ARTICLE IN PRESS

Journal of Computational and Applied Mathematics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation

V. Shutyaev [a,b,*], F.-X. Le Dimet [c], E. Parmuzin [a,b]

[a] Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina 8, Moscow 119333, Russia
[b] Moscow Institute for Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russia
[c] LJK, Université Grenoble Alpes, 700 Avenue Centrale, 38401 Domaine Universitaire de Saint-Martin-d'Hères, France

ARTICLE INFO

Article history:
Received 28 December 2018
Received in revised form 18 July 2019

Keywords:
Data assimilation
Optimal control
Parameter estimation
Sensitivity
Sea thermodynamics model

ABSTRACT

The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find simultaneously unknown parameters and initial state of the model. A response function is considered as a functional of the optimal solution after assimilation. The sensitivity of the response function to the observation data is studied. The gradient of the response function with respect to observations is related to the solution of a non-standard problem involving the coupled system of direct and adjoint equations. Based on the Hessian of the original cost function, the solvability of the non-standard problem is studied. An algorithm to compute the gradient of the response function with respect to observation data is formulated and justified. A numerical example is presented for variational data assimilation problem for the Baltic Sea thermodynamics model.

1. Introduction

The methods of variational data assimilation have become a very important tool for state observation and parameter estimation for geophysical models. The problems of variational data assimilation can be formulated as optimal control problems (e.g. [1–6]) to find unknown model parameters such as initial and boundary conditions, right-hand sides in the model equations, distributed coefficients, based on minimization of the cost function related to observations. A necessary optimality condition reduces an optimal control problem to an optimality system which involves the model equations, the adjoint problem, and input data functions. The optimal solution depends on the observation data, which may contain uncertainties, and for the forecasts it is very important to study the sensitivity of the optimal solution and its functionals with respect to observation errors [7].

The necessary optimality condition is related to the gradient of the original cost function, thus to study the sensitivity of the optimal solution, one should differentiate the optimality system with respect to observations. In this case, we come to the so-called second-order adjoint problem [8]. The first studies of sensitivity of the response functions after assimilation with the use of second-order adjoint were done by [9] for variational data assimilation problem aimed at restoration of initial condition, where sensitivity with respect to model parameters was considered. The equations of the forecast sensitivity to observations in a four-dimensional (4D-Var) data assimilation were derived by [10]. Based on these results, a practical computational approach was given by [11] to quantify the effect of observations in 4D-Var data assimilation.

* Corresponding author at: Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina 8, Moscow 119333, Russia.
E-mail address: victor.shutyaev@mail.ru (V. Shutyaev).

https://doi.org/10.1016/j.cam.2019.112368
0377-0427/

Sensitivity of the optimal solution is related to its statistical properties (see [12–16]). General sensitivity analysis in variational data assimilation with respect to observations for a nonlinear dynamic model was given by [17] to control the initial-value function. The paper [18] presented the sensitivity analysis with respect to observations in variational data assimilation aimed at restoration of unknown parameters of a dynamic model.

This paper generalizes the results of [17] and [18] and presents the sensitivity analysis with respect to observations in variational data assimilation aimed at simultaneous restoration of unknown parameters and initial state of a dynamic model. The problems of parameter estimation are common inverse problems considered in geophysics and in engineering applications (see [19–28]). Last years an interest is rising to the joint initial state and parameter estimation using 4D-Var [29–31].

We consider a dynamic formulation of variational data assimilation problem for joint parameter and initial state estimation in a continuous form, but the presented sensitivity analysis formulas with respect to observations do not follow from our previous results [17] and [18] and constitute a novelty of this paper. The main contribution of the paper, as compared to [17] and [18], is a derivation of new formulas for the gradient of a response function with respect to observations in variational data assimilation problem aimed at joint parameter and initial state estimation for a general nonlinear dynamic model.

This paper is organized as follows. In Section 2, we give the statement of the variational data assimilation problem for a nonlinear evolution model to estimate simultaneously the model parameters and the initial state. In Section 3, sensitivity of the response function after assimilation with respect to observations is studied, and the theorem is proved to relate the gradient to the solution of a non-standard problem. An algorithm to compute the gradient of the response function is formulated, based on an operator equation involving the Hessian of the original cost function, and the solvability of the non-standard problem is studied. In Section 4, we consider a simple example with known manufactured exact solution and present some numerical results. Section 5 presents an application of the theory to the data assimilation problem for the Baltic Sea thermodynamics model with a numerical example. The main results are discussed in the Conclusions.

2. Statement of the problem

Consider the mathematical model of a physical process that is described by the nonlinear evolution problem

$$\begin{cases} \dfrac{\partial\varphi}{\partial t} = F(\varphi,\lambda)+f, \quad t\in(0,T) \\ \varphi\big|_{t=0} = u, \end{cases} \tag{1}$$

where the initial state u is supposed to be from a Hilbert space X, the unknown function $\varphi=\varphi(t)$ belongs to $Y=L_2(0,T;X)$ with the norm $\|\varphi\|_Y=(\varphi,\varphi)_Y^{1/2}=(\int_0^T\|\varphi(t)\|_X^2dt)^{1/2}$, F is a nonlinear operator mapping $Y\times Y_p$ into Y, Y_p is a Hilbert space (space of model parameters), $f\in Y$. We suppose that for given $u\in X, f\in Y$ and $\lambda\in Y_p$ there exists a unique solution $\varphi\in Y$ to (1) with $\dfrac{\partial\varphi}{\partial t}\in Y$. The function λ is an unknown model parameter, and we suppose that the initial state u is also unknown.

We introduce the cost function as a functional on $X\times Y_p$ in the form

$$J(u,\lambda)=\frac{1}{2}(V_1(u-u_b),u-u_b)_X+\frac{1}{2}(V_2(\lambda-\lambda_b),\lambda-\lambda_b)_{Y_p}+\frac{1}{2}(V_3(C\varphi-\varphi_{obs}),C\varphi-\varphi_{obs})_{Y_{obs}}, \tag{2}$$

where $u_b\in X, \lambda_b\in Y_p$ are prior (background) functions, $\varphi_{obs}\in Y_{obs}$ is a prescribed function (observational data), Y_{obs} is a Hilbert space (observation space), $C:Y\to Y_{obs}$ is a linear bounded operator (observation operator), $V_1:X\to X, V_2:Y_p\to Y_p$ and $V_3:Y_{obs}\to Y_{obs}$ are symmetric positive definite bounded operators. Usually, V_1,V_2,V_3 are chosen as inverse covariance operators of observation and background errors [7,30].

Let us consider the following data assimilation problem with the aim to find the initial value u and the parameter λ: for given $f\in Y, \varphi_{obs}\in Y_{obs}, u_b\in X, \lambda_b\in Y_p$, find $u\in X, \lambda\in Y_p$ and $\varphi\in Y$ such that they satisfy (1), and on the set of solutions to (1), the functional $J(u,\lambda)$ takes the minimum value, i.e.

$$\begin{cases} \dfrac{\partial\varphi}{\partial t} = F(\varphi,\lambda)+f, \quad t\in(0,T) \\ \varphi\big|_{t=0} = u, \\ J(u,\lambda) = \inf\limits_{w\in X, v\in Y_p} J(w,v). \end{cases} \tag{3}$$

We suppose that the solution of (3) exists. Let us note that the solvability of the parameter estimation problems (or identifiability) has been addressed, e.g., in [32,33]. To derive the optimality system, we assume the solution φ and the operator $F(\varphi,\lambda)$ in (1)–(2) are regular enough, and for $w\in X$, $v\in Y_p$ find the gradient of the functional J with respect to u and λ:

$$J'_u(u,\lambda)w=(V_1(u-u_b),w)_X+(C^*V_3(C\varphi-\varphi_{obs}),\tilde{\phi})_Y, \tag{4}$$

$$J'_\lambda(u,\lambda)v=(V_2(\lambda-\lambda_b),v)_{Y_p}+(C^*V_3(C\varphi-\varphi_{obs}),\phi)_Y, \tag{5}$$

where ϕ is the solution to the problem:

$$\begin{cases} \dfrac{\partial \phi}{\partial t} &= F'_\varphi(\varphi,\lambda)\phi + F'_\lambda(\varphi,\lambda)v, \quad t\in(0,T), \\ \phi\big|_{t=0} &= 0, \end{cases} \tag{6}$$

and $\tilde{\phi}$ is the solution to the problem:

$$\begin{cases} \dfrac{\partial \tilde{\phi}}{\partial t} &= F'_\varphi(\varphi,\lambda)\tilde{\phi}, \quad t\in(0,T), \\ \phi\big|_{t=0} &= w. \end{cases} \tag{7}$$

Here $F'_\varphi(\varphi,\lambda): Y\to Y,\ F'_\lambda(\varphi,\lambda): Y_p\to Y$ are the Fréchet derivatives of F [34] with respect to φ and λ, correspondingly, and C^* is the adjoint operator to C defined by $(C\varphi,\psi)_{Y_{obs}} = (\varphi, C^*\psi)_Y,\ \varphi\in Y, \psi\in Y_{obs}$.

