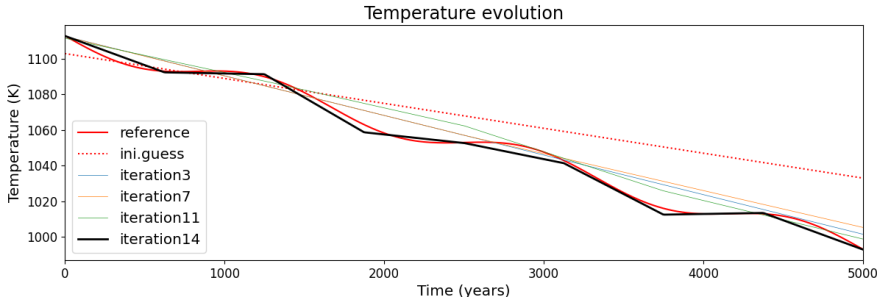

FIGURE 6.4. Test case 1: estimated temperature evolution (coarse resolution), along the iterations of the algorithm.

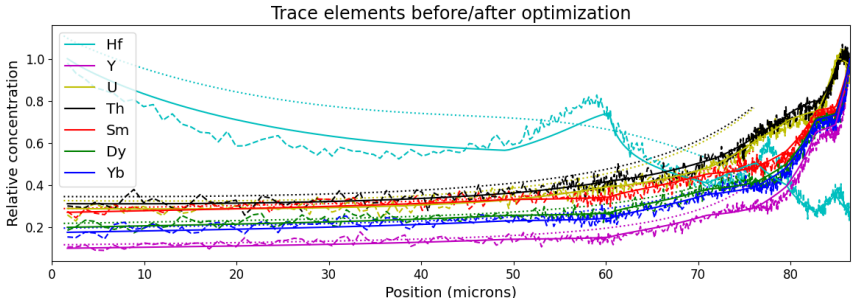

FIGURE 6.5. Test case 1: observations (dashed), initial guess (dotted) and estimations (continuous) for the concentration profiles (coarse resolution).

The reconstruction of the temperature evolution follows the coarse-to-fine dyadic refinement described above up to 16 intervals (fine resolution) or 8 intervals (coarse resolution). As initial guess we have on purpose chosen a linear ramp between temperature values that are significantly different from the true initial and final temperatures. The final estimated temperature profile, and the predicted concentration profiles are presented on Figures 6.2, 6.3, 6.4 and 6.5. This test case documents the ability of the algorithm to reconstruct a temperature pattern that comprises a few oscillations. The total computation time is 8.6 min for the fine resolution (16 time intervals, 17 control points) where we used 20 iterations. The computation time is 3.8 min for the coarse resolution (8 time intervals, 9 control points), where we used 15 iterations.

6.4. **Test case 2: total cooling time and temperature profile estimation**

In the second test case, we first optimize w.r.t. the total cooling time, using the sensitivity analysis described in Section 4.3 and Algorithm 1. Then the time interval is fixed and we search for a profile using dyadic refinement, where at each scale we use Algorithm 2.

The synthetic data are generated as follows. We generate trace elements concentration with a decaying and oscillating temperature profile. These concentration are then perturbed by additive gaussian noise, see Figure 6.8.

For the inverse problem solution, the initial guess for the total cooling time is on purpose chosen far from the true value (60 000 years vs 5000 years) in order to document the abilities of the proposed method. The total cooling time estimation requires 10 iterations (14 s), the resulting temperature evolution is denoted "best ramp", see Figure 6.7. The reconstruction using the coarse-to-fine approach is stopped when the number of intervals is 32 (6.2 min). The results are presented on Figures 6.6–6.7 (temperature evolution) and 6.8–6.9 (trace elements concentration). This test case documents the ability of the algorithm to recontruct the time scale of the temperature decrease, even with a far-off initial guess.

