HAL
open science

Optimal solution error quantification in variational data assimilation involving imperfect models

Victor Shutyaev, Igor Gejadze, Arthur Vidard, Francois-Xavier Le Dimet

To cite this version:

Victor Shutyaev, Igor Gejadze, Arthur Vidard, Francois-Xavier Le Dimet. Optimal solution error quantification in variational data assimilation involving imperfect models. International Journal for Numerical Methods in Fluids, 2017, 83 (3), pp.276-290. 10.1002/fld.4266 . hal-01411666

HAL Id: hal-01411666

https://inria.hal.science/hal-01411666v1

Submitted on 16 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

Optimal solution error quantification in variational data assimilation involving imperfect models

V. SHUTYAEV[1], I. GEJADZE[2], A. VIDARD[3], F.-X. LE DIMET[4]

Abstract. The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find the initial condition. If the model is 'perfect', the optimal solution (analysis) error rises due to the presence of the input data errors (background and observation errors). Then, this error is quantified by the covariance matrix, which can be approximated by the inverse Hessian of an auxiliary control problem. If the model is not perfect, the optimal solution error includes an additional component due to the presence of the model error. In this paper we study the influence of the model error on the optimal solution error covariance, considering strong and weak constraint data assimilation approaches. For the latter, an additional equation describing the model error dynamics is involved. Numerical experiments for the 1D Burgers equation illustrate the presented theory.

1 Introduction

Variational data assimilation (DA) is a preferred method used for weather and ocean forecasting in major operational centers around the globe. The method is usually understood in the framework of the optimal control theory, but it also represents a special case of the maximum a posteriory probability (MAP) estimator. No intrinsic uncertainty estimation and propagation procedures are involved with this method, however it is important to make such tools available for statistical post-processing of results. A *perfect model* assumption leads to the *strong constraint* formulation of variational DA. In this case (and within the Gaussian framework) the optimal solution error can be quantified by the covariance matrix, which is approximated by the inverse Hessian of an auxiliary control problem [1, 2]. The perfect model assumption, however, is not valid in many realistic applications. In this case the *weak constraint* formulation [3, 4] should be preferred. In this formulation the model error is regarded as a forcing term subjected to control along with the initial condition. A rigorous justification of the weak constraint formulation from the Bayesian perspective is given in [5]. The effect of the model error on the optimal solution error for the strong constraint formulation is analyzed in [6, 7]. In a special case when the model error is due to discretization of model equations (namely, the advection equation), a rigorous error analysis can be found in [8, 9]. In this paper we investigate the optimal solution error covariance in case of using imperfect models, both for the strong and weak constraint DA formulations. Below we present some introductory materials on variational DA and use these materials for substantive literature review.

[1]Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow Institute for Physics and Technology, 119333 Gubkina 8, Moscow, Russia; e-mail: shutyaev@inm.ras.ru
[2]UMR G-EAU, IRSTEA-Montpellier, 361 Rue J.F. Breton, BP 5095, 34196, Montpellier, France; e-mail: igor.gejadze@irstea.fr
[3]LJK, Université Grenoble Alpes, BP 51, 38051 Grenoble Cedex 9, France; e-mail: arthur.vidard@imag.fr
[4]LJK, Université Grenoble Alpes, BP 51, 38051 Grenoble Cedex 9, France; e-mail: ledimet@imag.fr

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

1.1 Standard strong constraint variational DA

Let us consider a dynamical system which is described by its *state vector* $\mathbf{x} \in \mathcal{X}$, where $\mathcal{X}$ is the *state space*. This vector includes different physical variables, some of them are spatially distributed vector or scalar fields (e.g. velocity, temperature, elevation of the free surface, salinity, concentration of biological or chemical species, etc.). The evolution of $\mathbf{x}$ is governed by a system of differential equations:

$$\begin{cases} \dfrac{d\mathbf{x}(t)}{dt} &= M(\mathbf{x}(t)), \quad t \in (0,T], \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{cases} \tag{1}$$

where $M : \mathcal{X} \to \mathcal{X}$ is a model operator. We suppose that for a given $\mathbf{x}_0$ the problem (1) has a unique solution $\mathbf{x}(t) \in \mathcal{Y} = L_2[0,T;\mathcal{X}]$, which is called the *trajectory*. The perfect model assumption means that the model (1) simulates the natural phenomena under consideration perfectly.

Let $x_0^b \in \mathcal{X}$ be the best available guess (prior, background) on $\mathbf{x}_0$. We aim to improve the prior by assimilating incomplete observations of the trajectory $\mathbf{y}(t) = H(x(t)) \in \mathcal{O}$, $t \in [0,T]$, where $\mathcal{O}$ is the *observation space* and $H : \mathcal{Y} \to \mathcal{O}$ is the *observation operator*. Under the Gaussian assumption on the observation and the background error probability density functions (pdf), maximizing the posterior pdf is equivalent to minimizing the following cost-function

$$J(\mathbf{x}_0) = \frac{1}{2}\int_0^T \left\| R^{-1/2}(\mathbf{y}(t) - H(\mathbf{x}(\mathbf{x}_0,t))) \right\|^2 dt \; + \frac{1}{2}\left\| B^{-1/2}(\mathbf{x}_0 - \mathbf{x}_0^b) \right\|^2, \tag{2}$$

where R and B are the observation and the background error covariances, respectively. The second term in (2) can also be regarded as a regularization term in the Tikhonov sense [10]. Thus, the variational DA problem can be written as follows:

$$\mathbf{x}_0^a = \text{argmin } J(\mathbf{x}_0). \tag{3}$$

The necessary condition for $\mathbf{x}_0^a$ to be a solution of (3) is

$$\nabla J(\mathbf{x}_0^a) = 0, \tag{4}$$

where ∇J is the gradient of $J(\mathbf{x}_0)$ with respect to the control variable $\mathbf{x}_0$. This condition also becomes a sufficient condition if J is convex. The gradient ∇J can be expressed in the form

$$\nabla J = \mathbf{x}^*(0) + B^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b), \tag{5}$$

where $\mathbf{x}^*(t)$ is the *adjoint variable* which satisfies the *adjoint model*

$$\begin{cases} -\dfrac{d\mathbf{x}^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \mathbf{x}^* = \left[\dfrac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}(H(\mathbf{x}) - \mathbf{y}), \quad t \in [0,T), \\ \mathbf{x}^*(T) = 0. \end{cases} \tag{6}$$

The derivation of the system (6) can be found in [11]. The model (1), the adjoint model (6) along with equations (5) and (4) form the *optimality system* (OS). Let us note that for different (non-Gaussian) assumptions on the observation and the background error pdfs, different forms of the cost-function (2) can be obtained.

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

1.2 Weak constraint formulation

In reality, perfect models are rarely available. For example, the model error may appear due to unresolved scales, unaccounted physical processes, crude discretization of model equations and/or misspecification of the model parameters. For a dynamical system in the form (1) the model error can be defined as

$$\boldsymbol{\eta}(t) = \mathcal{M}(\mathbf{x}(t)) - M(\mathbf{x}(t)), \tag{7}$$

where $\mathcal{M}(\cdot)$ is a hypothetical perfect model operator and $M(\cdot)$ is the available inexact model operator. Should we know $\boldsymbol{\eta}(t)$, the perfect model can be defined using the inexact operator $M(\cdot)$ as follows:

$$\begin{cases} \dfrac{d\mathbf{x}(t)}{dt} &= M(\mathbf{x}(t)) + \boldsymbol{\eta}(t), \quad t \in (0, T], \\ \mathbf{x}(0) &= \mathbf{x}_0. \end{cases} \tag{8}$$

Since $\boldsymbol{\eta}(t)$ is actually unknown, we aim to estimate its value along with the initial condition $\mathbf{x}_0$. For example, in conventional strong constraint variational DA involving the global-scale atmospheric and ocean models in the form (1) the initial state of the system is considered as control. However, because the actual model is in the form (8), the initial state control remains useful within a limited time period, whereas beyond this period $\boldsymbol{\eta}(t)$ becomes a dominant factor influencing the system behavior.