Let us introduce the adjoint operator $(F'_\varphi(\varphi,\lambda))^*: Y\to Y$ and consider the adjoint problem:

$$\begin{cases} \dfrac{\partial \varphi^*}{\partial t} + (F'_\varphi(\varphi,\lambda))^*\varphi^* &= C^*V_3(C\varphi-\varphi_{obs}), \quad t\in(0,T) \\ \varphi^*\big|_{t=T} &= 0. \end{cases} \tag{8}$$

The problem (8) is adjoint with respect to the linearized (tangent linear) problems (6), (7), therefore, it is also linear in φ^*, however, it is still nonlinear in φ.

In the below consideration, we assume that the direct and adjoint linear problems of the form

$$\begin{cases} \dfrac{\partial \phi}{\partial t} - F'_\varphi(\varphi,\lambda)\phi &= p, \quad t\in(0,T) \\ \phi\big|_{t=0} &= q, \end{cases}$$
$$\begin{cases} -\dfrac{\partial \phi^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\phi^* &= g, \quad t\in(0,T) \\ \phi^*\big|_{t=T} &= 0 \end{cases}$$

with $p,g\in Y, q\in X$ have the unique solutions $\phi,\phi^*\in Y$ and $\dfrac{\partial\phi}{\partial t}, \dfrac{\partial\phi^*}{\partial t}\in Y$. From (4)–(8) we get

$$J'_u(u,\lambda)w = (V_1(u-u_b),w)_X - (\varphi^*\big|_{t=0}, w)_X, \tag{9}$$

$$J'_\lambda(u,\lambda)v = (V_2(\lambda-\lambda_b),v)_{Y_p} - (\varphi^*, F'_\lambda(\varphi,\lambda)v)_Y = (V_2(\lambda-\lambda_b),v)_{Y_p} - ((F'_\lambda(\varphi,\lambda))^*\varphi^*, v)_{Y_p}, \tag{10}$$

where $(F'_\lambda(\varphi,\lambda))^*: Y\to Y_p$ is the operator adjoint to $F'_\lambda(\varphi,\lambda)$. Thus, the gradient of J is defined by

$$J'_u(u,\lambda) = V_1(u-u_b) - \varphi^*\big|_{t=0},\quad J'_\lambda(u,\lambda) = V_2(\lambda-\lambda_b) - (F'_\lambda(\varphi,\lambda))^*\varphi^*.$$

The necessary optimality condition [1] is $gradJ = 0$, therefore, $J'_u(u,\lambda)=0, J'_\lambda(u,\lambda)=0$. From (3)–(10) we obtain the optimality system:

$$\begin{cases} \dfrac{\partial \varphi}{\partial t} &= F(\varphi,\lambda)+f, \quad t\in(0,T), \\ \varphi\big|_{t=0} &= u, \end{cases} \tag{11}$$

$$\begin{cases} \frac{\partial \varphi^*}{\partial t} + (F'_\varphi(\varphi,\lambda))^*\varphi^* &= C^*V_3(C\varphi-\varphi_{obs}), \quad t\in(0,T) \\ \varphi^*\big|_{t=T} &= 0, \end{cases} \tag{12}$$

$$V_1(u-u_b) - \varphi^*\big|_{t=0} = 0, \tag{13}$$

$$V_2(\lambda-\lambda_b) - (F'_\lambda(\varphi,\lambda))^*\varphi^* = 0. \tag{14}$$

We suppose that the system (11)–(14) has a unique solution $\varphi,\varphi^*\in Y, u\in X, \lambda\in Y_p$. The system (11)–(14) may be considered as a generalized model of the form $\mathcal{A}(U)=0$ with the state variable $U=(\varphi,\varphi^*,u,\lambda)$, and it contains the information on the observation data $\varphi_{obs}\in Y_{obs}$. Below we study the sensitivity of functionals of the optimal solution with respect to the observation data.

3. Sensitivity of response functions with respect to observations

In many applications the observation data cannot be measured precisely, and therefore, it is important to be able to estimate the impact of uncertainties in observations on the outputs of the model after assimilation. Such outputs may be response functions considered as functionals of the optimal solution.

We introduce a response function $G(\varphi,u,\lambda)$, which is supposed to be a real-valued function and can be considered as a functional on $Z = Y\times X\times Y_p$. We are interested in the sensitivity of G with respect to φ_{obs}, with φ, u and λ obtained from the optimality system (11)–(14). By definition, the sensitivity is defined by the gradient of G with respect to φ_{obs}:

$$\frac{dG}{d\varphi_{obs}} = \frac{\partial G}{\partial\varphi}\frac{\partial\varphi}{\partial\varphi_{obs}} + \frac{\partial G}{\partial\lambda}\frac{\partial\lambda}{\partial\varphi_{obs}} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial\varphi_{obs}}, \tag{15}$$

where $\frac{\partial G}{\partial \varphi} : Z \to Y$, $\frac{\partial G}{\partial \lambda} : Z \to Y_p$, $\frac{\partial G}{\partial u} : Z \to X$, and $\frac{\partial \varphi}{\partial \varphi_{obs}}, \frac{\partial \lambda}{\partial \varphi_{obs}}, \frac{\partial u}{\partial \varphi_{obs}}$ are the Gâteaux derivatives of φ, λ, u with respect to φ_{obs}.

Let $\delta\varphi_{obs}$ be a perturbation on φ_{obs}, then we obtain from the optimality system (11)–(14):

$$\begin{cases} \dfrac{\partial \delta\varphi}{\partial t} &= F'_\varphi(\varphi,\lambda)\delta\varphi + F'_\lambda(\varphi,\lambda)\delta\lambda, \quad t \in (0,T) \\ \delta\varphi\big|_{t=0} &= \delta u, \end{cases} \tag{16}$$

$$\begin{cases} -\dfrac{\partial \delta\varphi^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\delta\varphi^* - (F''_{\varphi\varphi}(\varphi,\lambda)\delta\varphi)^*\varphi^* &= (F''_{\varphi\lambda}(\varphi,\lambda)\delta\lambda)^*\varphi^* \\ & \quad -C^*V_3(C\delta\varphi - \delta\varphi_{obs}), \\ \delta\varphi^*\big|_{t=T} &= 0, \end{cases} \tag{17}$$

$$V_1\delta u - \delta\varphi^*\big|_{t=0} = 0, \tag{18}$$

$$V_2\delta\lambda - (F''_{\lambda\varphi}(\varphi,\lambda)\delta\varphi)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)\delta\lambda)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*\delta\varphi^* = 0, \tag{19}$$

and

$$\left(\frac{dG}{d\varphi_{obs}}, \delta\varphi_{obs}\right)_{Y_{obs}} = \left(\frac{\partial G}{\partial \varphi}, \delta\varphi\right)_Y + \left(\frac{\partial G}{\partial \lambda}, \delta\lambda\right)_{Y_p} + \left(\frac{\partial G}{\partial u}, \delta u\right)_X, \tag{20}$$

where $\delta\varphi$, $\delta\varphi^*$, $\delta\lambda$, δu are the solutions of (16)–(19).

The following statement is valid.

Theorem 1. *Let $P_1, P_2 \in Y, P_3 \in Y_p, P_4 \in X$ be the solutions of the following system of equations*

$$\begin{cases} -\dfrac{\partial P_1}{\partial t} - (F'_\varphi(\varphi,\lambda))^*P_1 - (F''_{\varphi\varphi}(\varphi,\lambda)P_2)^*\varphi^* &= (F''_{\lambda\varphi}(\varphi,\lambda)P_3)^*\varphi^* - C^*V_3CP_2 + \dfrac{\partial G}{\partial \varphi}, \\ P_1\big|_{t=T} &= 0, \end{cases} \tag{21}$$

$$\begin{cases} \dfrac{\partial P_2}{\partial t} - F'_\varphi(\varphi,\lambda)P_2 - F'_\lambda(\varphi,\lambda)P_3 &= 0, \quad t \in (0,T) \\ P_2\big|_{t=0} - P_4 &= 0, \end{cases} \tag{22}$$

$$V_1P_4 - P_1\big|_{t=0} = \frac{\partial G}{\partial u}, \tag{23}$$

$$V_2P_3 - (F''_{\varphi\lambda}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)P_3)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*P_1 = \frac{\partial G}{\partial \lambda}, \tag{24}$$

where $\varphi, \varphi^* \in Y, u \in X, \lambda \in Y_p$ are the solution of the optimality system (11)–(14). Then the gradient of G with respect to φ_{obs} is given by

$$\frac{dG}{d\varphi_{obs}} = V_3CP_2. \tag{25}$$

Proof of this Theorem is presented in the Appendix.