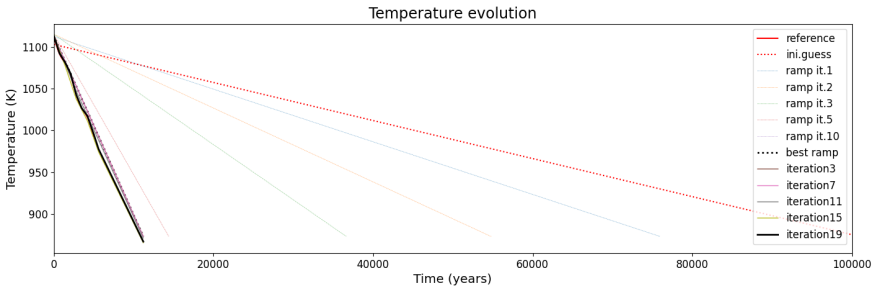

Figure 6.6. Test case 2: estimated temperature evolution along the iterations of the algorithm.

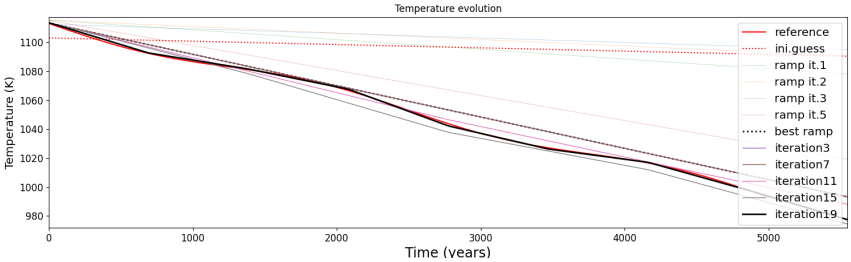

Figure 6.7. Test case 2: estimated temperature evolution along the iterations of the algorithm (zoom).

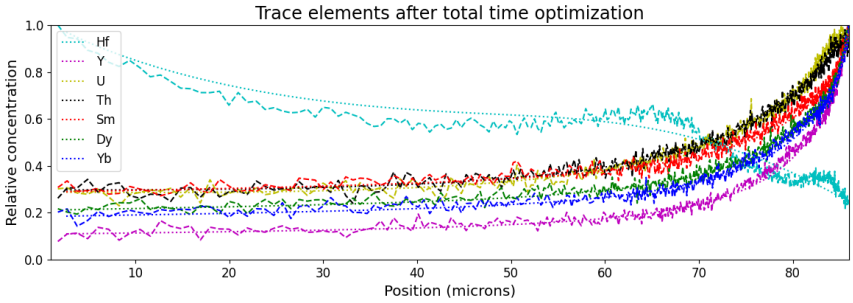

Figure 6.8. Test case 2: Observed (dashed) and estimated (continuous) concentration profiles after the first step (Algorithm 1).

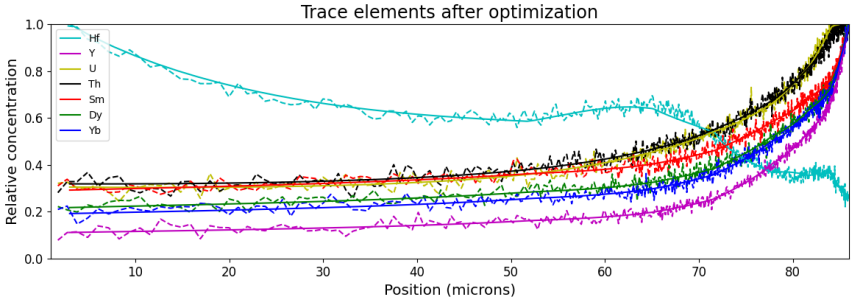

Figure 6.9. Test case 2: Observed (dashed) and estimated (continuous) concentration profiles after the second step (Algorithm 2).