Thus, the essence of the weak constraint variational DA approach is that the model error $\boldsymbol{\eta}(t)$ is considered as additional control. Under the Gaussian assumption on the model error pdf we consider the following cost-function:

$$J(\mathbf{x}_0, \boldsymbol{\eta}(t)) = \frac{1}{2}\int_0^T \left\| R^{-1/2}(\mathbf{y}(t) - H(\mathbf{x}(t))) \right\|^2 dt \; + \frac{1}{2}\left\| B^{-1/2}(\mathbf{x}_0 - \mathbf{x}_0^b) \right\|^2 +$$

$$+ \frac{1}{2}\int_0^T \int_0^T \boldsymbol{\eta}^T(t')\, Q^{-1}(t', t)\, \boldsymbol{\eta}(t)\, dt'\, dt, \tag{9}$$

where Q is the model error covariance. The weak constraint variational DA problem can be written as a minimization problem involving the above cost-function, i.e.

$$(\mathbf{x}_0^a, \boldsymbol{\eta}^a(t)) = \text{argmin } J(\mathbf{x}_0, \boldsymbol{\eta}(t)), \; t \in (0, T]. \tag{10}$$

Note that under different assumption on the model error pdf a different form of the corresponding term in (9) can be obtained (see e.g. [12]).

One major issue of the weak constraint variational DA method is how to define Q. The simplest approach whereby Q is taken to be proportional to the background error covariance, i.e. $Q = \alpha B$ is considered in [13]. The tendency-based model error covariance, which is obtained as a by-product of running the 4D-Var ensemble is described in [14]. If the model error is due to misspecification of model parameters, Q is constructed based on the Jacobian of M with respect to these parameters and by using a short time approximation of the resulting covariance operator [15]. An approach dealing specifically with the model error due to discretization is presented in [16].

The major technical difficulty is related to the size of the extended control vector $(\mathbf{x}_0, \boldsymbol{\eta}(t))$, which is equal to the dimension of the state vector multiplied by the number of time integration steps (after time discretization). To circumvent this problem several approaches have been proposed. The

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

simplest one is, of course, to assume that the model error is constant in time [6, 14]. A spectral form of the model error

$$\boldsymbol{\eta}(t) = \boldsymbol{\eta}_0 + \boldsymbol{\eta}_1 \sin(2\pi t/\tau) + \boldsymbol{\eta}_2 \cos(2\pi t/\tau),$$

where τ is the time scale on which the model error is expected to vary and $\boldsymbol{\eta}_{0,1,2}$ are spatially distributed values, is suggested in [17]. Another possibility is to use a separate dynamical model for the model error evolution, in which case one must control its initial condition only [18, 17, 19, 20]. This evolution model reads as follows:

$$\left\{ \begin{array}{rcl} \dfrac{d\boldsymbol{\eta}(t)}{dt} & = & N(\mathbf{x}(t))\boldsymbol{\eta}(t) + \epsilon(t), \quad t \in (0,T], \\ \boldsymbol{\eta}(0) & = & \boldsymbol{\eta}_0, \end{array} \right. \tag{11}$$

where N is an operator dependent on $\mathbf{x}$, ϵ is an unbiased stochastic noise (which can possibly be controlled but is neglected in what follows). We consider the operator N to be given in advance, a choice of N is discussed, for example, in [20]. Then, the cost function becomes

$$J(\mathbf{x}_0, \boldsymbol{\eta}_0) = \frac{1}{2}\int_0^T \left\| R^{-1/2}(\mathbf{y}(t) - H(\mathbf{x}(\mathbf{x}_0, t))) \right\|^2 dt \; + \frac{1}{2}\left\| B^{-1/2}(\mathbf{x}_0 - \mathbf{x}_0^b) \right\|^2 + \frac{1}{2}\left\| Q^{-1/2}\boldsymbol{\eta}_0 \right\|^2. \tag{12}$$

This approach (called sometimes 'control of systematic error') has been successfully employed in pre-operational conditions (realistic model, actual observations) for the ocean [21]. It is shown in [19], on a simple test case, that the model error estimate could be used to correct the drift of the subsequent forecast.

A different option for reducing the size of the extended control vector is to use a projection S onto a reduced basis as proposed in [22, 23, 19, 24]. In that case, equation (8) reads

$$\left\{ \begin{array}{rcl} \dfrac{d\mathbf{x}(t)}{dt} & = & M(\mathbf{x}(t)) + S^*\boldsymbol{\eta}_r(t), \quad t \in (0,T], \\ \mathbf{x}(0) & = & \mathbf{x}_0, \end{array} \right. \tag{13}$$

with $\boldsymbol{\eta}_r(t)$ being defined in a low-dimensional space and S^* being the adjoint to S. One can consider an approach combining both above-mentioned strategies, as illustrated in [22] on a simple test case.

Another approach is suggested in [25]. Let us discretize the time domain $[0,T]$ into a set of time discretization nodes $\{t_i\}$, $i = 1, \ldots, I$ and consider an ordered subset of this set $\{t_{k(j)}\}$, $j = 1, \ldots, J$. Then, according to [25] the cost-function (9) is modified as follows:

$$J(\mathbf{x}(t)) = \frac{1}{2}\sum_{i=1}^{I} \left\| R_i^{-1/2}(\mathbf{y}(t_i) - H_i(\mathbf{x}(t_i))) \right\|^2 + \frac{1}{2}\left\| B^{-1/2}(\mathbf{x}_0 - \mathbf{x}_0^b) \right\|^2 +$$

$$+\frac{1}{2}\sum_{j=1}^{J} \left\| Q_{i(j)}^{-1/2}(\mathcal{M}_{k(j)}\mathbf{x}(t_{k(j-1)}) - \mathbf{x}(t_{k(j)})) \right\|^2. \tag{14}$$

Next, it is proposed to control a subset of snapshots $\mathbf{x}(t_{k(j)})$, $j = 1, \ldots, J$ by penalizing the deviation between the current snapshot $\mathbf{x}(t_{k(j)})$ and the model state integrated from the previous controlled snapshot $\mathbf{x}(t_{k(j-1)})$. By limiting the number of the controlled snapshots the computational cost is reduced to a practically acceptable level. Moreover, the computation of the gradient is

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

independent for each time sub window $t \in [t_{k(j-1)}, t_{k(j)}]$ and, therefore, can be performed in parallel. This approach is similar to the *moving horizon estimator* with *direct multiple shoot*, well-known in the automatic control community and described in [26], for example.

The control of model error or weak constraint variational DA has some advantages. One of them is that it allows for larger assimilation windows [21]. Indeed, if the model drifts, it becomes difficult to fit all observations by adjusting the initial condition only. Another aspect is that it allows for a greater parallelism. This aspect is further enhanced by the use of saddle-point methods as proposed by [27] and further improved by [28] where the 4DVar method and the dual method (4D-PSAS) are combined.

In this paper we study the optimal solution error covariance for the strong constraint formulation, Subsect. 2.1 and 2.2 (without and with the model error, correspondingly), and for the weak constraint formulation, Subsect. 2.3. The details of numerical implementation for the 1D Burgers' equation are given in Sect. 3, and Sect. 4 presents the results of numerical experiments. The paper ends by the Conclusions.

2 Optimal solution error covariances

2.1 Standard strong constraint variational DA

The content of this subsection replicates the derivations presented in [2]. Here we accept the 'perfect model' assumption, i.e. $M = \mathcal{M}$.