We obtain a coupled system of two differential Eqs. (21) and (22) of the first order with respect to time, with additional conditions (23)–(24). To study this non-standard problem (21)–(24) with mutually dependent initial conditions for P_1, P_2, we reduce it to a single operator equation involving the Hessian of the original cost function.

Let us introduce the auxiliary variables $v = P_3 \in Y_p$, $w = P_4 \in X$ and rewrite the non-standard problem (21)–(24) in an equivalent form:

$$\begin{cases} \dfrac{\partial P_2}{\partial t} - F'_\varphi(\varphi,\lambda)P_2 &= F'_\lambda(\varphi,\lambda)v, \quad t \in (0,T) \\ P_2\big|_{t=0} &= w, \end{cases} \tag{26}$$

$$\begin{cases} -\dfrac{\partial P_1}{\partial t} - (F'_\varphi(\varphi,\lambda))^*P_1 - (F''_{\varphi\varphi}(\varphi,\lambda)P_2)^*\varphi^* &= (F''_{\lambda\varphi}(\varphi,\lambda)v)^*\varphi^* - C^*V_3CP_2 + \dfrac{\partial G}{\partial \varphi}, \\ P_1\big|_{t=T} &= 0, \end{cases} \tag{27}$$

$$V_1w - P_1\big|_{t=0} = \frac{\partial G}{\partial u}, \tag{28}$$

$$V_2v - (F''_{\varphi\lambda}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)v)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*P_1 = \frac{\partial G}{\partial \lambda} \tag{29}$$

with the four unknowns: $w \in X, v \in Y_p, \ P_1, P_2 \in Y$. Let us write (26)–(29) in the form of an operator equation for $U = (w, v)^T$. We define the operator $\mathcal{H} : X \times Y_p \to X \times Y_p$, which acts on U belonging to $X \times Y_p$, by the successive solution of the following problems:

$$\begin{cases} \dfrac{\partial \phi}{\partial t} - F'_\varphi(\varphi,\lambda)\phi &= F'_\lambda(\varphi,\lambda)v, \quad t \in (0,T) \\ \phi\big|_{t=0} &= w, \end{cases} \tag{30}$$

$$\begin{cases} -\dfrac{\partial \phi^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\phi^* - (F''_{\varphi\varphi}(\varphi,\lambda)\phi)^*\varphi^* &= (F''_{\lambda\varphi}(\varphi,\lambda)w)^*\varphi^* - C^*V_3C\phi, \\ \phi^*\big|_{t=T} &= 0, \end{cases} \tag{31}$$

$$\mathcal{H}U = \left(V_1 w - \phi^*\big|_{t=0},\ V_2 v - (F''_{\varphi\lambda}(\varphi,\lambda)\phi)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)w)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*\phi^* \right)^T, \tag{32}$$

where λ, u, φ and φ^* are the solutions of the optimality system (11)–(14). It is easily seen that (26)–(29) is equivalent to the following equation in $X \times Y_p$:

$$\mathcal{H}U = \mathcal{F} \tag{33}$$

with $\mathcal{F} \in X \times Y_p$ defined by

$$\mathcal{F} = \left(\frac{\partial G}{\partial u} + \tilde{\phi}^*\big|_{t=0},\ \frac{\partial G}{\partial \lambda} + (F'_\lambda(\varphi,\lambda))^*\tilde{\phi}^* \right)^T, \tag{34}$$

where $\tilde{\phi}^* \in Y$ is the solution to the adjoint problem:

$$\begin{cases} -\dfrac{\partial \tilde{\phi}^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\tilde{\phi}^* &= \dfrac{\partial G}{\partial \varphi}, \quad t \in (0,T) \\ \tilde{\phi}^*\big|_{t=T} &= 0. \end{cases} \tag{35}$$

It is easy to make sure that the operator $\mathcal{H}$ defined by (30)–(32) is the Hessian of the original functional J considered on the optimal solution u, λ of the problem (11)–(14): $J''(u,\lambda) = \mathcal{H}$.

Lemma 1. *Under the assumption that $\mathcal{H}$ is positive definite, the operator Eq. (33) is well posed: for every $\mathcal{F} \in X \times Y_p$ there exists a unique solution $U \in X \times Y_p$ and the estimate is valid:*

$$\|U\|_{X\times Y_p} \le c\|\mathcal{F}\|_{X\times Y_p}, \quad c = const > 0. \tag{36}$$

Proof. If the operator $\mathcal{H}$ is positive definite, then for any $U \in X \times Y_p$

$$(\mathcal{H}U, U)_{X\times Y_p} \ge \gamma (U,U)_{X\times Y_p}, \quad \gamma = const > 0.$$

Hence,

$$\|\mathcal{H}U\|_{X\times Y_p} \ge \gamma \|U\|_{X\times Y_p}, \tag{37}$$

and it means that Eq. (33) is uniquely and correctly solvable in $X \times Y_p$ [35].

By definition, $\mathcal{H}$ is self-adjoint, i.e. $\mathcal{H}^* = \mathcal{H}$. Then, the adjoint equation is also correctly solvable, which implies that Eq. (33) is everywhere solvable [35], i.e. for every $\mathcal{F} \in X \times Y_p$ there exists a unique solution $U \in X \times Y_p$.

Let $U \in X \times Y_p$ be the solution of (33) with the right-hand side $\mathcal{F}$, then (37) gives (36) with $c = 1/\gamma$. The lemma is proved.

Therefore, under the assumption that $J''(u,\lambda)$ is positive definite on the optimal solution, the non-standard problem (21)–(24) has a unique solution $P_1, P_2 \in Y, P_3 \in Y_p, P_4 \in X$.

From the above consideration, we come to the following algorithm to compute the gradient of the response function G:

(1) For $\frac{\partial G}{\partial \lambda} \in Y_p,\ \frac{\partial G}{\partial \varphi} \in Y,\ \frac{\partial G}{\partial u} \in X$ solve the adjoint problem

$$\begin{cases} -\dfrac{\partial \tilde{\phi}^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\tilde{\phi}^* &= \dfrac{\partial G}{\partial \varphi}, \quad t \in (0,T) \\ \tilde{\phi}^*\big|_{t=T} &= 0 \end{cases} \tag{38}$$

and put

$$\mathcal{F} = \left(\frac{\partial G}{\partial u} + \tilde{\phi}^*\big|_{t=0},\ \frac{\partial G}{\partial \lambda} + (F'_\lambda(\varphi,\lambda))^*\tilde{\phi}^* \right)^T.$$

(2) Find $U = (w, v)^T$ by solving

$$\mathcal{H}U = \mathcal{F}$$

with the Hessian of the original functional J defined by (30)–(32).

(3) Solve the direct problem

$$\begin{cases} \dfrac{\partial P_2}{\partial t} - F'_\varphi(\varphi,\lambda)P_2 &= F'_\lambda(\varphi,\lambda)v, \quad t \in (0,T) \\ P_2\big|_{t=0} &= w. \end{cases} \tag{39}$$

Please cite this article as: V. Shutyaev, F.-X. Le Dimet and E. Parmuzin, Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation, Journal of Computational and Applied Mathematics (2019) 112368, https://doi.org/10.1016/j.cam.2019.112368.

(4) Compute the gradient of the response function as

$$\frac{dG}{d\varphi_{obs}} = V_3 C P_2. \tag{40}$$

The last formula allows us to estimate the sensitivity of the response functions related to the optimal solution after assimilation, with respect to observation data.

4. Simple example

Let us consider a simple evolution problem for the ordinary differential equation

$$\begin{cases} \dfrac{d\varphi}{dt} + a\varphi = \lambda g, \quad t \in (0,T) \\ \varphi\big|_{t=0} = u, \end{cases} \tag{41}$$

where $u \in \mathbb{R}$; $a, \lambda \in \mathbb{R}$, $g = g(t) \geq 0$. Here, in the notations of Section 2, we have $X = \mathbb{R}$, $Y = L_2(0,T)$, $Y_p = \mathbb{R}$, $F(\varphi, \lambda) = -a\varphi + \lambda g$, $f = 0$. Let us formulate the data assimilation problem to find the initial state u and the parameter λ if we have observation data for φ at the end of the time interval $t = T$. We will minimize the cost function

$$J(u, \lambda) = \inf_{w, v \in \mathbb{R}} J(w, v), \tag{42}$$

where $J(u,\lambda) = \dfrac{\alpha}{2}\left|u - u_b\right|^2 + \dfrac{1}{2}\left|\varphi|_{t=T} - \varphi_{obs}\right|^2$, $\alpha > 0$, and φ is the solution to (41).

Thus, here we have $V_1 = \alpha$, $V_2 = 0$, $V_3 = 1$, $C\varphi = \varphi|_{t=T}$.