6.5. **Test case 3: reconstruction from experimental data**

The third test case consists of experimental data that were presented in [13] for the zircon crystal from early Fish Canyon Tuff. In this dataset the errors in the concentration measurement are considered to be gaussian, with a standard deviation for Hf/Y/U/Th of respectively 52/59/59/68 ppm. These standard deviations are used to weight the different terms in the cost function J, see (5.1).

In order to estimate the temperature history, we follow the methodology from the second test case, namely Algorithm 1 is used in a first step to estimate the best linear decrease (best ramp). The initial guess is a decrease from 1088 K to 1018 K in 4000 years. The optimal linear decrease obtained is a decrease from 1099 K to 1018 K in 2600 years. The covariance matrix of the error on the retrieved parameters is obtained from Equation (6.1). It indicates that the standard deviation for the initial temperature is 2.1 K and for the final temperature it is 25.4 K, which amounts to a standard deviation for the cooling time of 830 years (using Equation (4.3)). These standard deviations are indicated in Figure 6.10. Note that the latter is a non negligible fraction of the total cooling time.

In a second step, we apply Algorithm 2 on subspaces with growing dimension, obtained by successive dyadic refinement. The initial guess on the first subspace is the best ramp obtained in the first step. A number of 4 iterations are performed on each resolution, therefore the number of iterations is related to the number of iterations: 8 iterations correspond to a subdivision of the time interval into 4 subintervals (5 control points), 12 iterations correspond to 9 control points and 16 iterations correspond to 17 control points. The results of the temperature history at different resolutions is presented in Figure 6.10, and the comparison between the predictions and the observations are presented in Figure 6.11. This test case documents the ability of the algorithm to process experimental data, although the error in the measurements lead to relatively large uncertainties in the retrieved temperature evolution.

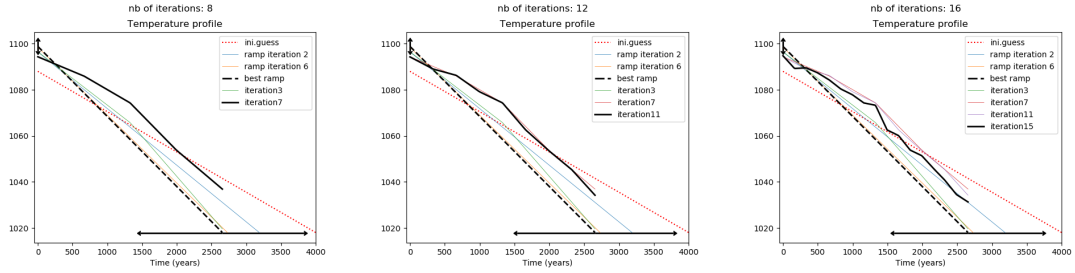

Figure 6.10. Test case 3: temperature history at several resolutions. Standard deviation for the best ramp are indicated with error bars

