



**HAL**  
open science

# Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model

I. Gejadze, Pierre-Olivier Malaterre

## ► To cite this version:

I. Gejadze, Pierre-Olivier Malaterre. Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model. *International Journal for Numerical Methods in Fluids*, 2017, 83 (5), pp.405-430. 10.1002/fld.4273 . hal-01720342

**HAL Id: hal-01720342**

**<https://hal.science/hal-01720342>**

Submitted on 1 Mar 2018

**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.

# Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model

I. Yu. GEJADZE <sup>1</sup>, P.-O. MALATERRE

UMR G-EAU, IRSTEA-Montpellier, 361 Rue J.F. Breton, BP 5095, 34196, Montpellier, France.

**Abstract.** Estimating river discharge from *in-situ* and/or remote sensing data is a key issue for evaluation of water balance at local and global scales and for water management. Variational data assimilation (DA) is a powerful approach used in operational weather and ocean forecasting, which can also be used in this context. A distinctive feature of the river discharge estimation problem is the likely presence of significant uncertainty in principal parameters of a hydraulic model, such as bathymetry and friction, which have to be included into the control vector alongside the discharge. However, the conventional variational DA method being used for solving such extended problems often fails. This happens because the control vector iterates (i.e. approximations arising in the course of minimization) result into hydraulic states not supported by the model. In this paper we suggest a novel version of the variational DA method specially designed for solving estimation-under-uncertainty problems, which is based on the ideas of iterative regularization.

The method is implemented with SIC<sup>2</sup>, which is a full Saint-Venant based 1D-network model. The SIC<sup>2</sup> software is widely used by research, consultant and industrial communities for modelling river, irrigation canal and drainage network behavior. The adjoint model required for variational DA is obtained by means of automatic differentiation. This is likely to be the first stable consistent adjoint of the 1D-network model of a commercial status in existence.

The DA problems considered in this paper are offtake/tributary estimation under uncertainty in the cross-device parameters, and inflow discharge estimation under uncertainty in the bathymetry defining parameters and the friction coefficient. Numerical tests have been designed to understand identifiability of discharge given uncertainty in bathymetry and friction. The developed methodology and software seem useful in the context of the future SWOT satellite mission.

*Keywords:* 1D hydraulic network model, Saint-Venant equations, SIC<sup>2</sup>, variational data assimilation, estimation under uncertainty, adjoint problem, automatic differentiation

## 1 Introduction

Hydraulic and hydrological modelling is important for understanding and monitoring the fresh water cycle, local and trans-boundary management in flood and drought context and evaluation of water balance in the global scale [8]. Information about river discharges plays a key role in this modelling. Unfortunately, *in-situ* measurements of water elevation and discharge are relatively rare on most rivers because of limited accessibility and associated costs. However, the future satellite missions, such as SWOT mission (Surface Water and Ocean Topography, to be launched in 2020, see <http://smc.cnes.fr/SWOT/index.htm>, <https://swot.jpl.nasa.gov/mission/>), will be able to provide a complete 2D map of water levels, water longitudinal slope and width with

---

<sup>1</sup>Corresponding author. Email address: [igor.gejadze@irstea.fr](mailto:igor.gejadze@irstea.fr)

a precision which seems sufficient for the river discharge estimation purpose [35]. Proper use of this type of data, either in combination with *in-situ* data or without it, is an important scientific and practical challenge.

For most rivers a 1D-network model based on the Saint-Venant equations (also referred to as 1.5D model) offers a good compromise between physical and computational complexity levels. It supports a set of important hydraulic phenomena (tide, backwater, artificial structures, control devices), at the same time allowing relatively inexpensive computations which may cover long-term behavior of an entire river system. Of course, for irrigation canals, drainage and sewage systems, the 1D-network models represent the major modelling tool. A number of highly developed models of this type are currently available, such as: Carima (Artelia, France), Mascaret (EDF, France), Mike11 (DHI Water and Environment, Denmark), DufLOW and Sobek (Delft, Netherlands), Isis (Halcrow, UK), Infoworks (HR Wallingford, UK), HEC-RAS (US corps of Engineers, USA), to mention a few. These models are widely used within academic, industrial and consultant communities. The 1D-network model SIC<sup>2</sup> (Simulation and Integration of Control for Canals) has been under development at IRSTEA (previously CEMAGREF) for about 30 years, succeeding the former CEMAGREF hydraulic models (Talweg-Fluvia-Sirene). An innovative feature of SIC<sup>2</sup> as compared to the models mentioned above is its adaptation to irrigation systems and to management and control issues. Currently SIC<sup>2</sup> has about 200 registered users around the globe.

The methods of data assimilation (DA) have become an important tool for analysis of complex physical phenomena in various fields of science and technology. These methods allow us to combine mathematical models, data resulting from instrumental measurements and prior information. The problems of variational DA can be formulated as optimal control problems (e.g. [26, 25]) to find unknown model parameters such as initial and/or boundary conditions, right-hand sides in the model equations (forcing terms), distributed and/or lumped coefficients. The solution method usually implies iterative minimization of the cost function. Equivalently, variational DA can be considered as a special case of the maximum a-posteriori probability estimator (MAP) in the Bayesian framework. For some time, variational DA (in the form of 'incremental 4D-Var', see [7]) has been the only feasible method for solving high-dimensional data assimilation problems in meteorology and oceanography. It still remains a preferable method for operational forecasting (also, in the form of 'ensemble 4D-Var'), even though ensemble Kalman filtering (EnKF, see [15]) is becoming increasingly popular.

The problem considered in this paper is the discharge estimation problem for river or canal network. A distinctive feature of this problem is the likely presence of significant uncertainty in model parameters, such as bathymetry, friction or in those defining behavior of hydraulic structures. Indeed, properties of many rivers are known with quite a limited precision, and even for once well-studied rivers they may evolve in time due to erosion, sedimentation or structures being damaged. This uncertainty, if not taken into account, could degrade the estimated discharge accuracy very noticeably. The estimation-under-uncertainty problems are solved here by including the uncertainty-bearing functions into the control vector. We shall call it a *composite control vector* to underline that it combines elements of a very different nature and weight. We found that with such control vector the conventional variational DA method often fails. This usually happens because some of the control vector iterates constructed during the minimization process may initialize a hydraulic scenario which results in 'unsupported' hydraulic state (locally dry bed, for example). In this case the model execution stops with a

critical error message and the iterative process must be somehow re-launched. In this paper we suggest a novel version of the variational DA approach which allows this difficulty to be alleviated in the framework of unconstrained minimization.