In this case $F'_\varphi(\varphi,\lambda) = a$, $F'_\lambda(\varphi,\lambda) = g$, and the optimality system (11)–(14) has the form:

$$\begin{cases} \dfrac{d\varphi}{dt} + a\varphi = \lambda g, \quad t \in (0,T) \\ \varphi\big|_{t=0} = u, \end{cases} \tag{43}$$

$$\begin{cases} \dfrac{d\varphi^*}{dt} - a\varphi^* = 0, \quad t \in (0,T) \\ \varphi^*\big|_{t=T} = \varphi_{obs} - \varphi\big|_{t=T}, \end{cases} \tag{44}$$

$$\alpha(u - u_b) - \varphi^*\big|_{t=0} = 0, \tag{45}$$

$$(g, \varphi^*) = \int_0^T g(t)\varphi^*(t)dt = 0. \tag{46}$$

It is easy to see that the problem of data assimilation (41)–(42) has a unique solution

$$\lambda = \lambda_{opt} = \frac{\varphi_{obs} - \varphi_0 u_b}{\varphi_1}, \quad u = u_{opt} = u_b, \tag{47}$$

where $\varphi_0 = e^{-aT}$, $\varphi_1 = \int_0^T e^{-a(T-t')} g(t')dt'$.

Indeed, if u, λ have the form (47), the solution of problem (41) satisfies $\varphi|_{t=T} = \varphi_{obs}$, and the functional J from (42) attains its minimal value $J = 0$. In this case $\varphi^* = 0$, and the optimality system (43)–(46) is satisfied. Also, we will see below that the Hessian of J is positive definite, and it means the uniqueness of the solution u, λ.

Let us consider the response function in the form

$$G(\varphi, \lambda, u) = \int_0^T \varphi(t)dt. \tag{48}$$

Let $a \neq 0$. After assimilation, taking into account the solution of problem (41), we have

$$G(\varphi, \lambda, u) = \frac{u_{opt}}{a}(1 - e^{-aT}) + \frac{\lambda_{opt}}{a}\left(\int_0^T g(t)dt - \varphi_1\right), \tag{49}$$

where u_{opt}, λ_{opt} are given by (47). Then, by direct differentiation of G with respect to φ_{obs} we have the gradient

$$\frac{dG}{d\varphi_{obs}} = \frac{1}{a\varphi_1}\left(\int_0^T g(t)dt - \varphi_1\right). \tag{50}$$

Let us now apply the algorithm (38)–(40) to compute the gradient of the function G. Since $\dfrac{\partial G}{\partial \varphi} = 1$, $(F'_\varphi(\varphi,\lambda))^* = -a$, then on the first step of the algorithm, we solve the problem (38) and get the solution

$$\tilde{\phi}^*(t) = \frac{1}{a}(1 - e^{-a(T-t)}). \tag{51}$$

Taking into account that $\partial G/\partial\lambda = \partial G/\partial u = 0$ and $(F'_\lambda(\varphi,\lambda))^*\tilde\phi^* = (g,\tilde\phi^*)$, we get $\mathcal{F} = \left(\tilde\phi^*(0),(g,\tilde\phi^*)\right)^T$, i.e., $\mathcal{F} = (\tilde f_0,\tilde f)^T$, where

$$\tilde f_0 = \tilde\phi^*\Big|_{t=0}, \quad \tilde f = \int_0^T g\tilde\phi^* dt = \frac{1}{a}\left(\int_0^T g(t)dt - \varphi_1\right). \tag{52}$$

On the second step of the algorithm, one need to solve the equation $\mathcal{H}U = \mathcal{F}$ with the Hessian $\mathcal{H}$ defined by the formulas (30)–(32). Since all the second order derivatives of $F(\varphi,\lambda)$ equal zero, then it is easily seen that $\mathcal{H}$ in this case is defined by

$$\mathcal{H}U = \left(\alpha w - \phi^*\big|_{t=0}, -\int_0^T g(t)\phi^*(t)dt\right)^T, \quad U = (w,v)^T,$$

where ϕ^* is the solution of the adjoint problem

$$\begin{cases} \dfrac{d\phi^*}{dt} - a\phi^* = 0, \quad t\in(0,T) \\ \phi^*\big|_{t=T} = -\phi\big|_{t=T}, \end{cases} \tag{53}$$

and ϕ is the solution of the forward problem

$$\begin{cases} \dfrac{d\phi}{dt} + a\phi = vg, \quad t\in(0,T) \\ \phi\big|_{t=0} = w. \end{cases} \tag{54}$$

Since $\phi\big|_{t=T} = w\varphi_0 + v\varphi_1$ and

$$\int_0^T g(t)\phi^*(t)dt = -\phi\big|_{t=T}\int_0^T e^{-a(T-t)}g(t)dt = -\phi\big|_{t=T}\varphi_1,$$
$$\phi^*\big|_{t=0} = -\phi\big|_{t=T}e^{-aT} = -\phi\big|_{t=T}\varphi_0,$$

we get

$$\mathcal{H}U = \left(\alpha w + \varphi_0^2 w + \varphi_0\varphi_1 v, \varphi_0\varphi_1 w + \varphi_1^2 v\right)^T,$$

hence $\mathcal{H}$ is the 2×2 matrix

$$\mathcal{H} = \begin{pmatrix} \alpha+\varphi_0^2 & \varphi_0\varphi_1 \\ \varphi_0\varphi_1 & \varphi_1^2 \end{pmatrix}. \tag{55}$$

For $\alpha > 0$, $\varphi_0,\varphi_1 \neq 0$ the matrix $\mathcal{H}$ is positive definite, which confirms the existence and uniqueness of the solution to problem (42).

The solution of the system $\mathcal{H}U = \mathcal{F}$ has the explicit form

$$w = -\frac{\varphi_0\tilde f}{\alpha\varphi_1} + \frac{\tilde f_0}{\alpha}, \quad v = \frac{(\alpha+\varphi_0^2)}{\alpha\varphi_1^2}\tilde f - \frac{\varphi_0}{\alpha\varphi_1}\tilde f_0. \tag{56}$$

On the third step of the algorithm, we need to solve problem (39). Since $F'_\lambda(\varphi,\lambda) = g$, the solution of this problem for $t = T$ has the form

$$P_2\big|_{t=T} = w\varphi_0 + v\varphi_1 = -\frac{\varphi_0^2\tilde f}{\alpha\varphi_1} + \frac{(\alpha+\varphi_0^2)\tilde f}{\alpha\varphi_1} = \frac{\tilde f}{\varphi_1}.$$

Finally, using (40), we get the gradient of G with respect to φ_{obs}:

$$\frac{dG}{d\varphi_{obs}} = CP_2 = P_2|_{t=T} = \frac{\tilde f}{\varphi_1}. \tag{57}$$

From (52) and (57) we have

$$\frac{dG}{d\varphi_{obs}} = \frac{1}{a\varphi_1}\left(\int_0^T g(t)dt - \varphi_1\right). \tag{58}$$

Thus, the gradient obtained by the algorithm (38)–(40) exactly coincides with the value of the gradient obtained in (50) by direct differentiation, which is the expected result.

For a numerical example, we consider the problem (41)–(42) and the response function G in the form (48) for $a = 1, g(t) = 1, \alpha = 10^{-5}$. The exact value of the gradient $\frac{dG}{d\varphi_{obs}}$ is defined by the formula (50). Easy to see that for $a = 1, g(t) = 1$ it has the explicit form:

$$\frac{dG}{d\varphi_{obs}} = \frac{T}{1-e^{-T}} - 1,$$

Table 1
The experiment with different assimilation windows, $\alpha = 10^{-5}$.

Assimilation window	$T = 1$	$T = 2$	$T = 5$	$T = 10$	$T = 100$
Exact gradient	0.5819	1.313	4.034	9	99
Approximate gradient, $\tau = 0.1$	0.5193	1.246	3.978	8.95	98.949
Approximate gradient, $\tau = 0.01$	0.5815	1.302	4.027	8.995	98.997

Table 2
The experiment with different parameters α, $T = 10$, $\tau = 0.01$.

Parameter α	$\alpha = 10^{-5}$	$\alpha = 10^{-2}$	$\alpha = 1$
Exact gradient	9	9	9
Approximate gradient	8.995	8.995	8.995

and it does not require a numerical integration. Approximate values of the gradient were obtained numerically with the help of the algorithm (38)–(40), where the problems (38), (39) were solved using the simplest explicit scheme in time. Table 1 presents for comparison the exact and approximate values of the gradient for different lengths T of the assimilation window and for different time steps τ. One can see from the table that for each case the gradient values obtained by the considered algorithm coincide with the exact values of the gradient with the accuracy $\mathcal{O}(\tau)$. Besides, the gradient $\frac{dG}{d\varphi_{obs}}$ rises with the increase of T, therefore, the sensitivity of the response function with respect to observation errors is increasing, which is natural for a larger assimilation window T.

Note that in this example, the exact gradient defined by the formula (50) does not depend on α, however, the algorithm (38)–(40) involve α as a parameter. Table 2 shows that the resulting approximate values of the gradient obtained by the algorithm also do not change with α.