Let $\mathbf{x}_0^t$ be the true initial state. Then, the true trajectory $\mathbf{x}^t(t),\ t \in (0, T]$ is given by the model

$$\left\{ \begin{array}{rcl} \dfrac{d\mathbf{x}^t(t)}{dt} & = & M(\mathbf{x}^t(t)), \quad t \in (0, T] \\ \mathbf{x}^t(0) & = & \mathbf{x}_0^t. \end{array} \right. \tag{15}$$

Let us define the initial state error $\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^t$ and the state error along the trajectory $\delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}^t(t)$. Then, for small errors, the state error evolution is given by the tangent linear model

$$\left\{ \begin{array}{rcl} \dfrac{d\delta\mathbf{x}(t)}{dt} & = & \dfrac{dM(\mathbf{x}^t)}{d\mathbf{x}}\delta\mathbf{x}(t), \quad t \in (0, T] \\ \delta\mathbf{x}(0) & = & \delta\mathbf{x}_0. \end{array} \right. \tag{16}$$

We define the input data as follows: the background function $\mathbf{x}_0^b \in \mathcal{X}$, $\mathbf{x}_0^b = \mathbf{x}_0^t + \boldsymbol{\xi}_x$ and the observations $\mathbf{y} \in \mathcal{O}$, $\mathbf{y} = H(\mathbf{x}^t) + \boldsymbol{\xi}_y$, where $\boldsymbol{\xi}_x \sim \mathcal{N}(0, \mathbf{B})$ and $\boldsymbol{\xi}_y \sim \mathcal{N}(0, \mathbf{R})$. Taking into account the latter definitions we note that

$$H(\mathbf{x}) - \mathbf{y} = H(\mathbf{x}) - H(\mathbf{x}^t) - \boldsymbol{\xi}_y \approx \frac{dH(\mathbf{x}^t)}{d\mathbf{x}}\delta\mathbf{x} - \boldsymbol{\xi}_y \tag{17}$$

$$\mathbf{x}_0 - \mathbf{x}_0^b = \mathbf{x}_0 - \mathbf{x}_0^t - \boldsymbol{\xi}_x = \delta\mathbf{x}_0 - \boldsymbol{\xi}_x. \tag{18}$$

Next, we substitute (17) into the adjoint model (6) and (18) into the expression for the gradient (5) and obtain the following optimality system for errors:

$$\left\{ \begin{array}{rcl} \dfrac{d\delta\mathbf{x}(t)}{dt} & = & \dfrac{dM(\mathbf{x}^t)}{d\mathbf{x}}\delta\mathbf{x}(t), \quad t \in (0, T] \\ \delta\mathbf{x}(0) & = & \delta\mathbf{x}_0, \end{array} \right. \tag{19a}$$

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

$$\begin{cases} -\dfrac{d\mathbf{x}^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \mathbf{x}^* = \left[\dfrac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\left(\dfrac{dH(\mathbf{x}^t)}{d\mathbf{x}}\delta\mathbf{x} - \boldsymbol{\xi}_y\right), \quad t \in [0,T), \\ \mathbf{x}^*(T) = 0, \end{cases} \tag{19b}$$

$$\mathbf{x}^*(0) + B^{-1}(\delta\mathbf{x}_0 - \boldsymbol{\xi}_x) = 0. \tag{19c}$$

Our goal now is to express explicitly the optimal solution error $\delta\mathbf{x}_0$ via the data errors $\boldsymbol{\xi}_y$ and $\boldsymbol{\xi}_x$ from the system (19). Having this in mind, we introduce an operator $\mathbf{K} : \mathcal{X} \to \mathcal{Y}$

$$\mathbf{K}(\mathbf{x})v = \varphi(t), \tag{20}$$

such that $v \in \mathcal{X}$ and $\varphi(t) \in \mathcal{Y}$ are related via the tangent linear model

$$\begin{cases} \dfrac{d\varphi(t)}{dt} &= \dfrac{dM(\mathbf{x})}{d\mathbf{x}}\varphi(t), \quad t \in (0,T] \\ \varphi(0) &= v. \end{cases}$$

The corresponding adjoint operator $\mathbf{K}^* : \mathcal{Y} \to \mathcal{X}$

$$\mathbf{K}^*(\mathbf{x})\varphi(t) = v, \tag{21}$$

satisfies the following adjoint model:

$$\begin{cases} -\dfrac{d\varphi^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \varphi^* = \varphi(t), \quad t \in [0,T), \\ \varphi^*(T) = 0, \\ v = \varphi^*(0). \end{cases}$$

Taking into account the above definition for $\mathbf{K}^*$, from the adjoint equation (19b) we obtain:

$$\mathbf{x}^*(0) = \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\frac{dH(\mathbf{x}^t)}{d\mathbf{x}}\,\delta\mathbf{x}(t) - \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\boldsymbol{\xi}_y. \tag{22}$$

Taking into account the definition for $\mathbf{K}$ we can write

$$\delta\mathbf{x}(t) = \mathbf{K}(\mathbf{x}^t)\,\delta\mathbf{x}_0 \tag{23}$$

and, subsequently, the expression (22) becomes

$$\mathbf{x}^*(0) = \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\frac{dH(\mathbf{x}^t)}{d\mathbf{x}}\,\mathbf{K}(\mathbf{x}^t)\,\delta\mathbf{x}_0 - \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\boldsymbol{\xi}_y.$$

Assuming $\mathbf{x} = \mathbf{x}^t$ we substitute the above expression into the optimality condition (19c) and, finally, obtain the error equation

$$\mathcal{H}(\mathbf{x}^t)\delta\mathbf{x}_0 = \boldsymbol{\xi}(\mathbf{x}^t), \tag{24}$$

where

$$\mathcal{H}(\mathbf{x}) = B^{-1} + \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\frac{dH(\mathbf{x})}{d\mathbf{x}}\mathbf{K}(\mathbf{x}) \tag{25}$$

and

$$\boldsymbol{\xi}(\mathbf{x}) = B^{-1}\boldsymbol{\xi}_x + \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\boldsymbol{\xi}_y. \tag{26}$$

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

Next, we express $\delta\mathbf{x}_0$ from (24) and compute its covariance matrix $P = E[\delta\mathbf{x}_0\delta\mathbf{x}_0^T]$. It can be easily seen that if the data errors are uncorrelated, i.e. $E[\boldsymbol{\xi}_x\boldsymbol{\xi}_y^T] = 0$, then

$$E[\boldsymbol{\xi}(\mathbf{x}^t)\boldsymbol{\xi}^T(\mathbf{x}^t)] = \mathcal{H}(\mathbf{x}^t),$$

and, therefore

$$P = E[\mathcal{H}^{-1}\boldsymbol{\xi}\boldsymbol{\xi}^T\mathcal{H}^{-1}] = \mathcal{H}^{-1}E[\boldsymbol{\xi}\boldsymbol{\xi}^T]\mathcal{H}^{-1} = \mathcal{H}^{-1}\mathcal{H}\mathcal{H}^{-1} = \mathcal{H}^{-1}(\mathbf{x}^t). \tag{27}$$

Going back to the original formulation in terms of differential equations, the action of $\mathcal{H}$ on a function $v \in \mathcal{X}$ is defined by the successive solution of the tangent linear and adjoint problems as follows:

$$\left\{ \begin{array}{rcl} \dfrac{d\varphi}{dt} & = & \dfrac{dM(\mathbf{x}^t)}{d\mathbf{x}}\varphi(t), \quad t \in (0,T] \\ \varphi(0) & = & v, \end{array} \right. \tag{28}$$

$$\left\{ \begin{array}{rcl} -\dfrac{d\varphi^*}{dt} - \left[\dfrac{dM(\mathbf{x}^t)}{d\mathbf{x}}\right]^* \varphi^* & = & -\left[\dfrac{dH(\mathbf{x}^t)}{d\mathbf{x}}\right]^* R^{-1}\dfrac{dH(\mathbf{x}^t)}{d\mathbf{x}}\varphi, \quad t \in [0,T) \\ \varphi^*(T) & = & 0, \end{array} \right. \tag{29}$$

$$\mathcal{H}(x^t)v = B^{-1}v - \varphi^*(0). \tag{30}$$

Let us note that $\mathcal{H}$ is the Hessian of the cost function

$$J_1(v) = \frac{1}{2}\int_0^T \left\| R^{-1/2}\left(\frac{dH(\mathbf{x}^t)}{d\mathbf{x}}\varphi - \boldsymbol{\xi}_y\right)\right\|^2 dt + \frac{1}{2}\left\|B^{-1/2}(v - \boldsymbol{\xi}_x)\right\|^2, \tag{31}$$

where v and φ are related by the tangent linear model (28). Hence, the name: *Hessian of an auxiliary control problem.* After discretization, the sought covariance P can be obtained in the limited-memory form using the Hessian-vector product $\mathcal{H}v$ by the Lanczos method, or by minimizing the cost function (31) using the BFGS method [2].