A major technical difficulty with variational DA is that it requires the *adjoint* model to be constructed. The advantage of EnKF is that the adjoint concept is not involved, which partly explains its more frequent use within the engineering community. The EnKF may also require far less computational time for producing a single current state estimate. Yet, there are good reasons for considering variational approach as a DA method useful with the 1D-network hydraulic model. Firstly, this model can certainly be a high-dimensional model. Secondly, the flow behavior in the river or canal network is largely defined by the boundary conditions and source terms, i.e. the problem of estimating the time-dependent controls could be of a primary interest here. For this type of problems variational approach is superior to EnKF in terms of the estimation quality, though a comparable quality can possibly be achieved by using the Ensemble Kalman Smoother (EnKS, see [16, 5]). Thirdly, variational DA is a convenient framework for analysis of identifiability, which is a key issue for problems with a *composite* control vector [36, 6]. This issue can be investigated via the eigenvalue decomposition of the Hessian, see e.g. [42, 19]. In this paper we report on a successful construction of the adjoint of SIC<sup>2</sup>, which is, to the best of our knowledge, the first consistent adjoint of the full complexity 1D-network hydraulic model of an commercial status in existence. This result has been achieved by applying the automatic differentiation (AD) engine TAPENADE [21] to a slightly modified version of the existing SIC<sup>2</sup> software.

References on data assimilation with the full Saint-Venant network models are scarce. There are only a few papers reporting on the use of EnKF in this context [14, 44], whereas this method is used more frequently with different simplified hydraulic models, see e.g. [13], [3] or [35], or with 2D models [41]. Concerning variational DA, there are examples for idealized setups (single reach configuration) [39, 2, 12], or for simplified (linearized) Saint-Venant model [45]. The only paper reporting on the variational control for the full Saint-Venant network model we can refer is [11]. The adjoint model in the latter reference is derived analytically, then implemented numerically ('optimize-then-discretize' approach, which yields an 'inconsistent' adjoint). However, none of the full Saint-Venant commercial models listed at the beginning of this section has the variational DA option. This is largely due to the difficulty of constructing adjoint models. One reason is a complexity of the 1D-network models, which is higher than the one of 2D models. An example of the 2D shallow water model for which the adjoint has been generated by means of AD is DASSFLOW, see <http://www.math.univ-toulouse.fr/DassFlow>. Another difficulty is related to the implicit (or semi-implicit) time discretization schemes commonly used in Saint-Venant solvers.

In this paper two different DA problems with SIC<sup>2</sup> have been considered: the offtake estimation problem under uncertainty in the cross-device properties and the discharge estimation problem under uncertainty in parameters defining the canal bathymetry and friction. The observations include dense-in-time surface elevation measurements. The problems have been solved by the novel version of the variational DA approach, whereas the conventional approach has failed in most cases considered. Numerical tests have been designed to understand the identifiability issue.

The paper is organized as follows. A brief mathematical statement of the 1D-network model is given in Sect.2, and the DA problem statement in Sect.3. Distinctive features of

the DA problem under consideration and the solution method which takes into account these features are presented in Sect.4. The most important implementation details are presented in Sect.5 and the numerical results for two test problems in Sect.6. The main results of the paper are summarized in the Conclusions. Appendix I provides a brief description of the SIC<sup>2</sup> model and Appendix II contains the list of notations.

## 2 Model statement

The hydraulic network is represented by a set of closed-line segments or 'reaches' connected at nodes  $N_k$ , see e.g. Fig.1. In reality 'reach' is a continuous extent water, especially a stretch of a river or a canal between two bends, oftakes/tributaries or structures (weirs, locks, mills). The spatial discretization along reach number  $i$  produces a set of co-ordinates (also called longitudinal abscissas)  $x_{i,j}$ ,  $j = 1, \dots, \bar{K}_i$ , where  $\bar{K}_i$  is the set dimension. Each  $x_{i,j}$  has the associated global index  $k$  and its own position vector  $\vec{r}_k = (x'_k, y'_k, z'_k)$  in the global co-ordinate system (bathymetry). Given  $\vec{n}_k$  is a pre-dominant flow direction at  $x_{i,j}$ , a hydraulic cross-section  $S_{i,j}$  is defined by a set of points on a plane  $\vec{n}_k \cdot (\vec{r} - \vec{r}_k) = 0$  describing the bed profile, which are evaluated from a design sketch or from a topographical survey. For each section this data allows the wetted area  $A(Z, p_g)$ , the wetted perimeter  $P(Z, p_g)$ , the hydraulic radius  $R(Z, p_g)$  and the top width  $L(Z, p_g)$  to be computed for any given water level line  $Z$ . Here  $p_g$  are geometric parameters of the corresponding computational cross-section. For a given reach,  $p_g$  are functions of the longitudinal abscissa.

For a 'regular' section, the shallow water flow in the longitudinal direction  $x$  is described by the Saint-Venant equations:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = Q_L, \quad (2.1)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial Q^2/A}{\partial x} + gA \frac{\partial Z}{\partial x} = -gAS_f + C_L Q_L v, \quad (2.2)$$

$$t \in (0, T),$$

where  $Q(x, t)$  is the discharge,  $Z(x, t)$  is the water level,  $A(Z(x, t), p_g(x))$  is the wetted area,  $v(x, t) = Q/A$  is the mean velocity,  $S_f$  is the friction term dependent on the Strickler coefficient  $C_s(x)$  and on the wetted perimeter  $P(Z, p_g)$ ,  $Q_L(x, t)$  is the lateral discharge and  $C_L(x)$  is the lateral discharge coefficient. The initial condition for equations (2.1)-(2.2) is

$$Z(x, 0) = Z_0(x), \quad Q(x, 0) = Q_0(x). \quad (2.3)$$

For an internal node we consider the mass balance equation alongside the condition of local elevations or 'heads' ( $h = v^2/2g + Z$ ) equality, for all connected reaches. On the network example presented in Fig.1 these equations are

$$q_1 = -Q|_{S_{1,k1}} - Q|_{S_{2,k2}} + Q|_{S_{3,1}}, \quad (2.4)$$

$$Z|_{S_{1,k1}} = Z|_{S_{3,1}}, \quad Z|_{S_{2,k2}} = Z|_{S_{3,1}}, \quad (2.5)$$

or

$$h|_{S_{1,k1}} = h_{S_{3,1}}, \quad h|_{S_{2,k2}} = h_{S_{3,1}}, \quad (2.6)$$

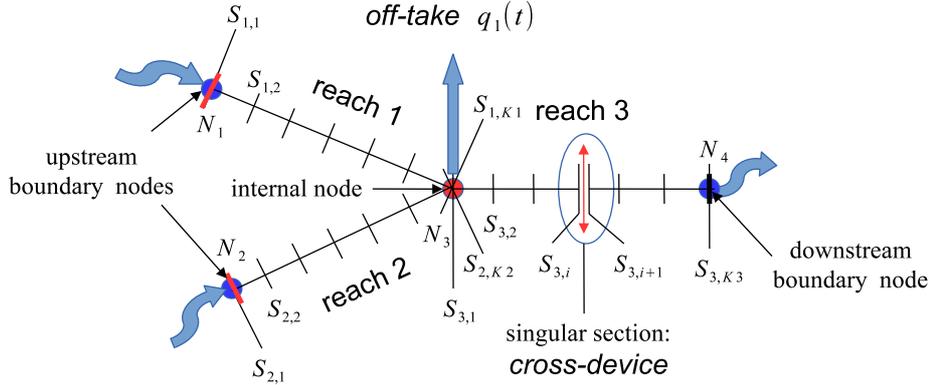


Figure 1: River or canal network conceptual scheme

where  $q_1$  is the offtake or tributary at node  $N_3$ .