5. Application: data assimilation problem for a sea thermodynamics model

We consider the sea thermodynamics problem in the form [36]:

$$\begin{aligned}
&T_t + (\bar{U}, \mathrm{Grad})T - \mathrm{Div}(\hat{a}_T \cdot \mathrm{Grad}\, T) = f_T \quad \text{in} \quad D \times (t_0, t_1), \\
&T = T_0 \text{ for } t = t_0 \text{ in } D, \\
&-\nu_T \frac{\partial T}{\partial z} = Q \text{ on } \Gamma_S \times (t_0, t_1), \qquad \frac{\partial T}{\partial n} = 0 \text{ on } \Gamma_{w,c} \times (t_0, t_1), \\
&\bar{U}_n^{(-)} T + \frac{\partial T}{\partial n} = Q_T \text{ on } \Gamma_{w,op} \times (t_0, t_1), \\
&\frac{\partial T}{\partial n} = 0 \text{ on } \Gamma_H \times (t_0, t_1),
\end{aligned} \tag{59}$$

where $T = T(x, y, z, t)$ is an unknown temperature function, $t \in (t_0, t_1)$, $(x, y, z) \in D = \Omega \times (0, H)$, $\Omega \subset R^2$, $H = H(x, y)$ is the function of the bottom relief, $Q = Q(x, y, t)$ is the total heat flux, $\bar{U} = (u, v, w)$, $\widehat{a}_T = \mathrm{diag}((a_T)_{ii})$, $(a_T)_{11} = (a_T)_{22} = \mu_T$, $(a_T)_{33} = \nu_T$, $f_T = f_T(x, y, z, t)$ are given functions. The boundary of the domain $\Gamma \equiv \partial D$ is represented as a union of four disjoint parts Γ_S, $\Gamma_{w,op}$, $\Gamma_{w,c}$, Γ_H, where $\Gamma_S = \Omega$ (the unperturbed sea surface), $\Gamma_{w,op}$ is the liquid (open) part of vertical lateral boundary, $\Gamma_{w,c}$ is the solid part of the vertical lateral boundary, Γ_H is the sea bottom, $\bar{U}_n^{(-)} = (|\bar{U}_n| - \bar{U}_n)/2$, and $\bar{U}_n$ is the normal component of $\bar{U}$. The other notations and a detailed description of the problem statement can be found in [37].

Problem (59) can be written in the form of an operator equation:

$$\begin{aligned}
T_t + LT &= \mathcal{F} + BQ, \quad t \in (t_0, t_1), \\
T &= T_0, \qquad t = t_0,
\end{aligned} \tag{60}$$

where the equality is understood in the weak sense, namely,

$$(T_t, \widehat{T}) + (LT, \widehat{T}) = \mathcal{F}(\widehat{T}) + (BQ, \widehat{T}) \ \ \forall \widehat{T} \in W_2^1(D), \tag{61}$$

in this case L, $\mathcal{F}$, B are defined by the following relations:

$$(LT, \widehat{T}) \equiv \int_D (-T\mathrm{Div}(\bar{U}\widehat{T}))dD + \int_{\Gamma_{w,op}} \bar{U}_n^{(+)} T\widehat{T} d\Gamma + \int_D \widehat{a}_T \mathrm{Grad}(T) \cdot \mathrm{Grad}(\widehat{T}) dD,$$

$$\mathcal{F}(\widehat{T}) = \int_{\Gamma_{w,op}} Q_T \widehat{T} d\Gamma + \int_D f_T \widehat{T} dD, \quad (T_t, \widehat{T}) = \int_D T_t \widehat{T} dD, \quad (BQ, \widehat{T}) = \int_\Omega Q\widehat{T}\big|_{z=0} d\Omega,$$

and the functions $\widehat{a}_T$, Q_T, f_T, Q are such that equality (61) makes sense. The properties of the operator L were studied by [37].

Due to (61), Eq. (60) is considered in $Y = L_2(t_0, t_1; (W_2^1(D))^*)$, and the operator $B : L_2(\Omega \times (t_0, t_1)) \to Y$ maps the function $Q \in L_2(\Omega \times (t_0, t_1))$ into the function $BQ \in Y$ such that $(BQ, \widehat{T}) = \int_\Omega Q\widehat{T}\big|_{z=0} d\Omega,\ \forall \widehat{T} \in W_2^1(D)$. Problem (59) is linear in T, Q, however, written in the form (60), it is a particular case of the original problem (1), and all the reasoning and the methodology presented in Sections 2–3 are easily transferred to the case of problem (60), understood in a weak sense.

We consider the data assimilation problem for the sea surface temperature (see [37]). Suppose that the functions $Q \in L_2(\Omega \times (t_0, t_1))$ and $T_0 \in L_2(D)$ are unknown in problem (59). Let also $T_{obs}(x, y, t) \in L_2(\Omega \times (t_0, t_1))$ be the function on Ω obtained for $t \in (t_0, t_1)$ by processing the observation data, and this function in its physical sense is an approximation to the surface temperature function on Ω, i.e. to $T\big|_{z=0}$. We admit the case when $T_{\rm obs}$ is defined only on some subset of $\Omega \times (t_0, t_1)$ and denote the indicator (characteristic) function of this set by m_0. For definiteness sake, we assume that T_{obs} is zero outside this subset.

Consider the data assimilation problem for the surface temperature in the following form: find T_0 and Q such that

$$\left\{\begin{array}{rcl} T_t + LT & = & \mathcal{F} + BQ \quad \text{in } D \times (t_0, t_1), \\ T & = & T_0, \qquad t = t_0 \\ J(T_0, Q) & = & \inf\limits_{w,v} J(w, v), \end{array}\right. \tag{62}$$

where

$$J(T_0, Q) = \frac{\alpha}{2}\int_{t_0}^{t_1}\int_\Omega |Q - Q^{(0)}|^2 d\Omega dt + \frac{\beta}{2}\int_D |T_0 - T^{(0)}|^2 dD + \frac{1}{2}\int_{t_0}^{t_1}\int_\Omega m_0 |T\big|_{z=0} - T_{obs}|^2 d\Omega dt, \tag{63}$$

and $Q^{(0)} = Q^{(0)}(x, y, t)$, $T^{(0)} = T^{(0)}(x, y, z)$ are given functions, $\alpha, \beta = const > 0$.

Lemma 2. *For $\alpha, \beta > 0$ the variational data assimilation problem (62) has a unique solution.*

Proof. Let T_0^n, Q^n be a sequence minimizing $J(T_0, Q)$, i.e. $J(T_0^n, Q^n) \to \inf_{T_0, Q} J(T_0, Q)$, $n \to \infty$. Since

$$J(T_0, Q) \ge \frac{\alpha}{2}\int_{t_0}^{t_1}\int_\Omega |Q - Q^{(0)}|^2 d\Omega dt + \frac{\beta}{2}\int_D |T_0 - T^{(0)}|^2 dD, \ \forall T_0 \in L_2(D), Q \in L_2(\Omega \times (t_0, t_1)),$$

then, for $\alpha, \beta > 0$, the sequence T_0^n, Q^n is bounded: $\|T_0^n\|_{L_2(D)} \le const$, $\|Q^n\|_{L_2(\Omega\times(t_0,t_1))} \le const$. Hence, there exists a weakly convergent subsequence (we denote it also by T_0^n, Q^n). The Hilbert spaces $L_2(D)$ and $L_2(\Omega \times (t_0, t_1))$ are weakly closed, therefore, there exist elements $T_0 \in L_2(D), Q \in L_2(\Omega \times (t_0, t_1))$ such that $T_0^n \to T_0$ weakly in $L_2(D)$, and $Q^n \to Q$ weakly in $L_2(\Omega \times (t_0, t_1))$, i.e. $(T_0^n, p)_{L_2(D)} \to (T_0, p)_{L_2(D)}$, $(Q^n, q)_{L_2(\Omega\times(t_0,t_1))} \to (Q, q)_{L_2(\Omega\times(t_0,t_1))} \forall p \in L_2(D), q \in L_2(\Omega \times (t_0, t_1))$. Let T^n and T be the solutions of problem (60) for T_0^n, Q^n and T_0, Q, respectively. Then, for the difference we have

$$\begin{array}{c} (T^n - T)_t + L(T^n - T) = B(Q^n - Q), \quad t \in (t_0, t_1), \\ T^n - T = T_0^n - T_0, \qquad t = t_0. \end{array} \tag{64}$$

The solution to problem (64) continuously depends on the initial value $T_0^n - T_0$ and the flux $Q^n - Q$ (*a priori* estimates are valid in the corresponding functional spaces) [1], therefore, $T^n \to T$ weakly in Y, and $T^n|_{z=0} \to T|_{z=0}$ weakly in $L_2(\Omega \times (t_0, t_1))$. The functional $S(\cdot) = \|\cdot\|^2$ is known [1] to be lower semi-continuous in the weak topology, then $\lim \inf J(T_0^n, Q^n) \ge J(T_0, Q)$, and, therefore, $\inf_{w,v} J(w, v) \ge J(T_0, Q)$. Hence, $\inf_{w,v} J(w, v) = J(T_0, Q)$, that is, T_0, Q gets the minimum to the functional J. This proves the lemma.