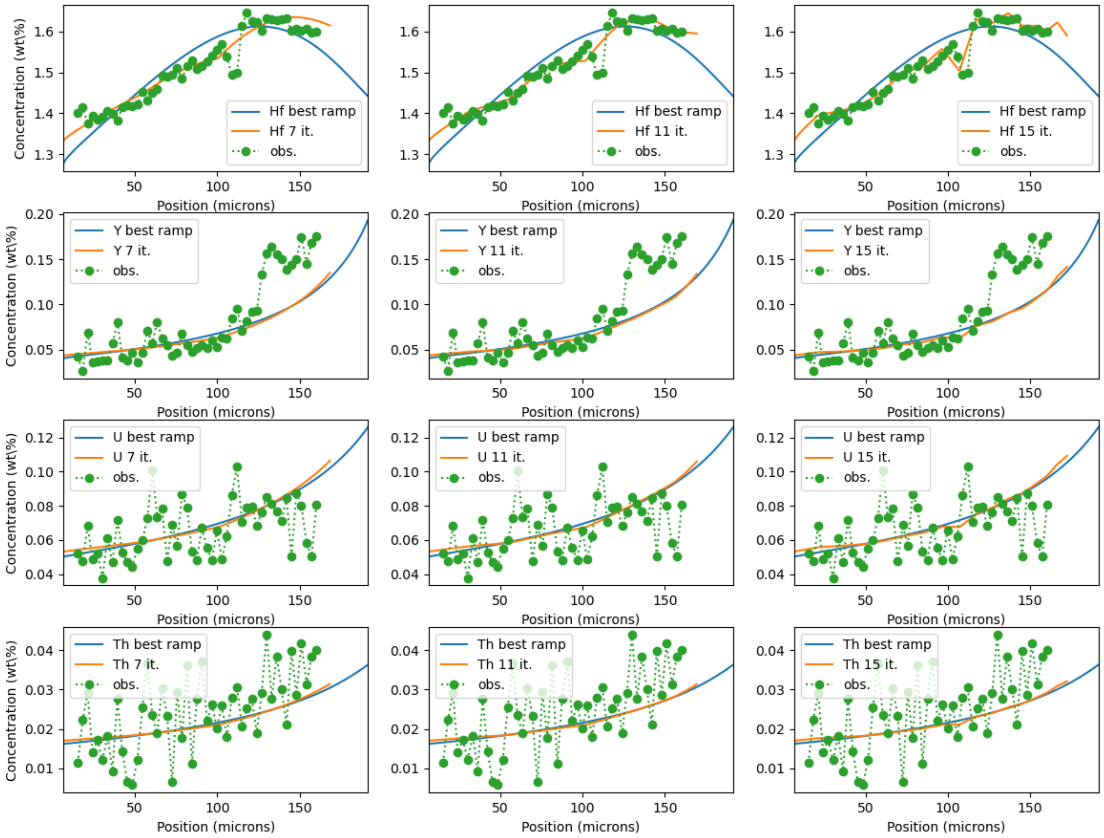

Figure 6.11. Test case 3: trace element concentration at several resolutions (left column: 5 control points, center: 9 control points, right: 17 control points): observations vs predictions of the model. From top to bottom row: Hf, Y, U, Th. The best ramp obtained after Algorithm 1 is also indicated in blue.

7. Summary and Discussion

In the present work, we have proposed a variational framework to reconstruct temperature evolution from the observation of trace elements concentrations in zircon crystals. This required to compute the tangent model of a one-phase Stefan problem and the analytical derivation of a nonlinear observation operator. The numerical experiments showed a reasonable ability to

reconstruct the temperature evolution in a small dimensional subspace (piecewise linear evolution), without the need of incorporating regularization terms or other a-priori information. The quadratic minimization was performed using a Gauss–Newton method and convergence is achieved within a small number of iterations. We now discuss the contribution of this work for the geological community.

Test case 1 based on synthetic data shows the ability of the algorithm to reconstruct a complicated temperature history that mimics realistic thermal conditions of zircon crystallization in an incrementally building magmatic intrusion [14] by periodic injection of bathes of hot magma. There are several limitations of the method that naturally come from the physics of zircon growth in magmas. First, the temperature history can be evaluated only above the solid state of the system (solidus) when crystallization stops. Second, thermal information can be lost if significant temperature increase leads to partial zircon dissolution or diffusional re-equilibration at high temperature. Third, if crystallization occurs on timescales much larger than timescales of trace element diffusion inside a melt cell around the growing zircon, trace element incorporation is only controlled by a partition coefficient. In the later case the temperature variation can be reconstructed directly from the mass balance of a trace element within the cell but no timescales can be retrieved.

The algorithm also shows good performance on natural samples. Note that the model does not fit perfectly the behavior of both Hf and Y, which are the principal trace elements in the considered crystal. This may come from imperfection in the data collection or in the model and further studies in controlled experiments should provide hints towards an improvement. The temperature range estimated by the best fit model is in a good agreement with existing numerical modelling [4].