Boundary conditions are defined at boundary nodes. For the upstream nodes we usually use the inflow discharge  $Q(t)$  or elevation  $Z(t)$ , for example <sup>2</sup>:

$$Q(t)|_{S_{1,1}} = Q_1(t) \quad \vee \quad Z(t)|_{S_{1,1}} = Z_1(t), \quad (2.7)$$

whereas for the downstream nodes it is the elevation  $Z(t)$  or the rating curve  $Q = f(Z, p_{rc})$ , where  $p_{rc}$  are the rating curve parameters, for example:

$$Z(t)|_{S_{3,k3}} = Z_3(t), \quad \vee \quad Q|_{S_{3,k3}} = f(Z|_{S_{3,k3}}, p_{rc}). \quad (2.8)$$

We also consider a singular section (which consists of the collocated upstream and downstream sections), where the Saint-Venant equations are no longer valid and are replaced by other equations. It is mainly used to represent artificial structures (cross-devices), such as gates, weirs, bridges etc., but it can also be used to describe an abruptly changing natural bathymetry. For a singular section we consider the mass balance equation alongside the equation relating the elevations (or 'heads'), for example:

$$Q|_{S_{3,i}} - Q|_{S_{3,i+1}} = 0 \quad (2.9)$$

$$Q|_{S_{3,i}} = \mathcal{F}(Z|_{S_{3,i}}, Z|_{S_{3,i+1}}, C_d|_{S_{3,i}}), \quad (2.10)$$

where  $C_d$  is the cross-device discharge coefficient or the expansion head loss coefficient.

Let  $\mathcal{U}$  be a space of the input variables of the model (2.1)-(2.10). Let us also assume that the specified network configuration includes  $K_r$  reaches,  $K_{bn}$  boundary nodes,  $K_s$  sections,  $K_{ss}$  singular sections and  $K_{in}$  internal nodes where the part of the network can be cut off and substituted by sink/source terms (offtakes/tributaries). Then, the full control vector  $U \in \mathcal{U}$  looks as follows:

$$U = (Z_0, Q_0, Z, Q, q, Q_L, C_s, C_k, C_d, p_{rc}, p_g)^T, \quad (2.11)$$

<sup>2</sup>In the text below  $\vee$  stands for logical 'or' and  $\wedge$  for logical 'and'

where by  $(Z_0, Q_0)$  we mean a set of initial conditions for all reaches, i.e.

$$(Z_0, Q_0) = \{(Z_{0,k}(x), Q_{0,k}(x)), k = 1, \dots, K_r\},$$

by  $(\mathcal{Z}, \mathcal{Q})$  - a set of inflow discharges or elevations at all boundary nodes, i.e.

$$(\mathcal{Z}, \mathcal{Q}) = \{(\mathcal{Z}_k(t) \vee \mathcal{Q}_k(t)), k = 1, \dots, K_{bn}\}$$

by  $q$  - a set of all offtakes/tributaries, i.e.  $q = \{q_k(t), k = 1, \dots, K_{in}\}$ , by  $p_g$  - a set of geometric parameters for all sections, i.e.  $p_g = \{p_{g,k}, k = 1, \dots, K_s\}$ , and, similarly, for the remaining variables in (2.11), each having its own dimension. For given  $U$ , by solving the model equations (2.1)-(2.10) simultaneously for all network reaches, we obtain the flow fields  $(Z, Q) = \{(Z_k(x, t), Q_k(x, t)), i = 1, \dots, K_r\}$ ,  $(Z, Q) \in \mathcal{X}$ , where  $\mathcal{X}$  is the state space.

Let us note that the general description presented above does not cover all the variety of hydraulic options implemented in SIC<sup>2</sup>, but it is sufficient to demonstrate our results. The basic features of this model are presented in Appendix I.

### 3 Data assimilation problem statement

For modelling the behavior of a hydraulic network the 'true' input vector (2.11) must be specified. In reality, some components of  $U$  contain uncertainties  $\xi_b$ . Thus, instead of  $U$  we use its best available guess

$$U_b = U + \xi_b, \quad (3.12)$$

which is known as a 'background' in variational DA or as a 'prior' in statistics. Because of the presence of  $\xi_b$ , the flow field  $(Z, Q)$  evaluated on  $U_b$  also contains an error.

Let us assume that we observe the state variables in the form

$$Y = C(Z, Q) \in \mathcal{Y}, \quad (3.13)$$

where  $\mathcal{Y}$  is the observation space, and  $C : \mathcal{X} \rightarrow \mathcal{Y}$  is the observation operator. For example, the water surface elevation measured by the gauge stations, located at the specified sections of the specified reaches, may be available. We shall denote by  $I_o$  the array defining the indices of these reaches and sections. Usually, such measurements are recorded with sufficiently small time step, so we can treat them as nearly continuous in time. Then the observation operator is as follows:

$$Y = C(Z, Q) = \{Z(t)|_{s_{i,j}}, (i, j) \in I_o\}.$$

Now we can define the control-to-observation map

$$\mathcal{R}(U) = Y \quad (3.14)$$

where  $\mathcal{R} : \mathcal{U} \rightarrow \mathcal{Y}$  is a nonlinear operator. The actual observations usually contain noise (observation uncertainty), i.e.

$$Y^* = Y + \xi_o. \quad (3.15)$$

The aim of data assimilation is to use observations  $Y^*$  to improve the prior  $U_b$  and, subsequently, the model solution for  $(Z, Q)$ .

In the Bayesian framework the posterior probability density of  $U$  conditioned on observations  $Y^*$  is given by the Bayes formula

$$p(U|Y^*) = \frac{p(Y^*|U)p(U)}{p(Y^*)}. \quad (3.16)$$

Looking for the mode of the posterior density  $p(U|Y^*)$ , i.e. maximizing  $p(U|Y^*)$ , is the essence of variational DA. Under a Gaussian assumption on the prior and observation uncertainties, i.e.  $\xi_b \sim N(0, \mathcal{B})$ ,  $\xi_o \sim N(0, O)$ , where  $\mathcal{B}$  is the background error covariance and  $O$  - the observation error covariance, maximizing  $p(U|Y^*)$  is equivalent to minimizing the cost function

$$J(U) = \frac{1}{2} \|O^{-1/2}(\mathcal{R}(U) - Y^*)\|^2 + \frac{1}{2} \|\mathcal{B}^{-1/2}(U - U_b)\|^2.$$

The covariance  $\mathcal{B}$  is block-diagonal, with each block corresponding to a component of  $U$  in (2.11). A method for generating those blocks is discussed in detail in Sect.5.4. For a given sensor the observation error is considered as an uncorrelated time series, thus the covariance  $O$  is diagonal with elements  $\sigma_O^2(t)$ .