The optimality system determining the solution of the formulated variational data assimilation problem according to the necessary condition grad$J = 0$ has the form:

$$\begin{array}{c} T_t + LT = \mathcal{F} + BQ \quad \text{in } D \times (t_0, t_1), \\ T = T_0, \qquad t = t_0, \end{array} \tag{65}$$

$$\begin{array}{c} -(T^*)_t + L^*T^* = Bm_0(T_{\rm obs} - T) \text{ in } D \times (t_0, t_1), \\ T^* = 0, \qquad t = t_1, \end{array} \tag{66}$$

$$\alpha(Q - Q^{(0)}) - T^* = 0 \quad \text{on } \Omega \times (t_0, t_1), \tag{67}$$

$$\beta(T_0 - T^{(0)}) - T^*\big|_{t=0} = 0 \quad \text{in } D, \tag{68}$$

where L^* is the operator adjoint to L.

Here the boundary-value function Q plays the role of λ from Section 2, $\varphi = T$, the operator F has the form $F(T, Q) = -LT + BQ$, and $F'_T = -L$, $F'_Q = B$. Since the operator $F(T, Q)$ is linear in this case and $F''_{TT} = F''_{QQ} = F''_{QT} = 0$, the Hessian $\mathcal{H}$ acting on some $U = (w, \psi)^T$, $w \in L_2(D)$, $\psi \in L_2(\Omega \times (t_0, t_1))$ is defined by the successive solution of the following problems:

$$\left\{\begin{array}{rcl} \dfrac{\partial \phi}{\partial t} + L\phi & = & B\psi, \quad t \in (t_0, t_1) \\ \phi\big|_{t=t_0} & = & w, \end{array}\right. \tag{69}$$

Please cite this article as: V. Shutyaev, F.-X. Le Dimet and E. Parmuzin, Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation, Journal of Computational and Applied Mathematics (2019) 112368, https://doi.org/10.1016/j.cam.2019.112368.

$$\begin{cases} -\dfrac{\partial \phi^*}{\partial t} + L^*\phi^* &= -Bm_0\phi, \quad t \in (t_0, t_1) \\ \phi^*\big|_{t=t_1} &= 0, \end{cases} \tag{70}$$

$$\mathcal{H}U = (\beta w - \phi^*\big|_{t=0}, \alpha\psi - B^*\phi^*)^T. \tag{71}$$

To illustrate the above-presented theory, we consider the problem of sensitivity of functionals of the optimal solution Q to the observations T_{obs}. Let us introduce the following response function:

$$G(T) = \int_{t_0}^{t_1} dt \int_{\Omega} k(x, y, t)T(x, y, 0, t)d\Omega, \tag{72}$$

where $k(x, y, t)$ is a weight function related to the temperature field on the sea surface $z = 0$. For example, if we are interested in the mean temperature of a specific region of the sea ω for $z = 0$ in the interval $\bar{t} - \tau \le t \le \bar{t}$, then as k we take the function

$$k(x, y, t) = \begin{cases} 1\Big/(\tau \text{mes } \omega) & \text{if } (x, y) \in \omega,\ \bar{t} - \tau \le t \le \bar{t} \\ 0 & \text{else,} \end{cases} \tag{73}$$

where mes ω denotes the area of the region ω. Thus, the functional (72) is written in the form:

$$G(T) = \frac{1}{\tau}\int_{\bar{t}-\tau}^{\bar{t}} dt \left(\frac{1}{\text{mes } \omega} \int_{\omega} T(x, y, 0, t)d\Omega \right). \tag{74}$$

Formula (74) represents the mean temperature averaged over the time interval $\bar{t} - \tau \le t \le \bar{t}$ for a given region ω. The response functions of this type are of most interest in the theory of climate change [34,38].

In our notations the functional (72) may be written as

$$G(T) = \int_{t_0}^{t_1} (Bk, T)dt = (Bk, T)_Y, \quad Y = L_2(D \times (t_0, t_1)).$$

We are interested in the sensitivity of the response function $G(T)$, obtained for T after data assimilation, with respect to the observation function T_{obs}.

By definition, the sensitivity is given by the gradient of G with respect to T_{obs}:

$$\frac{dG}{dT_{obs}} = \frac{\partial G}{\partial T}\frac{\partial T}{\partial T_{obs}}. \tag{75}$$

Since $\dfrac{\partial G}{\partial T} = Bk$, then according to the theory presented in Section 5, to compute the gradient (75) we need to perform the following steps:

(1) For k defined by (73) solve the adjoint problem

$$\begin{cases} -\dfrac{\partial \tilde{\phi}^*}{\partial t} + L^*\tilde{\phi}^* &= Bk, \quad t \in (t_0, t_1) \\ \tilde{\phi}^*\big|_{t=t_1} &= 0 \end{cases} \tag{76}$$

and put $\Phi = (\tilde{\phi}^*\big|_{t=0}, B^*\tilde{\phi}^*)^T$.

(2) Find $U = (w, v)^T$ by solving $\mathcal{H}U = \Phi$ with the Hessian defined by (69)–(71).

(3) Solve the direct problem

$$\begin{cases} \dfrac{\partial P_2}{\partial t} + LP_2 &= Bv, \quad t \in (t_0, t_1) \\ P_2\big|_{t=t_0} &= w. \end{cases} \tag{77}$$

(4) Compute the gradient of the response function as

$$\frac{dG}{dT_{obs}} = m_0 P_2\big|_{z=0}. \tag{78}$$

The last formula allows us to estimate the sensitivity of the functionals related to the mean temperature after data assimilation, with respect to the observations on the sea surface.

For numerical experiments have used the three-dimensional numerical model of the Baltic Sea hydrothermodynamics developed at the INM RAS on the base of the splitting method [39] and supplied with the assimilation procedure [37] for the surface temperature T_{obs} with the aim to reconstruct the heat fluxes Q and the initial state T_0.

The parameters of the considered domain of the Baltic Sea and its geographic coordinates can be described as follows: σ-grid is 336 × 394 × 25 (the latitude, longitude, and depth, respectively). The first point of the "grid C" [39] has the coordinates 9.406° E and 53.64° N. The mesh sizes in x and y are constant and equal to 0.0625 and 0.03125 degrees. The

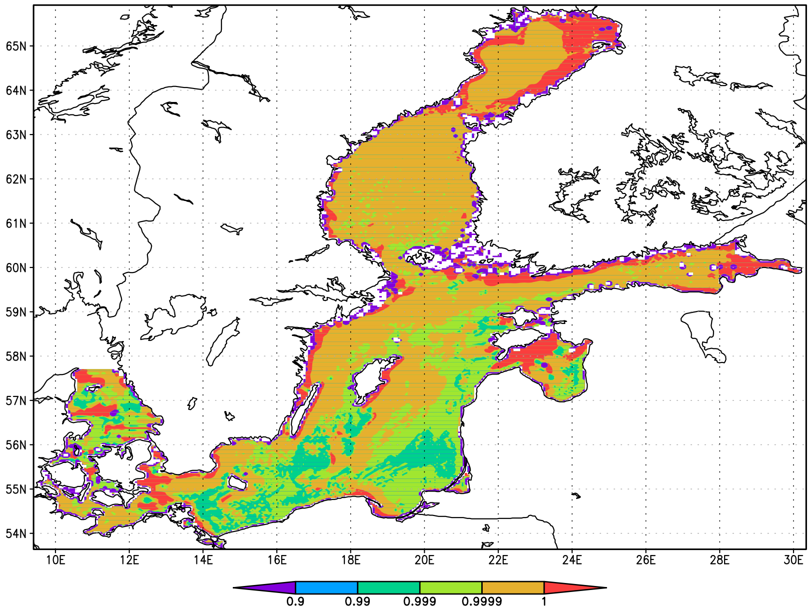

Fig. 1. The gradient of the response function $G(T)$.

time step is $\Delta t = 5$ minutes. The assimilation procedure worked only during some time windows. To start the assimilation procedure, the function $T^{(0)}$ was taken as a model forecast for the previous time interval.

The Baltic Sea daily-averaged nighttime surface temperature data were used for T_{obs}. These are the data of the Danish Meteorological Institute based on measurements of radiometers (AVHRR, AATSR and AMSRE) and spectroradiometers (SEVIRI and MODIS) [40]. Data interpolation algorithms were used [41] to convert observations on computational grid of the numerical model of the Baltic Sea thermodynamics. The mean climatic flux obtained from the NCEP (National Center for Environmental Prediction) reanalysis was taken for $Q^{(0)}$.