2.2 Data assimilation problem for imperfect model

Let us assume that the model (1) which involves an inexact operator $M \neq \mathcal{M}$ is used for DA. The perfect model, however, is replaced by (8), so the true state evolution is given by

$$\left\{ \begin{array}{rcl} \dfrac{d\mathbf{x}^t(t)}{dt} & = & M(\mathbf{x}^t(t)) + \boldsymbol{\eta}^t(t), \quad t \in (0,T] \\ \mathbf{x}^t(0) & = & \mathbf{x}_0^t, \end{array} \right. \tag{32}$$

where $\boldsymbol{\eta}^t(t)$ is a 'true' model error trajectory. As before we define the initial state error $\delta\mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^t$ and the state error along the trajectory $\delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}^t(t)$. Then, for small errors, the state error evolution is given by the tangent linear model with a source term

$$\left\{ \begin{array}{rcl} \dfrac{d\delta\mathbf{x}(t)}{dt} & = & \dfrac{dM(\mathbf{x}^t)}{d\mathbf{x}}\delta\mathbf{x}(t) - \boldsymbol{\eta}^t(t), \quad t \in (0,T] \\ \delta\mathbf{x}(0) & = & \delta\mathbf{x}_0. \end{array} \right. \tag{33}$$

The optimality system for errors is complemented by the equations (19b) and (19c).

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

As before, our goal is to express explicitly the optimal solution error $\delta\mathbf{x}_0$ via the data errors $\boldsymbol{\xi}_y$ and $\boldsymbol{\xi}_x$ and the model error $\boldsymbol{\eta}^t(t)$. Having this in mind, we introduce an operator $\mathbf{K}_1 : \mathcal{Y} \to \mathcal{Y}$

$$\mathbf{K}_1(\mathbf{x})\psi(t) = \varphi(t), \tag{34}$$

such that $\psi(t) \in \mathcal{Y}$ and $\varphi(t) \in \mathcal{Y}$ are related via the tangent linear model

$$\left\{ \begin{array}{rcl} \dfrac{d\varphi(t)}{dt} & = & \dfrac{dM(\mathbf{x})}{d\mathbf{x}}\varphi(t) + \psi(t), \quad t \in (0,T] \\ \varphi(0) & = & 0. \end{array} \right.$$

The corresponding adjoint operator $\mathbf{K}_1^* : \mathcal{Y} \to \mathcal{Y}$

$$\mathbf{K}_1^*(\mathbf{x})\varphi(t) = \psi(t), \tag{35}$$

satisfies the following adjoint model:

$$\left\{ \begin{array}{l} -\dfrac{d\varphi^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \varphi^* = \varphi(t), \quad t \in [0,T), \\ \varphi^*(T) = 0, \\ \psi(t) = \varphi^*(t). \end{array} \right.$$

Taking into account the definitions for $\mathbf{K}_1$ and $\mathbf{K}$ for the error trajectory we write

$$\delta\mathbf{x}(t) = \mathbf{K}(\mathbf{x}^t)\ \delta\mathbf{x}_0 - \mathbf{K}_1(\mathbf{x}^t)\ \boldsymbol{\eta}^t(t). \tag{36}$$

Since the adjoint model and the optimality condition in the error optimality system are the same as in Subsect. 2.1, the only difference from the previously considered case is due to the presence of the second term in (36). Thus, we substitute (36) into (22) and obtain the error equation

$$\mathcal{H}(\mathbf{x}^t)\delta\mathbf{x}_0 = \boldsymbol{\xi}_1(\mathbf{x}^t), \tag{37}$$

where

$$\boldsymbol{\xi}_1(\mathbf{x}) = \boldsymbol{\xi}(\mathbf{x}) + \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]\mathbf{K}_1(\mathbf{x})\boldsymbol{\eta}^t(t). \tag{38}$$

Note that $\mathcal{H}(\mathbf{x}^t)$ in (37) is defined by (25) or by (28)-(30), whereas $\boldsymbol{\xi}(\mathbf{x})$ in (38) - by (26).

We suppose that $\boldsymbol{\eta}^t(t)$ satisfies (11), that is

$$\left\{ \begin{array}{rcl} \dfrac{d\boldsymbol{\eta}^t(t)}{dt} & = & N(\mathbf{x}^t(t))\boldsymbol{\eta}^t(t), \quad t \in (0,T] \\ \boldsymbol{\eta}^t(0) & = & \boldsymbol{\eta}_0^t, \end{array} \right. \tag{39}$$

where $\boldsymbol{\eta}_0^t$ is the true initial model error. We also assume that $\boldsymbol{\eta}_0^t \sim \mathcal{N}(0,Q)$ where Q is the model error covariance. To express explicitly the model error evolution $\boldsymbol{\eta}^t(t)$ via the initial model error $\boldsymbol{\eta}_0^t$ we introduce an operator $\mathbf{K}_2 : \mathcal{X} \to \mathcal{Y}$

$$\mathbf{K}_2(\mathbf{x})v = \varphi(t), \tag{40}$$

such that $v \in \mathcal{X}$ and $\varphi(t) \in \mathcal{Y}$ are related via the model

$$\left\{ \begin{array}{rcl} \dfrac{d\varphi(t)}{dt} & = & N(\mathbf{x})\varphi(t), \ t \in (0,T] \\ \varphi(0) & = & v. \end{array} \right.$$

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

The corresponding adjoint operator $\mathbf{K}_2^* : \mathcal{Y} \to \mathcal{X}$

$$\mathbf{K}_2^*(\mathbf{x})\varphi(t) = v, \tag{41}$$

satisfies the following adjoint model:

$$\begin{cases} -\dfrac{d\varphi^*}{dt} - [N(\mathbf{x})]^*\varphi^* = \varphi(t), \quad t \in [0,T), \\ \varphi^*(T) = 0, \\ v = \varphi^*(0). \end{cases}$$

Taking into account (40) we note that $\boldsymbol{\eta}^t(t) = \mathbf{K}_2\boldsymbol{\eta}_0^t$ and re-write the right-hand side of the error equation in its final form as follows:

$$\boldsymbol{\xi}_1(\mathbf{x}) = \boldsymbol{\xi}(\mathbf{x}) + \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right] \mathbf{K}_1(\mathbf{x})\, \mathbf{K}_2(\mathbf{x})\boldsymbol{\eta}_0^t. \tag{42}$$

Now we express $\delta\mathbf{x}_0$ from (37) and compute its covariance matrix $P = E[\delta\mathbf{x}_0\delta\mathbf{x}_0^T]$. It can be easily seen that if the model error is not correlated to the data errors, i.e. $E[\boldsymbol{\eta}_0^t\boldsymbol{\xi}_x^T] = 0$ and $E[\boldsymbol{\eta}_0^t\boldsymbol{\xi}_y^T] = 0$, then

$$P = E[\mathcal{H}^{-1}\boldsymbol{\xi}_1\boldsymbol{\xi}_1^T\mathcal{H}^{-1}] = \mathcal{H}^{-1}(\mathbf{x}^t) + \mathcal{H}^{-1}\mathbf{K}_3 Q \mathbf{K}_3^*\mathcal{H}^{-1}, \tag{43}$$

where

$$\mathbf{K}_3 = \mathbf{K}^*(\mathbf{x})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right] \mathbf{K}_1(\mathbf{x})\, \mathbf{K}_2(\mathbf{x}).$$

Thus, according to (43) the algorithm for computing P consists of two steps: a) compute $\mathcal{H}^{-1}$, as described at the end of Subsect. 2.2; b) evaluate (by the Lanczos method) the leading eigenpairs of the symmetric positive-definite operator $\mathcal{H}^{-1}\mathbf{K}_3 Q \mathbf{K}_3^*\mathcal{H}^{-1}$ and use them to approximate the second term in (43). Note that all operators involved with $\mathbf{K}_3$ and $\mathbf{K}_3^*$ are defined in the above text.