The above expression is the conventional 4D-Var cost function as it is understood in meteorology and oceanography [7, 25]. Let us note that in these applications the initial condition is normally used as control. This is perfectly justifiable given the huge inertia of atmospheric and oceanic flows. Since the initial condition is a state of the system, asymptotic solutions of the model equations can be used for defining  $\mathcal{B}$ , see [9]. In our case the inertia of the system is relatively small, so the influence of the initial condition on the flow is limited. Instead, the flow is governed by the boundary conditions and source terms. For these types of controls it is only possible to specify a simplified  $\mathcal{B}$  using, for example, hydrological modelling. The same is true for the spatially distributed model coefficients  $p_g$  or  $C_s$ . Here one should consider the expert opinion on the accuracy of topological surveying. Since defining sufficiently accurate  $\mathcal{B}$  seems unreliable we consider the cost function in the form

$$J(U, \alpha) = \frac{1}{2} \|O^{-1/2}(\mathcal{R}(U) - Y^*)\|^2 + \frac{\alpha}{2} \|\mathcal{B}^{-1/2}(U - U_b)\|^2. \quad (3.17)$$

where  $\alpha > 0$  is a parameter responsible for the relative weight of the residual and penalty terms inside  $J(U, \alpha)$ , to be determined from an auxiliary condition. The above formulation is similar to the one used in the Tikhonov regularization (TR) method [43], where  $\alpha$  is called the 'regularization' parameter<sup>3</sup>.

In practice, the full input vector  $U$  is not used as a control vector in data assimilation. First, the components of  $U$  which do not contain any uncertainty must be excluded. The remaining components should be ranked by their influence on the model solution and the most influential ones may be considered as controls. Since we deal with incomplete observations, the reason for reducing the control vector dimension (degrees of freedom) is quite obvious. Thus, considering  $U$  as a set with elements  $\{U_i\}$ , we define  $V \subset U$ , which consists of a few components of interest, and  $U^0$ , the complement of  $V$  in  $U$ . The simplest approach is to ignore the remaining uncertainty-bearing components in  $U^0$  by fixing their values at the corresponding  $U_b$ . A more

---

<sup>3</sup>TR is the most commonly used method of regularization of ill-posed inverse problems, derived in the deterministic framework

sophisticated approach is to consider these components as 'nuisance' parameters, which implies a different expression for the posterior density (3.16) and, under the Gaussian assumption, a modified cost function.

Taking into account the above considerations we consider the data assimilation problem in the form

$$\hat{V}_\alpha = \underset{V}{\operatorname{argmin}} J(V, \alpha), \quad (3.18)$$

$$J(V, \alpha) = \frac{1}{2} \|O^{-1/2}(\mathcal{R}(V, U^0) - Y^*)\|^2 + \frac{\alpha}{2} \|B^{-1/2}(V - V_b)\|^2, \quad V \subset U, \quad U^0 = U \setminus V, \quad (3.19)$$

where  $V_b \subset U_b$  and  $B$  is a sub-matrix of  $\mathcal{B}$  relevant to  $V$ . The problem (3.18)-(3.19) can be solved by any gradient-based minimization algorithm, for example by a quasi-Newton method in the form:

$$V_{i+1} = V_i + \beta_i H_i^{-1} J'_V(V_i, \alpha), \quad V_0 = V_b, \quad (3.20)$$

where  $H_i^{-1}$  is an inverse Hessian approximation,  $\beta_i$  is the descent step and  $J'_V(V_i, \alpha)$  is the cost function gradient at point  $V_i$  at the  $i^{\text{th}}$ -iteration.

For the operator  $\mathcal{R}(U)$  we define the tangent linear operator  $\mathcal{R}'(U)$  (Gateaux derivative) and its adjoint  $(\mathcal{R}'(U))^*$  [32] as follows:

$$\mathcal{R}'(U)w = \lim_{t \rightarrow 0} \frac{\mathcal{R}(U + tw) - \mathcal{R}(U)}{t}, \quad (3.21)$$

$$(w, (\mathcal{R}'(U))^*w^*)_{\mathcal{U}} = (\mathcal{R}'(U)w, w^*)_{\mathcal{Y}}, \quad w \in \mathcal{U}, \quad w^* \in \mathcal{Y}. \quad (3.22)$$

Given the above operator definitions, the full gradient of  $J(u, \alpha)$  in (3.17) can be expressed in the form:

$$J'_U(U, \alpha) = (\mathcal{R}'(U))^*O^{-1}(\mathcal{R}(U) - Y^*) + \alpha B^{-1}(U - U_b), \quad (3.23)$$

whereas the sought partial gradient is simply a subset of the full gradient

$$J'_V(V, \alpha) \subset J'_U(U, \alpha).$$

Operator  $\mathcal{R}(U)$  is implemented as a FORTRAN subroutine involving the SIC<sup>2</sup> software. Subroutines for the tangent linear  $\mathcal{R}'(U)$  and adjoint  $(\mathcal{R}'(U))^*$  operators are produced by the automatic differentiation engine TAPENADE [21]. Analytic expressions for the tangent linear and adjoint operators for the Saint-Venant equations can also be found in the literature, e.g. [2, 39, 12, 11].

Let us mention that in solving estimation problems involving the 'composite' control vector, i.e. a vector which includes both driving conditions and model parameters, the quasi-Newton method should be preferred to the Gauss-Newton method because it allows the intrinsic nonlinearity of such problems to be better managed. This makes sense unless computing the gradient at each iteration is too costly.

## 4 Solving the DA problem: special features

The DA problem presented in Section 3 has a few special features making it different from the classical DA problems in weather and ocean forecasting. One difference has been already

mentioned: it is a boundary and source term control problem, rather than an initial value control problem. For these controls the background covariance is constructed on a more subjective basis. In order to attenuate the impact of the subjective judgment the regularization parameter  $\alpha$  is introduced, thus allowing a variable influence of the penalty term. Another important feature is the 'composite' nature of the control vector, which includes both the driving conditions and parameters defining basic properties of the system, such as bathymetry, friction, etc.

In order to use the gradient based unconstrained minimization of the cost function (3.19) we assume that the mapping  $\mathcal{R}(V, U^0)$  exists and is continuous everywhere in a vicinity of the reference (true) value  $V_{ref}$ , i.e. the domain  $D(\mathcal{R}) \subseteq \mathcal{U}$  is dense around the point  $(V_{ref} \cup U^0)$ . Then we believe that if  $(V_b \cup U^0) \in D(\mathcal{R})$ , then  $(V_i \cup U^0) \in D(\mathcal{R}), \forall i$ . Here we intuitively assume a 'regular convergence' property, which means that if the iterative process is converging in terms of the residual function, i.e.

$$\|O^{-1/2}(\mathcal{R}(V_{i+1}, U^0) - Y^*)\| \leq \|O^{-1/2}(\mathcal{R}(V_i, U^0) - Y^*)\|, \quad (4.24)$$

then it is also converging in terms of the distance

$$\|V_{i+1} - V_{ref}\| \leq \|V_i - V_{ref}\|. \quad (4.25)$$

In the numerical experiments conducted this convergence pattern has indeed been observed when controlling the driving conditions only, for example the inflow discharge  $Q(t)$ , or parameters only, for example, the Strickler coefficient  $C_s(x)$ .