Using the hydrothermodynamics model mentioned above, which is supplied with the assimilation procedure for the surface temperature T_{obs}, we have performed calculations for the Baltic Sea area where the assimilation algorithm worked only at certain time moments t_0; in this case $t_1 = t_0 + \Delta t$. The aim of the experiment was the numerical study of the sensitivity of functionals of the optimal solution Q, T_0 to observation errors in the interval (t_0, t_1).

We use the discretize-then-optimize approach, and for numerical experiments all the presented equations are understood in a discrete form, as finite-dimensional analogues of the corresponding problems, obtained after approximation. This allows us to consider model equations as a perfect model, with no approximation errors.

Let us present some results of numerical experiments.

The calculation results for t_0 = 41 h 40 min (500 time steps for the model) are presented in Fig. 1 showing the gradient of the response function $G(T)$ defined by (74) and related to the mean temperature after data assimilation, with respect to the observations on the sea surface, according to (76)–(78). Here $\omega = \Omega$, $\tau = \Delta t$, $\bar{t} = t_1$, $\alpha = \beta = 10^{-5}$.

We can see the sub-areas (in red) in which the response function $G(T)$ is most sensitive to errors in the observations during assimilation. The largest values of the gradient of $G(T)$ correspond to the points x, y lying near the boundary of the domain. This result is confirmed by the direct computation of the response function $G(T)$ according to (74) obtained after assimilation, by introducing perturbations into the observation data T_{obs}.

The above studies allow to determine the sea sub-areas in which the response function related to the optimal solution is most sensitive to errors in the observations during variational data assimilation.

6. Conclusions

Numerical algorithms are considered to study the sensitivity of functionals of the optimal solution of variational data assimilation problem aimed at the reconstruction of unknown parameters and initial state of the model. The optimal solution obtained as a result of assimilation depends on the observations that may contain uncertainties. Computing the gradient of the functionals with respect to observations reduces to the solution of a non-standard problem which is a coupled system involving direct and adjoint equations with mutually dependent variables. Solvability of the non-standard problem is related to the properties of the Hessian of the original cost function. An algorithm to compute the gradient of the response function is developed. Numerical example for variational data assimilation problem related to sea surface temperature for the Baltic Sea thermodynamics model demonstrates the result of the gradient computation of

Please cite this article as: V. Shutyaev, F.-X. Le Dimet and E. Parmuzin, Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation, Journal of Computational and Applied Mathematics (2019) 112368, https://doi.org/10.1016/j.cam.2019.112368.

the response function associated with the mean surface temperature. The presented algorithm may be used to determine the sea sub-areas in which the response functions of the optimal solution are most sensitive to errors in the observations during variational data assimilation.

Acknowledgments

This work was carried out within Russian Science Foundation project 19-71-20035 (studies in Sections 2–3), and the project 18-01-00267 of the Russian Foundation for the Basic Research. The authors are greatly thankful to the reviewers for providing valuable comments that helped to improve the paper.

Appendix. Proof of Theorem 1

Consider the system of perturbations (16)–(19). We have here 4 problems and 4 unknowns $\delta u, \delta\lambda, \delta\varphi, \delta\varphi^*$. The function $\delta\varphi_{obs}$ is included in the right-hand side of (17), therefore, all the unknowns $\delta u, \delta\lambda, \delta\varphi, \delta\varphi^*$ will depend on $\delta\varphi_{obs}$. The expression (20) involves $\delta\varphi_{obs}$ in the left-hand side, and we would like to represent the right-hand side of (20) through $\delta\varphi_{obs}$ also, to obtain the explicit formula for the gradient $\frac{dG}{d\varphi_{obs}}$. Let us introduce four adjoint variables $P_1 \in Y$, $P_2 \in Y$, $P_3 \in Y_p$ and $P_4 \in X$. By taking the inner product of (16) by P_1, (17) by P_2, (19) by P_3 and of (18) by P_4 and adding them, we get:

$$\left(\frac{\partial\delta\varphi}{\partial t} - F'_\varphi(\varphi,\lambda)\delta\varphi - F'_\lambda(\varphi,\lambda)\delta\lambda, P_1\right)_Y + \left(-\frac{\partial\delta\varphi^*}{\partial t} - (F'_\varphi(\varphi,\lambda))^*\delta\varphi^* - (F''_{\varphi\varphi}(\varphi,\lambda)\delta\varphi)^*\varphi^* - \right.$$
$$\left. - (F''_{\varphi\lambda}(\varphi,\lambda)\delta\lambda)^*\varphi^* + C^*V_3(C\delta\varphi - \delta\varphi_{obs}), P_2\right)_Y + \left(V_2\delta\lambda - (F''_{\lambda\varphi}(\varphi,\lambda)\delta\varphi)^*\varphi^* - \right.$$
$$\left. - (F''_{\lambda\lambda}(\varphi,\lambda)\delta\lambda)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*\delta\varphi^*, P_3\right)_{Y_p} + \left(V_1\delta u - \delta\varphi^*\big|_{t=0}, P_4\right)_X = 0.$$

Using integration by parts and adjoint operators, we obtain

$$\left(\delta\varphi, -\frac{\partial P_1}{\partial t} - (F'_\varphi(\varphi,\lambda))^*P_1 - (F''_{\varphi\varphi}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\varphi}(\varphi,\lambda)P_3)^*\varphi^* + C^*V_3CP_2\right)_Y +$$
$$+ \left(\delta\varphi\big|_{t=T}, P_1\big|_{t=T}\right)_X - \left(\delta u, P_1\big|_{t=0}\right)_X + \left(\delta\varphi^*, \frac{\partial P_2}{\partial t} - F'_\varphi(\varphi,\lambda)P_2 - F'_\lambda(\varphi,\lambda)P_3\right)_Y +$$
$$+ \left(\delta\varphi^*\big|_{t=0}, P_2\big|_{t=0}\right)_X + \left(\delta\lambda, V_2P_3 - (F''_{\varphi\lambda}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)P_3)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*P_1\right)_{Y_p}$$
$$- \left(\delta\varphi_{obs}, V_3CP_2\right)_{Y_{obs}} + \left(\delta u, V_1P_4\right)_X - \left(\delta\varphi^*\big|_{t=0}, P_4\right)_X = 0. \tag{79}$$

Hence,

$$\left(-\frac{\partial P_1}{\partial t} - (F'_\varphi(\varphi,\lambda))^*P_1 - (F''_{\varphi\varphi}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\varphi}(\varphi,\lambda)P_3)^*\varphi^* + C^*V_3CP_2, \delta\varphi\right)_Y +$$
$$+ \left(V_2P_3 - (F''_{\varphi\lambda}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)P_3)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*P_1, \delta\lambda\right)_{Y_p} +$$
$$+ \left(V_1P_4 - P_1|_{t=0}, \delta u\right)_X + \left(P_1\big|_{t=T}, \delta\varphi\big|_{t=T}\right)_X + \left(\frac{\partial P_2}{\partial t} - F'_\varphi(\varphi,\lambda)P_2 - F'_\lambda(\varphi,\lambda)P_3, \delta\varphi^*\right)_Y +$$
$$+ \left(P_2\big|_{t=0} - P_4, \delta\varphi^*\big|_{t=0}\right)_X = \left(V_3CP_2, \delta\varphi_{obs}\right)_{Y_{obs}}. \tag{80}$$

We would like the first three summands in the left-hand side of (80) be equal to the right-hand side of (20), keeping the others summands to be zero, therefore, we put

$$-\frac{\partial P_1}{\partial t} - (F'_\varphi(\varphi,\lambda))^*P_1 - (F''_{\varphi\varphi}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\varphi}(\varphi,\lambda)P_3)^*\varphi^* + C^*V_3CP_2 = \frac{\partial G}{\partial\varphi},$$
$$V_1P_4 - P_1\big|_{t=0} = \frac{\partial G}{\partial u},$$

and

$$V_2P_3 - (F''_{\varphi\lambda}(\varphi,\lambda)P_2)^*\varphi^* - (F''_{\lambda\lambda}(\varphi,\lambda)P_3)^*\varphi^* - (F'_\lambda(\varphi,\lambda))^*P_1 = \frac{\partial G}{\partial\lambda},\ P_1\big|_{t=T} = 0,$$

$$\frac{\partial P_2}{\partial t} - F'_\varphi(\varphi,\lambda)P_2 - F'_\lambda(\varphi,\lambda)P_3 = 0,\ \ P_2\Big|_{t=0} - P_4 = 0.$$

Thus, if P_1, P_2, P_3, P_4 are the solutions of (21)–(24), we get from (80):

$$\left(\frac{\partial G}{\partial \varphi}, \delta\varphi\right)_Y + \left(\frac{\partial G}{\partial \lambda}, \delta\lambda\right)_{Y_p} + \left(\frac{\partial G}{\partial u}, \delta u\right)_X = \left(V_3 C P_2, \delta\varphi_{obs}\right)_{Y_{obs}},$$

and due to (20) the gradient of G is given by (25). The theorem is proved.