Model calibration for a particular magmatic system requires an experimental determination of partition coefficients for trace elements between zircon and melts, and zirconium saturation in the melts as a function of temperature for the magma composition and temperature ranges typical for the given magmatic system. Additionally, diffusion coefficients for the trace elements must be measured for the melt at the representative temperatures. But, even for generally accepted values the model developed here gives good starting estimates for the thermal history that was experienced by the growing zircon crystal.

Future in situ investigation of magmatic reservoirs through deep drilling programs such as the Krafla Magma Testbed program in Northern Iceland is aimed at better constraining the kinetic effects of magmatic processes. These effects include the crystal growth and dissolution, magma mixing, and melt segregation necessary to form and evolve the continental crust. Planned future experiments on the kinetics of zircon growth under controlled conditions aim to understand how the rate of zircon crystallization affects trace element partitioning.

Acknowledgments

The authors thank the anonymous referees, whose relevant remarks helped improve the content and the presentation of this work.

References

[1] `https://perso.math.univ-toulouse.fr/fehren/software/`.

[2] `https://zircocalc.math.univ-toulouse.fr/`.

[3] Ilya Bindeman and Oleg Melnik. Zircon survival, rebirth and recycling during crustal melting, magma crystallization, and mixing based on numerical modelling. *J. Petrol.*, 57(3):437–460, 2016.

[4] Karoline Brückel, Craig Lundstrom, Mickael Ackerson, and Christopher Campe. Testing the Limits of Ti-in-Quartz Thermometry and Diffusion Modelling to Determine the Thermal History of the Fish Canyon Tuff. *J. Petrol.*, 64(12): article no. egad082, 2023.

[5] Abdellatif El Badia and F. Moutazaim. A one-phase inverse Stefan problem. *Inverse Probl.*, 15(6):1507–1522, 1999.

[6] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications (Dordrecht)*. Kluwer Academic Publishers, 1996.

[7] Nataliya Gol'dman. *Inverse Stefan Problems*, volume 412 of *Mathematics and its Applications (Dordrecht)*. Springer, 2012.

[8] Eugenia Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, 2003.

[9] Andreas Kirsch. *An introduction to the mathematical theory of inverse problems*, volume 120 of *Applied Mathematical Sciences*. Springer, 2011.

[10] François-Xavier Le Dimet and Olivier Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2):97–110, 1986.

[11] Jacques-Louis Lions. *Optimal control of systems governed by partial differential equations*, volume 170 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 1971.

[12] Dong Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Math. Program., Ser. B*, 45(3):503–528, 1989.

[13] Oleg Melnik and Ilya Bindeman. Modeling of trace elemental zoning patterns in accessory minerals with emphasis on the origin of micrometer-scale oscillatory zoning in zircon. *Am. Mineral.*, 103(3):355–368, 2018.

[14] Oleg Melnik, Ivan Utkin, and Ilya Bindeman. Magma Chamber Formation by Dike Accretion and Crustal Melting: 2D Thermo-Compositional Model With Emphasis on Eruptions and Implication for Zircon Records. *J. Geophys. Res. Solid Earth*, 126(12): article no. e2021JB023008, 2021.

[15] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Series in Operations Research. Springer, 1999.

[16] Rembert Reemtsen and Andreas Kirsch. A method for the numerical solution of the one-dimensional inverse Stefan problem. *Numer. Math.*, 45:253–273, 1984.

[17] Daniela Rubatto and Jörg Hermann. Experimental zircon/melt and zircon/garnet trace element partitioning and implications for the geochronology of crustal rocks. *Chem. Geol.*, 241(1):38–61, 2007.

[18] Olivier Talagrand. Assimilation of observations, an introduction (special issue data assimilation in meteology and oceanography: Theory and practice). *J. Meteorol. Soc. Jpn. Ser. II*, 75(1B):191–209, 1997.

[19] Sifan Wang and Paris Perdikaris. Deep learning of free boundary and Stefan problems. *J. Comput. Phys.*, 428: article no. 109914 (24 pages), 2021.