The situation changes dramatically when a composite control vector is considered, for example, when the inflow discharge  $Q(t)$  and the bathymetry defining parameters  $p_g(x)$  are simultaneously sought. In this case the convergence pattern looks far more erratic. That is, reduction of the residual norm does not necessarily imply reduction in the estimation error norm. It is, therefore, possible that during minimization some combinations of controls may result into the hydraulic conditions which cannot be supported by the model, such as dry bed or supercritical flow. Formally, it means that  $(V_i \cup U^0) \notin D(\mathcal{R})$ .

The reason for such behavior can be explained as follows. The control vector is weakly constrained by the penalty term, but the constraints start to take effect only when the relative weight of the two terms in the cost function becomes comparable. In the course of iterations the residual term value is decreasing and the penalty term value is increasing starting from zero. Thus, at initial iterations the residual term largely dominates the gradient and the control update remains essentially **unconstrained**. At the same time, introducing inequality constraints into the DA formulation looks highly undesirable.

In order to overcome this difficulty we intend to regularize the iterative process itself. The idea comes from the iterative regularization (IT) method [1, 23]<sup>4</sup>, where iterating in a Hilbert space with a stronger than  $L_2$ -norm is suggested. In our setup this can be achieved by using the change of variables

$$V = V_b + B^{1/2}W. \quad (4.26)$$

Then, the DA problem (3.18)-(3.19) can be written as follows:

$$\hat{W}_\alpha = \underset{W}{\operatorname{argmin}} J(W, \alpha), \quad (4.27)$$

---

<sup>4</sup>IR is the method of regularization of ill-posed inverse problems which relies on intrinsic properties of iterative minimization algorithms defined in Hilbert spaces, such that the iteration number serves as the regularization parameter

$$J(W, \alpha) = \frac{1}{2} \|O^{-1/2}(\mathcal{R}(V_b + B^{1/2}W, U^0) - Y^*)\|^2 + \frac{\alpha}{2} \|W\|^2. \quad (4.28)$$

Again, we are looking for  $W$  by a quasi-Newton minimization algorithm

$$W_{i+1} = W_i + \beta_i \tilde{H}_i^{-1} J'_W(W_i, \alpha), \quad W_0 = 0. \quad (4.29)$$

It is obvious that, if  $W$  belongs to the Euclidean space with  $L_2$ -norm,  $V$  belongs to the space with a stronger norm defined by the covariance  $B$  or, in different words,  $V$  always belongs to a compact set restricted by  $B$ , which is the main regularizing condition [43].

Let us note that the DA cost function in the form similar to (4.27)-(4.28) is sometimes used in variational DA (first introduced probably in [10]), however for a totally different reason: for high-dimensional problems it could sometimes be easier (or better substantiated) to define the operator-vector product  $B^{1/2}v$  rather than  $B^{-1/2}v$ . For example, it can be modelled by solving the heat conduction equation [33], or by using a recursive filter [28].

It is easy to show that the expression for the gradient of (4.28) becomes

$$\begin{aligned} J'_W(W, \alpha) &= (B^{1/2})^*(\mathcal{R}'(V_b + B^{1/2}W, U^0))^*O^{-1}(\mathcal{R}(V_b + B^{1/2}W, U^0) - Y^*) + \alpha W \\ &= (B^{1/2})^*J'_V(V, \alpha). \end{aligned}$$

Taking into account the above relationship, the regularized quasi-Newton iterative process for solving (3.18)-(3.19) takes the form:

$$\begin{aligned} W_{i+1} &= W_i + \beta_i \tilde{H}_i^{-1} (B^{1/2})^* J'_V(V_i, \alpha), \quad W_0 = 0, \quad V_0 = V_b, \\ V_{i+1} &= V_b + B^{1/2}W_{i+1}, \\ i &= 0, \dots \end{aligned} \quad (4.30)$$

After the iterations have converged, the solutions by (3.20) and (4.30) coincide because both satisfy the condition  $J'_V(V_i, \alpha) = 0$ . However, all iterates  $V_i = V_b + B^{1/2}W_i$  from the process (4.30) belong to a compact set restricted by  $B$ . Subsequently, the above iterative sequence exhibits a better 'regularity' in the sense (4.24)-(4.25). Let us also mention that, from the point of view of the IR concept, some other options to regularize the iterative process can be considered.

Optimal choice of the regularization parameter  $\alpha$  is the key issue in the Tikhonov regularization method. For example, in the deterministic framework the residual (discrepancy) principle is used [34]. Let us introduce the residual function  $r(W) = \mathcal{R}(V_b + B^{1/2}W, U^0) - Y^*$ . According to the residual principle, the optimal value of  $\alpha$  is the solution to the equation

$$\|r(\hat{W}_\alpha)\| = \Delta, \quad (4.31)$$

where  $\Delta$  is an estimation of the observation uncertainty norm, called the residual threshold. In the probabilistic framework we assume that  $\xi_o \sim N(0, O)$ . Therefore, instead of (4.31), one may consider a statistical hypothesis

$$r(\hat{W}_\alpha) \sim N(0, O),$$

which can be accepted with confidence level  $\gamma$  if

$$r^T(\hat{W}_\alpha)O^{-1}r(\hat{W}_\alpha) \sim \chi^2(m, \gamma), \quad (4.32)$$

where  $m$  is the observation space dimension. For large  $m$  the mode of  $\chi^2(m, \gamma)$  is approaching  $m$ .

The need for solving the problem (4.27)-(4.28) alongside the conditions (4.31) or (4.32) noticeably complicates the solution algorithm and may increase the computational cost. This difficulty can be avoided by using the iterative regularization (IR) method. This method is founded on special properties of some gradient-based algorithms, which is explained below. The solution to the problem (4.27)-(4.28) depends on  $\alpha$ , therefore the residual norm and the solution norm both depend on  $\alpha$ . Excluding the parameter one can build a plot  $\{\|r(\hat{W}_\alpha)\|, \|\hat{W}_\alpha\|\}_{\alpha \in (0, \infty)}$ , which is known as the  $L$ -curve, see [20]. For linear  $\mathcal{R}$  this is a monotonic decreasing function. In the IR method the cost function in the problem (4.27)-(4.28) is considered without the penalty term. If this modified problem is solved by an iterative method, we can obtain the residual norm and the solution norm as functions of the iteration number  $i$ ; subsequently a plot  $\{\|r(W_i)\|, \|W_i\|\}_{i=1, \dots}$ , which is called the discrete  $L$ -curve, can be constructed. The main postulate of the iterative regularization is that for certain family of iterative methods the discrete  $L$ -curve is close to the continuous  $L$ -curve, at least for  $\|r(W_i)\| \geq \Delta$ . This means that, if the iterations have been stopped upon satisfying the criterion

$$\|r(W_i)\| = \Delta, \quad (4.33)$$

then the solution norm  $\|W_i\|$  ought to be close to  $\|\hat{W}_\alpha\|$  satisfying the condition (4.31). Similarly, instead of (4.32) we can use

$$r^T(W_i)O^{-1}r(W_i) \sim \chi^2(m, \gamma). \quad (4.34)$$

In our numerical implementation, for the IR method the gradient of the cost function (3.19) is computed with  $\alpha$  being fixed as an infinitesimal real constant.