References

[1] J.-L. Lions, Contrôle Optimal des Systèmes Gouvernés par des Équations aux Dérivées Partielles, Dunod, Paris, France, 1968.
[2] F.-X. Le Dimet, O. Talagrand, Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus 38A (1986) 97–110.
[3] M. Asch, M. Bocquet, M. Nodet, Data Assimilation: Methods, Algorithms, and Applications, SIAM, Philadelphia, USA, 2016.
[4] S.J. Fletcher, Data Assimilation for the Geosciences: from Theory to Application, Elsevier, Amsterdam, Netherlands, 2017.
[5] E. Blayo, M. Bocquet, E. Cosme, L.F. Cugliandolo (Eds.), Advanced Data Assimilation for Geosciences, in: Lecture Notes of the Les Houches School of Physics, Oxford University Press, Oxford, UK, 2014.
[6] A. Carrassi, M. Bocquet, L. Bertino, G. Evensen, Data assimilation in the geosciences: An overview of methods, issues, and perspectives, WIREs Clim. Change 9 (5) (2018).
[7] N.L. Baker, R. Daley, Observation and background adjoint sensitivity in the adaptive observation-targeting problem, Q. J. R. Meteorol. Soc. 126 (565) (2000) 1431–1454.
[8] F.-X. Le Dimet, I.M. Navon, D.N. Daescu, Second-order information in data assimilation, Mon. Weather Rev. 130 (3) (2002) 629–648.
[9] F.-X. Le Dimet, H.E. Ngodock, B. Luong, J. Verron, Sensitivity analysis in variational data assimilation, J. Meteorol. Soc. 75 (1B) (1997) 245–255.
[10] D.N. Daescu, On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation, Mon. Weather Rev. 136 (8) (2008) 3050–3065.
[11] A. Cioaca, A. Sandu, E. de Sturler, Efficient methods for computing observation impact in 4D-Var data assimilation, Comput. Geosci. 17 (6) (2013) 975–990.
[12] I. Gejadze, F.-X. Le Dimet, V. Shutyaev, On analysis error covariances in variational data assimilation, SIAM J. Sci. Comput. 30 (4) (2008) 1847–1874.
[13] I.Yu. Gejadze, G.J.M. Copeland, F.-X. Le Dimet, V.P. Shutyaev, Computation of the analysis error covariance in variational data assimilation problems with nonlinear dynamics, J. Comput. Phys. 230 (2011) 7923–7943.
[14] I.Yu. Gejadze, V.P. Shutyaev, On gauss-verifiability of optimal solutions in variational data assimilation problems with nonlinear dynamics, J. Comput. Phys. 280 (2015) 439–456.
[15] I.Yu. Gejadze, V.P. Shutyaev, F.-X. Le Dimet, Analysis error covariance versus posterior covariance in variational data assimilation, Q. J. R. Meteorol. Soc. 139 (2013) 1826–1841.
[16] V.P. Shutyaev, F.-X. Le Dimet, I.Yu. Gejadze, G.J.M. Copeland, Optimal solution error covariance in highly nonlinear problems of variational data assimilation, Nonlinear Processes Geophys. 19 (2012) 177–184.
[17] V. Shutyaev, F.-X. Le Dimet, E. Shubina, Sensitivity with respect to observations in variational data assimilation, Russ. J. Numer. Anal. Math. Modell. 32 (1) (2017) 61–71.
[18] V.P. Shutyaev, F.-X. Le Dimet, E.I. Parmuzin, Sensitivity analysis with respect to observations in variational data assimilation for parameter estimation, Nonlinear Processes Geophys. 25 (2018) 429–439.
[19] O.M. Alifanov, E.A. Artyukhin, S.V. Rumyantsev, Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems, Begell House Publishers, Danbury, USA, 1996.
[20] N.-Z. Sun, Inverse Problems in Groundwater Modeling, Kluwer, Dordrecht, Netherlands, 1994.
[21] Y. Zhu, I.M. Navon, Impact of parameter estimation on the performance of the FSU global spectral model using its full-physics adjoint, Mon. Weather Rev. 127 (1999) 1497–1517.
[22] R.B. Storch, L.C.G. Pimentel, H.R.B. Orlande, Identification of atmospheric boundary layer parameters by inverse problem, Atmos. Environ. 41 (7) (2007) 1417–1425.
[23] L.W. White, B. Vieux, D. Armand, F.-X. Le Dimet, Estimation of optimal parameters for a surface hydrology model, Adv. Water Resour. 26 (3) (2003) 337–348.
[24] M. Bocquet, Parameter-field estimation for atmospheric dispersion: application to the Chernobyl accident using 4D-Var, Q. J. R. Meteorol. Soc. 138 (2012) 664–681.
[25] S. Schirber, D. Klocke, R. Pincus, J. Quaas, J.L. Anderson, Parameter estimation using data assimilation in an atmospheric general circulation model: From a perfect toward the real world, J. Adv. Modell. Earth Syst. 5 (2013) 58–70.
[26] V.I. Agoshkov, E.I. Parmuzin, V.B. Zalesny, V.P. Shutyaev, N.B. Zakharova, A.V. Gusev, Variational assimilation of observation data in the mathematical model of the Baltic Sea dynamics, Russ. J. Numer. Anal. Math. Modell. 30 (4) (2015) 203–212.
[27] W. Yuepeng, Ch. Yue, I.M. Navon, G. Yuanhong, Parameter identification techniques applied to an environmental pollution model, J. Indust. Manage. Optim. 14 (2) (2018) 817–831.
[28] V.I. Agoshkov, T.O. Sheloput, The study and numerical solution of some inverse problems in simulation of hydrophysical fields in water areas with ‘liquid’ boundaries, Russ. J. Numer. Anal. Math. Modell. 32 (3) (2017) 147–164.
[29] D.P. Dee, Bias and data assimilation, Q. J. R. Meteorol. Soc. 131 (2005) 3323–3343.
[30] P.J. Smith, G.D. Thornhill, S.L. Dance, A.S. Lawless, D.C. Mason, N.K. Nichols, Data assimilation for state and parameter estimation: application to morphodynamic modelling, Q. J. R. Meteorol. Soc. 139 (2013) 314–327.
[31] R. Herzog, I. Riedell, D. Uciński, Optimal sensor placement for joint parameter and state estimation problems in large-scale dynamical systems with applications to thermo-mechanics, Opt. Eng. 19 (2018) 591–627.
[32] G. Chavent, Local stability of the output least square parameter estimation technique, Math. Appl. Comp. 2 (1983) 3–22.
[33] I.M. Navon, Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, Dyn. Atmos. Oceans 27 (1998) 55–79.
[34] G.I. Marchuk, V.I. Agoshkov, V.P. Shutyaev, Adjoint Equations and Perturbation Algorithms in Nonlinear Problems, CRC Press Inc., New York, USA, 1996.
[35] M.M. Vainberg, Variational Methods for the Study of Nonlinear Operators, Holden-Day, San Francisco, USA, 1964.

Please cite this article as: V. Shutyaev, F.-X. Le Dimet and E. Parmuzin, Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation, Journal of Computational and Applied Mathematics (2019) 112368, https://doi.org/10.1016/j.cam.2019.112368.

[36] G.I. Marchuk, V.P. Dymnikov, V.B. Zalesny, Mathematical Models in Geophysical Hydrodynamics and Numerical Methods for Their Realization, Gidrometeoizdat, Leningrad, USSR, 1987.
[37] V.I. Agoshkov, E.I. Parmuzin, V.P. Shutyaev, Numerical algorithm of variational assimilation of the ocean surface temperature data, Comp. Math. Math. Phys. 48 (2008) 1371–1391.
[38] G.I. Marchuk, Adjoint Equations and Analysis of Complex Systems, Kluwer, Dordrecht, Netherlands, 1995.
[39] V. Zalesny, V. Agoshkov, R. Aps, V. Shutyaev, A. Zayachkovskiy, F. Goerlandt, P. Kujala, Numerical modeling of marine circulation, pollution assessment and optimal ship routes, J. Mar. Sci. Eng. 5 (3) (2017) 1–20, 27.
[40] I. Karagali, J. Hoyer, C.B. Hasager, SST diurnal variability in the North Sea and the Baltic Sea, Remote Sens. Environ. 121 (2012) 159–170.
[41] N.B. Zakharova, V.I. Agoshkov, E.I. Parmuzin, The new method of ARGO buoys system observation data interpolation, Russ. J. Numer. Anal. Math. Modell. 28 (1) (2013) 67–84.

Please cite this article as: V. Shutyaev, F.-X. Le Dimet and E. Parmuzin, Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation, Journal of Computational and Applied Mathematics (2019) 112368, https://doi.org/10.1016/j.cam.2019.112368.