2.3 Weak constraint formulation

In the two previous subsections the cost-function (2) has been considered. Here we consider the problem of minimization of the cost-function (12):

$$(\mathbf{x}_0^a, \boldsymbol{\eta}_0^a) = \text{argmin } J(\mathbf{x}_0, \boldsymbol{\eta}_0) \tag{44}$$

for the model which combines equations (8) and (11). Then, the optimality system for the weak constraint formulation is complemented by the adjoint models

$$\begin{cases} -\dfrac{d\mathbf{x}^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \mathbf{x}^* - \left[\dfrac{dN(\mathbf{x})}{d\mathbf{x}}\right]^* \boldsymbol{\eta}^* = \left[\dfrac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}(H(\mathbf{x}) - \mathbf{y}), \quad t \in [0,T), \\ \mathbf{x}^*(T) = 0, \end{cases} \tag{45}$$

$$\begin{cases} -\dfrac{d\boldsymbol{\eta}^*}{dt} - [N(\mathbf{x})]^*\boldsymbol{\eta}^* = \mathbf{x}^*, \quad t \in [0,T), \\ \boldsymbol{\eta}^*(T) = 0, \end{cases} \tag{46}$$

and by the optimality conditions

$$\begin{cases} \mathbf{x}^*(0) + B^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) = 0 \\ \boldsymbol{\eta}^*(0) + Q^{-1}\boldsymbol{\eta}_0 = 0. \end{cases} \tag{47}$$

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

Let us denote the initial error in the model error estimate $\boldsymbol{\delta\eta}_0 = \boldsymbol{\eta}_0 - \boldsymbol{\eta}_0^t$ and the trajectory of this error $\boldsymbol{\delta\eta}(t) = \boldsymbol{\eta}(t) - \boldsymbol{\eta}^t(t)$. The true state evolution is given by (32) and the model error evolution by (39). By subtracting (32) from (8) and (39) from (11) for small errors we obtain the extended error evolution model as follows:

$$\begin{cases} \dfrac{d\boldsymbol{\delta}\mathbf{x}}{dt} &= \dfrac{dM(\mathbf{x}^t)}{d\mathbf{x}}\boldsymbol{\delta}\mathbf{x}(t) + \boldsymbol{\delta\eta}(t), \quad t \in (0,T], \\ \boldsymbol{\delta}\mathbf{x}(0) &= \boldsymbol{\delta}\mathbf{x}_0, \\ \dfrac{d\boldsymbol{\delta\eta}}{dt} &= \dfrac{dN(\mathbf{x}^t)}{d\mathbf{x}}\boldsymbol{\delta}\mathbf{x}(t) + N(\mathbf{x}^t)\boldsymbol{\delta\eta}(t), \quad t \in (0,T], \\ \boldsymbol{\delta\eta}(0) &= \boldsymbol{\delta\eta}_0. \end{cases} \tag{48a}$$

Using expressions (17) and (18) in (45) and (47), respectively, we complement the optimality system for errors as follows:

$$\begin{cases} -\dfrac{d\mathbf{x}^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \mathbf{x}^* - \left[\dfrac{dN(\mathbf{x})}{d\mathbf{x}}\right]^* \boldsymbol{\eta}^* = \left[\dfrac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\left(\dfrac{dH(\mathbf{x}^t)}{d\mathbf{x}}\boldsymbol{\delta}\mathbf{x} - \boldsymbol{\xi}_y\right), \\ \mathbf{x}^*(T) = 0, \\ -\dfrac{d\boldsymbol{\eta}^*}{dt} - [N(\mathbf{x})]^*\boldsymbol{\eta}^* = \mathbf{x}^*, \quad t \in [0,T), \\ \boldsymbol{\eta}^*(T) = 0, \end{cases} \tag{48b}$$

$$\begin{cases} \mathbf{x}^*(0) + B^{-1}(\boldsymbol{\delta}\mathbf{x}_0 - \boldsymbol{\xi}_\mathbf{x}) = 0 \\ \boldsymbol{\eta}^*(0) + Q^{-1}(\boldsymbol{\delta\eta}_0 + \boldsymbol{\eta}_0^t) = 0. \end{cases} \tag{48c}$$

Again, our goal is to express explicitly the optimal solution error $\delta\mathbf{x}_0$ via the data errors $\boldsymbol{\xi}_y$ and $\boldsymbol{\xi}_x$ and the model error $\boldsymbol{\eta}_0^t$ from the system (48). Having this in mind, we introduce an operator $\mathbf{K} : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}$

$$\mathbf{K}(\mathbf{x}, \boldsymbol{\eta})(v_1, v_2)^T = \varphi(t), \tag{49}$$

such that $v_{1,2} \in \mathcal{X}$ and $\varphi(t) \in \mathcal{Y}$ are related via the extended tangent linear model

$$\begin{cases} \dfrac{d\varphi(t)}{dt} = \dfrac{dM(\mathbf{x})}{d\mathbf{x}}\varphi(t) + \psi(t), \quad t \in (0,T], \\ \dfrac{d\psi(t)}{dt} = \dfrac{dN(\mathbf{x})}{d\mathbf{x}}\varphi(t) + N(\mathbf{x})\psi(t), \quad t \in (0,T], \\ (\varphi(0), \psi(0))^T = (v_1, v_2)^T. \end{cases} \tag{50}$$

The corresponding adjoint operator $\mathbf{K}^* : \mathcal{Y} \to \mathcal{X} \times \mathcal{X}$

$$\mathbf{K}^*(\mathbf{x}, \boldsymbol{\eta})\varphi(t) = (v_1, v_2)^T \tag{51}$$

satisfies the following extended adjoint model:

$$\begin{cases} -\dfrac{d\varphi^*}{dt} - \left[\dfrac{dM(\mathbf{x})}{d\mathbf{x}}\right]^* \varphi^* - \left[\dfrac{dN(\mathbf{x})}{d\mathbf{x}}\right]^* \psi^* = \varphi(t), \quad t \in [0,T), \\ -\dfrac{d\psi^*}{dt} - [N(\mathbf{x})]^*\psi^* = \varphi^*, \quad t \in [0,T), \\ \varphi^*(T) = 0, \psi^*(T) = 0, \\ (v_1, v_2)^T = (\varphi^*(0), \psi^*(0))^T. \end{cases} \tag{52}$$

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

Taking into account the above definitions for $\mathbf{K}$ and $\mathbf{K}^*$, from the adjoint system (48b) we obtain

$$\begin{pmatrix} \mathbf{x}^*(0) \\ \boldsymbol{\eta}^*(0) \end{pmatrix} = \mathbf{K}^*(\mathbf{x},\boldsymbol{\eta})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\frac{dH(\mathbf{x}^t)}{d\mathbf{x}}\mathbf{K}(\mathbf{x}^t,\boldsymbol{\eta}^t)\begin{pmatrix} \delta\mathbf{x}_0 \\ \delta\boldsymbol{\eta}_0 \end{pmatrix} - \mathbf{K}^*(\mathbf{x},\boldsymbol{\eta})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\boldsymbol{\xi}_y.$$

Let us denote $\boldsymbol{\delta}\mathbf{X} = (\delta\mathbf{x}_0, \delta\boldsymbol{\eta}_0)^T$. Assuming $\mathbf{x} = \mathbf{x}^t$ we substitute the above expression into the optimality condition (48c) and obtain the error equation

$$\mathcal{H}(\mathbf{x}^t,\boldsymbol{\eta}^t)\boldsymbol{\delta}\mathbf{X} = \boldsymbol{\xi}(\mathbf{x}^t,\boldsymbol{\eta}^t), \tag{53}$$

where

$$\mathcal{H}(\mathbf{x},\boldsymbol{\eta}) = \begin{pmatrix} B^{-1} & 0 \\ 0 & Q^{-1} \end{pmatrix} + \mathbf{K}^*(\mathbf{x},\boldsymbol{\eta})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\frac{dH(\mathbf{x})}{d\mathbf{x}}\mathbf{K}(\mathbf{x},\boldsymbol{\eta}) \tag{54}$$

and

$$\boldsymbol{\xi}(\mathbf{x},\boldsymbol{\eta}) = \begin{pmatrix} B^{-1} & 0 \\ 0 & Q^{-1} \end{pmatrix}\begin{pmatrix} \boldsymbol{\xi}_x \\ -\boldsymbol{\eta}_0^t \end{pmatrix} + \mathbf{K}^*(\mathbf{x},\boldsymbol{\eta})\left[\frac{dH(\mathbf{x})}{d\mathbf{x}}\right]^* R^{-1}\boldsymbol{\xi}_y. \tag{55}$$