## 5 Implementation details

### 5.1 Initial condition treatment

The initial condition  $(Z_0, Q_0)$  is a model state at  $t = 0$ . As such it must be a model solution consistent with the parameters which define the fundamental properties of the model, such as bathymetry, friction, rating curve parameters and cross-device coefficients, and also with the previous values of time-dependent controls. Changing arbitrarily some of those parameters while keeping the initial condition intact leads to severe shocks in the flow fields at the initial time period. Furthermore, the difference between the observations and the model predictions during the initial time period dominates the gradient. The corresponding updates being introduced into the nonlinear system may lead to unsupported flow conditions. Even if the initial condition is consistent with the other parameters at the start of the iterative process, independent updates of time-dependent controls and parameters may lead to inconsistency again.

The way to deal with this issue is as follows. We notice that the influence of the initial state on the flow is very limited in time, then it is dominated by the driving conditions (boundary conditions and source terms). Therefore, we postulate that  $(Z_0, Q_0)$  is a **steady state** flow solution consistent with the initial value of the time-dependent controls and time-independent controls. This state is approached by performing a relaxation model run. By doing so we stop considering  $(Z_0, Q_0)$  as an independent control, but it becomes a unique function of other controls.

## 5.2 Development of the adjoint model

The adjoint model is a code for computing the cost function  $J(U)$  and its gradient  $J'_U(U)$ . It is generated by applying the automatic differentiation (AD) engine TAPENADE to the code which computes the hydraulic state over the time period  $t = 0, \dots, T$  and, given the data  $U_b, Y^*$ , evaluates the cost function. The first step is, therefore, to create a subroutine *forward*( $\Psi, U, (U_b, Y^*), J(U)$ ), which calls *transi* - the main computational routine of SIC<sup>2</sup>. Here  $\Psi$  is the complete input of SIC<sup>2</sup>,  $U$  is the control vector,  $(U_b, Y^*)$  is the data to be assimilated and  $J(U)$  is the cost function. The structure of this subroutine is presented in Fig.2. The SIC<sup>2</sup> model reads all inputs from data files before calling *transi*. Thus, the first operation inside *forward* is mapping  $U$  into the variables involved in this call. At this stage the penalty term is computed. Next, the relaxation is performed by running *transi* for a relatively short period of time  $T^*$  with the value of time dependent controls being fixed at their initial values, for example  $Q(t) = Q(0)$ , etc. At this stage the residual term is not computed. Then, the initial condition is equated to the final state from the relaxation run and *transi* is run again for the full simulation period  $T$ . After (or during) the main run the residual term of  $J(U)$  is computed.

Subroutine *forward* is specified as a 'head' object for processing (differentiation) by TAPENADE, mentioning  $U$  as an input and  $J(U)$  as an output. After processing we obtain subroutine *forward\_b*( $\Psi, U, U^*, (U_b, Y^*), J(U), J^*(U)$ ), where  $U^* := J'_U(U)$ , plus the adjoint counterparts to all subroutines involved with *forward*.

Despite a relatively simple basic concept, the practical implementation of the AD procedure has been a challenging task, since the SIC<sup>2</sup> code was not originally designed having the AD option in view. For successful AD, the connections between the input and output variables should be as transparent as possible. This is not always the case with older software, usually created under severe memory restrictions. For example, all EQUIVALENCE operators have been removed and the memory allocation correspondingly re-arranged. The logic of some iterative procedures used in the solver has also been revised.

## 5.3 Spline approximation of time-dependent controls

The time-dependent controls, such as the inflow discharge  $Q(t)$ , water elevation  $Z(t)$  at boundaries, offtakes/tributaries  $q_i(t)$  and the lateral discharge  $Q_L(x, t)$  are approximated in time by cubic splines. Thus, the control points for the chosen control variable can be arbitrarily distributed in time. Given the values of control at these points the spline coefficients are constructed. The values of the control variable at time instants required for the model numerical integration (usually at  $t = i \times dt$ , where  $dt$  is the integration time step) are evaluated as the corresponding spline values. This allows the number of control nodes to be significantly smaller than the number of integration time steps, which is useful given that the simulation period can be fairly long. Besides, the control nodes can be distributed more densely in the areas of fast dynamics and more sparsely in the areas of slow dynamics. This approach can also be considered as a preliminary regularization. The spline evaluation routine is a part of the block referred to as '*conformity mapping*' in Fig.2 and, as such, is also subjected to the AD.

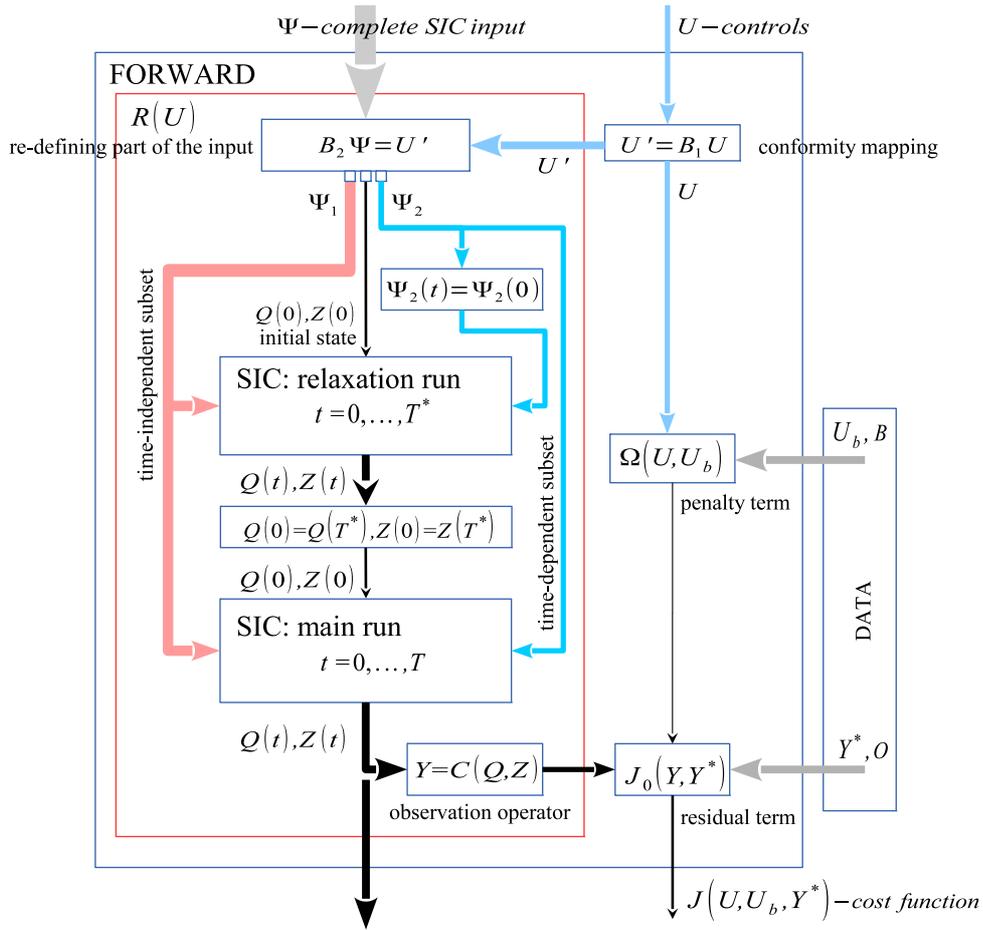


Figure 2: Structure of subroutine *forward*.

## 5.4 Defining the covariance

The controls in  $U$  can be divided into three groups: time dependent controls, e.g. the inflow discharge  $Q(t)$ , spatially distributed controls, e.g. the Strickler coefficient  $C_s(x)$  or bed elevation  $z(x) \in p_g$ , and lumped parameters, e.g. the cross-device discharge coefficients  $C_d$ . For the lumped parameters we can only prescribe the variance, for the distributed controls the covariance matrix must be specified.

Here we present a slightly modified version of the approach described in [19]. In solving ill-posed inverse problems the solution is often considered to be a smooth function which belongs to a Sobolev space of certain order, e.g.  $W_2^2$ . Let  $f(x)$ ,  $x \in (a, b)$  be a one-dimensional function of  $x$  and let us introduce two positive weight functions  $\mu_1(x), \mu_2(x)$ . We define the norm of  $f(x)$  in  $W_2^2$  as follows:

$$\|f(x)\|_{W_2^2[a,b]}^2 = \int_a^b \mu_1(x) f^2(x) + \mu_1(x) \left( \frac{\partial}{\partial x} \left( \mu_2(x) \frac{\partial f(x)}{\partial x} \right) \right)^2 dx.$$

To evaluate this integral numerically we discretize  $f(x)$  using a set of uniformly distributed

nodes  $\{x_i \in \bar{X}, i = 1, \dots, m\}$ , such that  $x_i = (i - 1)\Delta x$  and substitute the integral by the sum

$$\|f(x)\|_{W_2^2[m]}^2 = \Delta x \sum_{i=1}^m \mu_1(x_i) f^2(x_i) + \Delta x \sum_{i=1}^m \mu_1(x_i) \left( \frac{\partial}{\partial x} \left( \mu_2(x) \frac{\partial f(x)}{\partial x} \right) \right)^2 \Big|_{x=x_i}. \quad (5.35)$$

Numerical implementation of the second term depends on the boundary conditions imposed on  $f(x)$ ; in this paper we use the 'natural' boundary conditions, i.e.  $f''(a) = f''(b) = 0$ .

In practice, we consider a discrete function  $\bar{f}(\bar{x}_i)$ , where  $\{\bar{x}_i \in \bar{X}, i = 1, \dots, n\}$  are arbitrarily distributed nodes. Therefore, a mapping  $G : \bar{X} \rightarrow X$  must be constructed. Since we need the second derivative of  $f(x)$ , the cubic spline approximation of  $f(x_i)$  is sufficient. In this case the mapping is nonlinear, i.e.  $G := G(\bar{f})$ . The inverse of the covariance matrix  $B$  must satisfy the following condition

$$\|B^{-1/2} \bar{f}\|_{L_2[n]} \approx \|G(\bar{f})\|_{W_2^2[m]}. \quad (5.36)$$

Assuming that  $\bar{f}$  is reasonably close to the prior  $\bar{f}_b$ , the elements  $B_{i,j}^{-1}$  can be obtained by the following formula:

$$B_{i,j}^{-1} \approx \frac{\partial^2 (\|G(\bar{f})\|_{W_2^2[m]}^2)}{\partial \bar{f}_i \partial \bar{f}_j} \Big|_{\bar{f}=\bar{f}_b}.$$

The code for computing the elements  $B_{i,j}^{-1}$  is obtained by applying automatic differentiation (in the direct mode) twice to the subroutine evaluating  $\|G(\bar{f})\|_{W_2^2[m]}$ . The matrix  $B^{-1}$  is symmetric and narrow-banded. It can be easily factorized using Cholesky decomposition:

$$B^{-1} = B^{-1/2} (B^{-1/2})^T.$$

Given the factor  $B^{-1/2}$ , the product  $v = B^{1/2} w$  is defined via solving the equation  $B^{-1/2} v = w$ . Since  $B^{-1/2}$  is a triangular banded matrix, the solution procedure is simply a backward sweep. In the covariance matrix obtained by this method the functions  $\mu_1(x)$  and  $\mu_2(x)$  define the local variance and correlation radius, respectively. For any time-dependent control, the covariance is generated for the full time domain. It is slightly more complicated for the spatially distributed controls due to the presence of nodes and singular sections.

Let us denote by  $\Omega_1(f)$  and  $\Omega_2(f)$  the first and the second terms in (5.35), respectively. Then, equation (5.36) can be written in the form

$$\|B^{-1/2} \bar{f}\|_{L_2[n]} \approx \Omega_1(G(\bar{f})) + \Omega_2(G(\bar{f})). \quad (5.37)$$

Here  $\Omega_1$  is the normalized Euclidean norm and  $\Omega_2$  is a measure of auto-correlations in  $\bar{f}$ . It is reasonable to suggest that assimilating the observation noise may lead to 'decorrelation', i.e. to a sharp rise of  $\Omega_2$  against  $\Omega_1$ . Thus, for monitoring the data assimilation process we shall consider

$$\epsilon(\bar{f}) = \Omega_2(G(\bar{f})) / \Omega_1(G(\bar{f})). \quad (5.38)$$

For example,  $\epsilon(V - V_b)$  can be used for comparing solutions obtained by different methods: for an equal residual norm  $\|r(V)\|$  one should prefer the solution with the smallest  $\epsilon(V - V_b)$ .

The blocks of matrix  $\mathcal{B}$  for different functions involved in the control vector are generated using the method described above. While the method allows to vary the variance and correlation levels along the abscissa  $x$ , in the numerical examples considered below these are uniform almost everywhere (except near the boundaries). The correlation radius for the time dependent controls has been chosen as  $3h$ , and for the spatially distributed controls as  $1km$ . The background mean deviation values for different functions  $\sigma_B[\cdot]$  are given in Sect. 6.1 and Sect. 6.2.