Now we express $\boldsymbol{\delta}\mathbf{X}$ from (53) and compute its covariance matrix $P = E[\boldsymbol{\delta}\mathbf{X}\boldsymbol{\delta}\mathbf{X}^T]$. It can be easily seen that if the data errors are not correlated, i.e. $E[\boldsymbol{\xi}_x\boldsymbol{\xi}_y^T] = 0$ and, also, the model error is not correlated to the data errors, i.e. $E[\boldsymbol{\eta}_0^t\boldsymbol{\xi}_x^T] = 0$ and $E[\boldsymbol{\eta}_0^t\boldsymbol{\xi}_y^T] = 0$, then

$$E[\boldsymbol{\xi}(\mathbf{x}^t,\boldsymbol{\eta}_0^t)\boldsymbol{\xi}^T(\mathbf{x}^t,\boldsymbol{\eta}_0^t)] = \mathcal{H}(\mathbf{x}^t,\boldsymbol{\eta}_0^t),$$

and, therefore

$$P = E[\mathcal{H}^{-1}\boldsymbol{\xi}\boldsymbol{\xi}^T\mathcal{H}^{-1}] = \mathcal{H}^{-1}E[\boldsymbol{\xi}\boldsymbol{\xi}^T]\mathcal{H}^{-1} = \mathcal{H}^{-1}\mathcal{H}\mathcal{H}^{-1} = \mathcal{H}^{-1}(\mathbf{x}^t,\boldsymbol{\eta}_0^t). \tag{56}$$

The covariance matrix P has the block structure

$$P = \begin{pmatrix} P_{1,1} & P_{1,2} \\ P_{2,1} & P_{2,2} \end{pmatrix},$$

where $P_{1,1}$ is the covariance of $\mathbf{x}_0$, $P_{2,2}$ - the covariance of $\boldsymbol{\eta}_0$ and $P_{1,2} = P_{2,1}^T$ is the cross-covariance.

This result has been anticipated since the weak constraint formulation is equivalent to the strong constraint formulation involving an extended control vector. Let us note that $\mathcal{H}$ is now the Hessian of the auxiliary cost function

$$J_1(v) = \frac{1}{2}\int_0^T \left\| R^{-1/2}\left(\frac{dH(\mathbf{x}^t)}{d\mathbf{x}}\varphi - \boldsymbol{\xi}_y\right)\right\|^2 dt + \frac{1}{2}\left\|B^{-1/2}(v-\boldsymbol{\xi}_x)\right\|^2 + \frac{1}{2}\left\|Q^{-1/2}(\boldsymbol{\delta}\boldsymbol{\eta}_0 + \boldsymbol{\eta}_0^t)\right\|^2, \tag{57}$$

where φ, v and $\boldsymbol{\eta}_0^t$ are related by the tangent linear model (50), using $v_1 = v$ and $v_2 = \boldsymbol{\eta}_0^t$.

3 Details of numerical implementation

Let $\mathbf{x}(t)$ be a spatially distributed function $\theta(x,t)$. As a model for θ we use the 1D Burgers' equation including a nonlinear viscous term

$$\begin{cases} \dfrac{\partial\theta}{\partial t} + \dfrac{1}{2}\dfrac{\partial\theta^2}{\partial x} = \dfrac{\partial}{\partial x}\left(\mu(\theta)\dfrac{\partial\theta}{\partial x}\right) + \eta, \ t\in(0,T], \ x\in(0,1), \\ \theta(x,0) = \theta_0(x), \end{cases} \tag{58}$$

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

with the Neumann boundary conditions

$$(d\theta/dx)|_{x=0} = (d\theta/dx)|_{x=1} = 0 \tag{59}$$

and the viscosity coefficient

$$\mu(\theta) = \mu_0 + \mu_1 \left(d\theta/dx\right)^2, \quad \mu_0, \mu_1 = const > 0. \tag{60}$$

The Burgers' equation is sometimes considered in the DA context as a simple model describing elements of atmospheric motion. For modelling η according to (11) we use the reaction-diffusion equation

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + b\eta = \dfrac{\partial}{\partial x}\left(\mu(\theta)\dfrac{\partial \eta}{\partial x}\right), \quad t \in (0,T], \ x \in (0,1), \\ \eta(x,0) = \eta_0(x), \end{cases} \tag{61}$$

with $b = 2$ in numerical experiments. The reaction equation is often used for modelling the model error evolution (see [20]). We add the diffusion term to introduce some spatial correlations. In numerical tests we use $\mu_0 = 10^{-4}$, $\mu_1 = 10^{-6}$.

For solving (58) we use the implicit time discretization as follows

$$\frac{\theta^i - \theta^{i-1}}{h_t} + \frac{\partial}{\partial x}\left(\frac{1}{2}w(\theta^i)\theta^i - \mu(\theta^i)\frac{\partial \theta^i}{\partial x}\right) - \eta = 0, \quad i = 1, ..., n, \ x \in (0,1), \tag{62}$$

where i is the time integration index, $h_t = T/n$ is a time step. The spatial differential operator is discretized on a uniform grid (h_x is the spatial discretization step, $j = 1, ..., m$ is the node number, m is the total number of grid nodes) using the 'power law' first-order scheme as described in [29], which yields a stable discretization. For each time step we perform nonlinear iterations on coefficients $w(\theta) = \theta$ and $\mu(\theta)$, assuming initially that $\mu(\theta^i) = \mu(\theta^{i-1})$ and $w(\theta^i) = \theta^{i-1}$, and keep iterating until (62) is satisfied (i.e. the norm of the left-hand side in (62) becomes smaller than a threshold $\epsilon_1 = 10^{-12}m^{1/2}$). The same approach is applied to equation (61). In computations we use the following parameters: observation period $T = 0.312$, discretization steps $h_t = 0.004$, $h_x = 0.005$, state vector dimension $m = 200$.

The true initial state is given by the expression

$$\theta_0^t(x) = \begin{cases} 0.5[1 - \cos(8\pi x)], & 0 < x \le 0.4, \\ 0.5[\cos(4\pi(x-1)) - 1], & 0.6 \le x < 1, \\ 0, & \text{otherwise}. \end{cases} \tag{63}$$

The graph of $\theta_0^t(x)$ is presented at Fig. 1 in line 1. A general property of Burgers solutions is that a smooth initial state evolves into a state characterised by areas of severe gradients (or even shocks in the inviscid case), see the final state $\theta^t(x,T)$ in line 2. In addition, the initial condition for the true model error $\eta^t(x,0)$ is presented in line 3, and the final time error $\eta^t(x,T)$ in line 4. The observation scheme for each case consists of a set of stationary sensors located at: $x_k = (0.3,\ 0.4,\ 0.5,\ 0.6,\ 0.7)$, the observation error standard deviation is $\sigma_o = 0.01$.

The background error and the model error covariances B and Q are defined under the assumption that these errors belongs to the Sobolev space $W_2^2[0,1]$ (see [30], Section 5.1). For simplicity these two covariances are considered equal. The standard deviation (square root of the covariance matrix diagonal) is presented in dotted line (line 1) at Fig. 3, the correlation radius (at some distance from the boundaries) is $r = 0.1$.

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

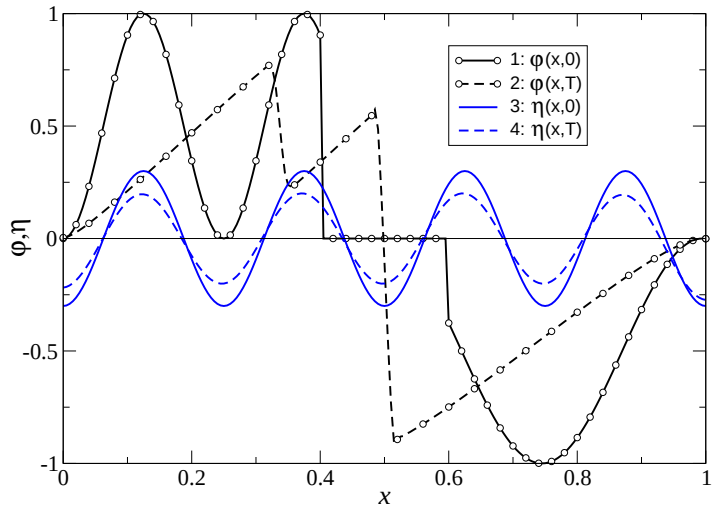

Figure 1: Initial and final states for the state variable θ (lines 1,2) and for the model error η (lines 3,4).

Truncated eigenvalue decomposition of symmetric operators involved in computations is evaluated by means of the Implicitly Restarted Arnoldi Method implemented in the ARPACK software [31]. Consistent tangent linear and adjoint models have been generated from the original forward solver implementing (62) by the Automatic Differentiation tool TAPENADE [32] and checked using the standard gradient test. The random series are produced by a pseudo-random generator, which uses the subroutines *'gasdev'* and *'run2'* provided in [33].

4 Numerical results

Here we describe the basic idea and the results of numerical experiments. We assume that the output of the extended model (8),(11) (controlled by the pair $(\mathbf{x}_0, \boldsymbol{\eta}_0)$)) is equivalent to the output of a hypothetical perfect model (controlled by $\mathbf{x}_0$). Thus, the extended model is used to generate synthetic observations y using $(\mathbf{x}_0^t, \boldsymbol{\eta}_0^t)$)). Initially we consider zero observation and background noise. First, we assimilate the synthetic data using the strong constraint variational DA. The corresponding optimal solution error is presented at Fig. 2 (left) in dashed line. This result shows the component of the error due to the model error η. Next we assimilate the same data using the weak constraint variational DA. The corresponding optimal solution error is presented in solid line. As expected, this error is largely eliminated here, but it is not zero because of the penalty constraints being imposed on $\boldsymbol{\eta}_0$ in the cost function (12). The computations have been also replicated using noisy data; the corresponding results are presented at Fig. 2 (right). Here, the optimal solution error rises from three different sources (observation, background and model errors) which interact nonlinearly. However, the reduction of the optimal solution error due to suppression of the model error is quite evident, most notably at $x \approx 0.3$.

Let us define the *standard deviation* (sd) vector $\sigma[V]$, such that its elements $\sigma_i[\cdot]$, $i = 1, \ldots, m$ are the square-roots of the diagonal entries $V_{i,i}$ of any symmetric matrix V. Each σ_i is associated with an element of the vector $\theta_0(x_i)$ or $\eta_0(x_i)$, hence, in Fig. 3 the sd vector is represented as

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

a function of x. In particular, line 1 shows the sd vector of the background error $\boldsymbol{\xi}_x$ and of the model error η_0^t, i.e. $\sigma[B]$ and $\sigma[Q]$, correspondingly. The two coincide in this particular numerical experiment.

First we consider the strong constraint DA case: line 2 shows $\sigma[P]$, where P is the optimal solution error covariance in (27), i.e. without the model error. Line 3 demonstrates $\sigma[P]$, where P is given by (43), i.e. including the model error. One can observe a significant rise in the optimal solution error due to the presence of the model error. The weak constraint DA case is represented by lines 4 and 5, which show $\sigma[P_{1,1}]$ and $\sigma[P_{2,2}]$ defined by (56), correspondingly. As expected, here the sd value is notably reduced. It is slightly larger than the values in line 2, but this is a price for achieving robustness with respect to the model error. Let us note that the reduction of the model error sd (difference between lines 1 and 5) is not too significant. This means that η_0 is not well estimated by itself.

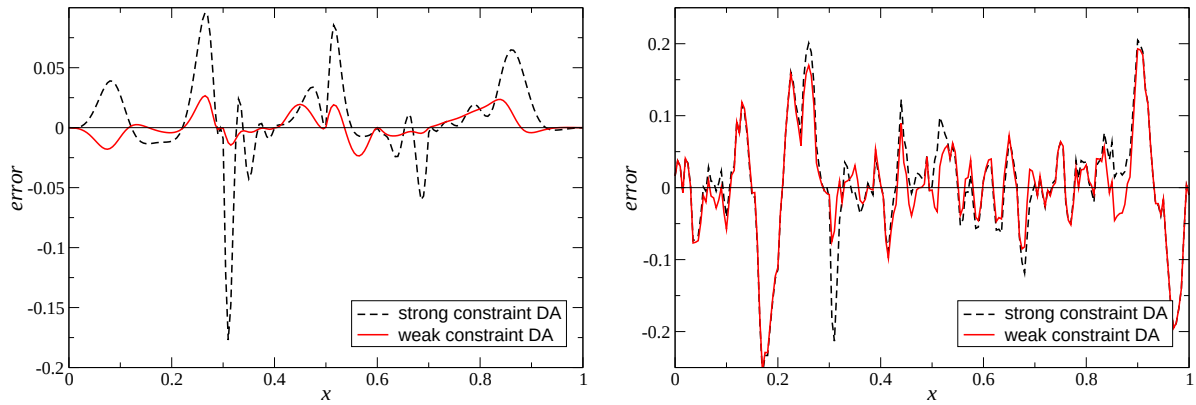

Figure 2: Optimal solution error involving imperfect model: left - exact data; right - noisy data.

5 Conclusions

In this paper the error quantification issue in variational DA problems involving imperfect models has been addressed. In case of a perfect model, the optimal solution error covariance is approximated by the inverse Hessian of an auxiliary control problem. This is a well established result. With an imperfect model, for the strong constraint variational DA formulation the result is given by (43). One may conclude that, due to the presence of the model error, the variance in this case is always larger and the difference may be significant. The formula (43) provides a generic tool for numerical assessment of the influence of the model error on the optimal solution accuracy. It is worth mentioning that the analytical assessment methods have a very limited applicability in practice. For the weak constraint variational DA the optimal solution error covariance is, again, approximated by the inverse Hessian (56) of an auxiliary control problem associated with the extended (augmented) DA formulation. This result has been expected. The error variance in this case is reduced as compared to the strong constraint case.

Let us finally make two remarks. First, the optimal solution error is unbiased if the model error is unbiased. In this case it can be quantified solely by the covariance. Otherwise, the frequency domain deterministic analysis presented in [6] seems more appropriate. Secondly, the results above

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

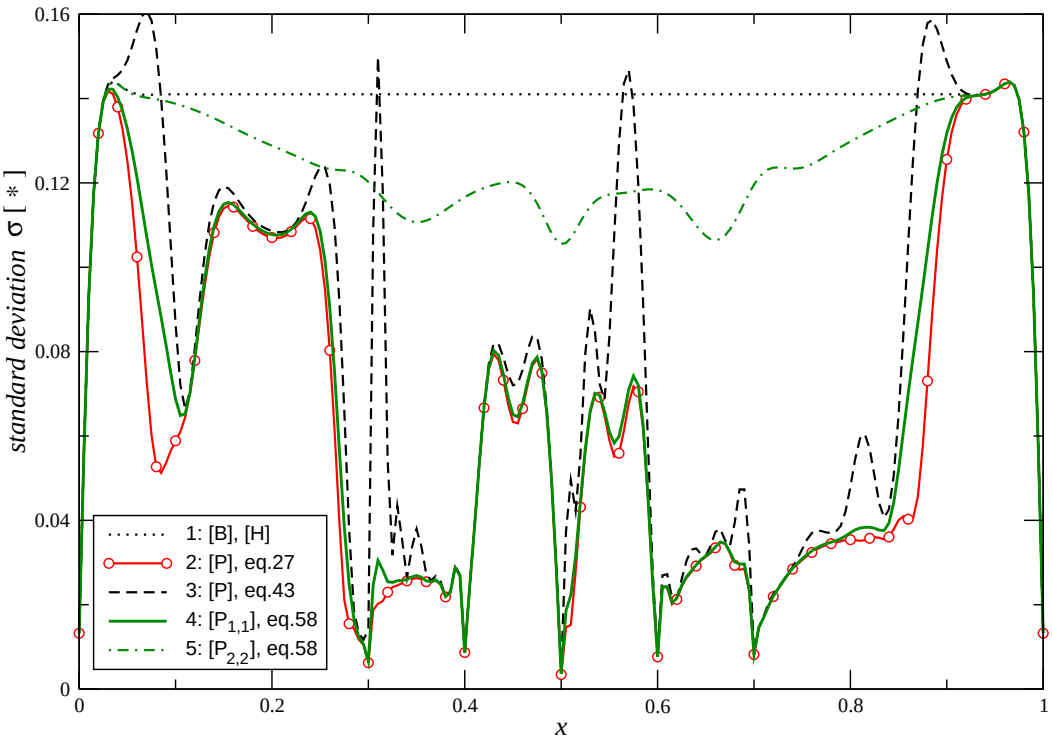

Figure 3: Standard deviation $\sigma[\cdot](x)$ for different cases.

are valid if the chosen model error evolution equation is adequate, which may be difficult to achieve in reality. The mentioned issues require a separate investigation.

Acknowledgments

This work was carried out within the SAMOVAR project (CNRS-RAS), Russian Science Foundation project 14-11-00609 (studies in Section 2), and the project 15-01-01583 of the Russian Foundation for the Basic Research.

References

[1] W. Thacker, "The role of the Hessian matrix in fitting models to measurements," *J. Geophys. Res.*, vol. 94, no. C5, pp. 6177–6196, 1989.

[2] I. Gejadze, F.-X. Le Dimet, and V. Shutyaev, "On analysis error covariances in variational data assimilation," *SIAM J. Sci. Computing*, vol. 30, no. 4, pp. 1847–1874, 2008.

[3] Y. Sasaki, "Some basic formalism in numerical variational analysis," *Month. Wea. Rev.*, vol. 98, no. 12, pp. 875–883, 1970.

[4] A. Jazwinski, *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

[5] S. Cotter, M. Dashti, J. Robinson, and A. Stuart, "Bayesian inverse problems for functions and applications to fluid mechanics," *Inverse Problems*, vol. 25, pp. 1–43, 2009.

[6] I. Gejadze and V. Shutyaev, "An optimal control problem of initial data restoration," *Computational Mathematics and Mathematical Physics*, vol. 39, no. 9, pp. 1416–1425, 1999.

[7] F.-X. Le Dimet and V. Shutyaev, "On deterministic error analysis in variational data assimilation," *Nonlinear Processes in Geophysics*, vol. 12, pp. 481–490, 2005.

[8] T. Vukićević, M. Steyscal, and M. Hecht, "Properties of advection algorithms in the context of variational data assimilation," *Q.J.R. Meteorol. Soc.*, vol. 129, pp. 1221–1231, 2001.

[9] S. Jenkins, C. Budd, M. Freitag, and N. Smith, "The effect of numerical model error on data assimilation," *Comp. Appl. Math.*, vol. 290, pp. 567–588, 2015.

[10] A. Tikhonov, "Regularization of incorrectly posed problems," *Soviet Math. Dokl.*, vol. 4, pp. 1624–1627, 1963.

[11] F.-X. Le Dimet and O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects," *Tellus A*, vol. 38A, no. 2, pp. 97–110, 1986.

[12] L. Ning, F. Carli, A. Ebtehaj, E. Foufoula-Georgiou, and T. Georgiou, " Coping with model error in variational data assimilation using optimal mass transport," *Water Resour. Res.*, vol. 50, pp. 5817–5830, 2014.

[13] J. Derber, "A variational continuous assimilation technique," *Mon. Weather Rev.*, vol. 117, pp. 2437–2446, 1989.

[14] Y. Trémolet, "Model-error estimation in 4D-Var," *Q.J.R. Meteorol. Soc.*, vol. 133, no. 626, pp. 1267–1280, 2007.

[15] A. Carrassi and S. Vannitsem, "Accounting for model error in variational data assimilation: a deterministic approach," *Month. Wea. Rev.*, vol. 138, pp. 875–883, 2010.

[16] D. Furbish, M. Y. Hussaini, F.-X. Le Dimet, P. Ngnepieba, and Y. Wu, "On discretization error and its control in variational data assimilation," *Tellus*, vol. 60A, pp. 979–991, 2008.

[17] A. K. Griffith and N. K. Nichols, "Adjoint methods in data assimilation for estimating model error," *Flow, Turbulence and Combustion*, vol. 65, no. 3/4, pp. 469–488, 2000.

[18] A. Griffith and N. Nichols, "Accounting for model error in data assimilation using adjoint methods," in *Computational Differentiation: Techniques, Applications and Tools* (M. Berz, C. Bischof, G. Corliss, and A. Greiwank, eds.), pp. 195–204, SIAM, Philadelphia, 1996.

[19] A. Vidard, A. Piacentini, and F.-X. Le Dimet, "Variational data analysis with control of the forecast bias," *Tellus A*, vol. 56, pp. 1–12, may 2004.

[20] S. Akella and I. Navon, "Different approaches to model error formulation in 4D-Var: a study with high resolution advection schemes," *Tellus*, vol. 61A, pp. 112–128, 2009.

Author-produced version of the article published in International Journal for Numerical Methods in Fluids, 2017, N°83(3), p.276-290.
The original publication is available at http://onlinelibrary.wiley.com
http://dx.doi.org/10.1002/fld.4266

[21] A. Vidard, *Vers une prise en compte des erreurs-modèle en assimilation de données 4D-variationnelle. Application à un modèle réaliste d'océan.* PhD thesis, Université Joseph Fourier, dec 2001.

[22] A. Vidard, E. Blayo, F.-X. Le Dimet, and A. Piacentini, "4D Variational Data Analysis with Imperfect Model," *Flow, Turbulence and Combustion*, vol. 65, no. 3-4, pp. 489–504, 2000.

[23] A. Vidard, F.-X. Le Dimet, and A. Piacentini, "Determination of optimal nudging coefficients," *Tellus A*, vol. 55, pp. 1–15, jan 2003.

[24] S. Durbiano, *Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données.* PhD thesis, Université Joseph Fourier, dec 2001.

[25] Y. Trémolet, "Accounting for an imperfect model in 4D-Var," *Q.J.R. Meteorol. Soc.*, vol. 132, no. 621, pp. 2483–2504, 2006.

[26] T. Kraus, P. Kuhl, L. Wirsching, H. G. Bock, and M. Diehl, "A moving horizon state estimation algorithm applied to the tennessee eastman benchmark process," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, (Heidelberg, Germany), pp. 377–382, sep 2006.

[27] T. Lagarde, *Nouvelle approche des méthodes d'assimilation de données: les algorithmes de point selle.* PhD thesis, Université Paul Sabatier, Toulouse, 2000.

[28] M. Fisher and H. Auvinen, "Long Window 4D-Var," in *Proceedings of the ECMWF Seminar Series on Data assimilation for atmosphere and ocean, 6-9 September 2011*, pp. 189–202, 2012.

[29] S. Patankar, *Numerical Heat Transfer and Fluid Flow.* Hemisphere Publishing Corporation, 1980.

[30] I. Gejadze, F.-X. Le Dimet, and V. Shutyaev, "On optimal solution error covariances in variational data assimilation problems," *Journal of Computational Physics*, vol. 229, no. 6, pp. 2159–2178, 2010.

[31] R. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods.* SIAM, 1998.

[32] L. Hascoët and V. Pascual, "TAPENADE 2.1 user's guide," *INRIA Technical Report*, no. 0300, pp. 1–78, 2004.

[33] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical recipes in Fortran 77: the art of scientific computing.* Cambridge Press, 1992.