## 5.5 Miscellaneous

**a) Bathymetry.** In Sect.2 the bathymetry is formally defined by geometric parameters  $p_g$ , entering the functions  $A(Z, p_g)$  and  $P(Z, p_g)$ . Here we present a more detailed description. For each section  $i$ , the elevation  $z(x_i)$  of the lowest point of the cross-section shape with respect to the reference horizontal level is given. The function  $z(x)$  is referred to as the bed elevation. Other parameters describe the cross-section shape itself. The dilation coefficient  $b(x_i)$  is introduced to modify the cross-section shape by scaling its horizontal dimensions. Subsequently, this affects the functions  $A(Z, p_g)$  and  $P(Z, p_g)$ . For example, in the case of trapezoidal shape,  $b(x)$  scales the trapezoid bases. The functions  $z(x)$  and  $b(x)$  are considered as the generalized bathymetry controls.

**b) Identical twin experiment.** In this paper the identical twin experiment approach is adopted: given a reference ('true') value of the control vector  $V_{ref}$ , for a chosen observation scheme the model predictions at the specified points (in space and time) are considered as 'exact' observations; after being corrupted by noise these observations are considered as 'noisy'. The task is to estimate the control vector using either 'exact' or 'noisy' data and to evaluate the estimation error  $dV = \hat{V} - V_{ref}$ .

In estimating boundary conditions involving dissipative models oscillatory solutions are often obtained. The oscillations, however, may have a limited effect on the averaged system behavior. Thus, in order to characterize the discharge estimation error, alongside

$$dQ(t) = \hat{Q}(t) - Q_{ref}(t) \quad (5.39)$$

we shall consider the spatially averaged discharge error

$$\langle dQ(t) \rangle = \left( \frac{1}{K_s} \sum_{i=1}^{K_s} (\hat{Q}(t)|_{S_i} - Q_{ref}(t)|_{S_i})^2 \right)^{1/2}. \quad (5.40)$$

**c) Minimization routine.** For minimization the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (LBFGS) [27] is used, with 7 secant pairs being retained. In all numerical tests conducted within this project utilizing the *unconstrained* minimization has been proved sufficient if applied in combination with the iterative regularization technique suggested in Sect.4. The LBFGS software has been downloaded from:

<http://users.iems.northwestern.edu/~nocedal/lbfgs.html>.

**d) Tables.** Some results of numerical tests are summarized in Tables 2-4. Firstly, this is the convergence information:  $N$  - number of iterations required to obtain an estimate,  $\|r(V_0)\|^2$  - the residual norm at iteration zero, i.e. before data assimilation,  $\|r(V_N)\|^2$  - the residual norm at iteration  $N$ , i.e. after data assimilation. Similarly, the values of the gradient norm  $\|J'_V(V_0)\|^2$  and  $\|J'_V(V_N)\|^2$  are presented. Secondly, this is the decorrelation level in the estimate, expressed by  $\epsilon$ , see (5.38). The latter is important for monitoring and stopping the iterative process. Also, Tab.3 provides  $\delta C_{d,k}$  - the estimation error in the cross-device discharge coefficients.

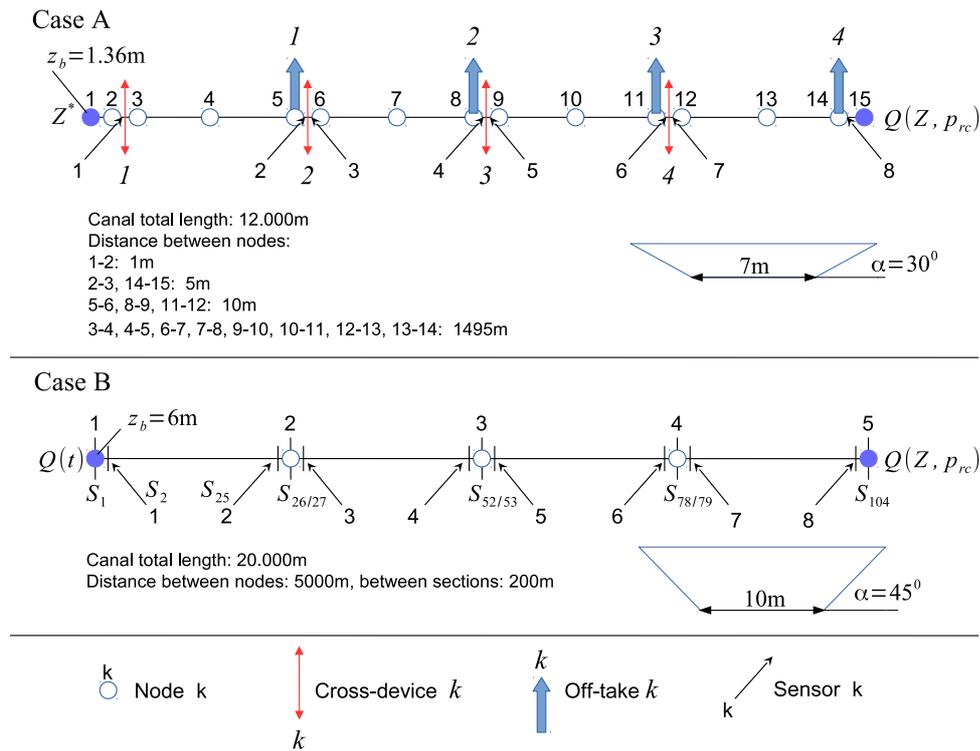


Figure 3: Testing configurations.

## 6 Numerical results

### 6.1 Case A

In this case the offtake estimation problem under uncertainty in the cross-device properties is considered. This problem is important in the framework of the irrigation canal network management.

The computational scheme and the geometry of the canal are presented in Fig.3(A). It consists of four 3km-long reaches. The canal is fed from a large upstream reservoir with a constant water level  $Z = Z^*$ . This provides the inflow discharge of about  $7m^3/sec$ . At the inlet and between reaches there are the cross-devices (of the 'undershot gate' type) which regulate the discharge and the water levels along the canal. The key parameter describing the cross-device hydraulic properties is the discharge coefficient  $C_d$ , see <http://sic.g-eau.net> for details. At the canal outlet the rating curve  $Q(Z, p_{rc})$  is defined. The bed elevation at the canal inlet is  $z = 1.36m$ , at the outlet - zero; the slope is, therefore, about  $11cm/km$ . Water can be taken in via the side canals or by pumps, which are implemented as offtakes  $q_k(t)$  at nodes 5, 8, 11, 14. The offtake functions are approximated by cubic splines constructed on a set of evenly distributed time points with the step  $dt_c = 1h$ . The model integration time step is  $dt = 5min$ .

The task is to estimate functions  $q_i(t)$  using observations of water levels  $Z$  made by stationary sensors. These are located just in front and behind the gates, as presented in the scheme

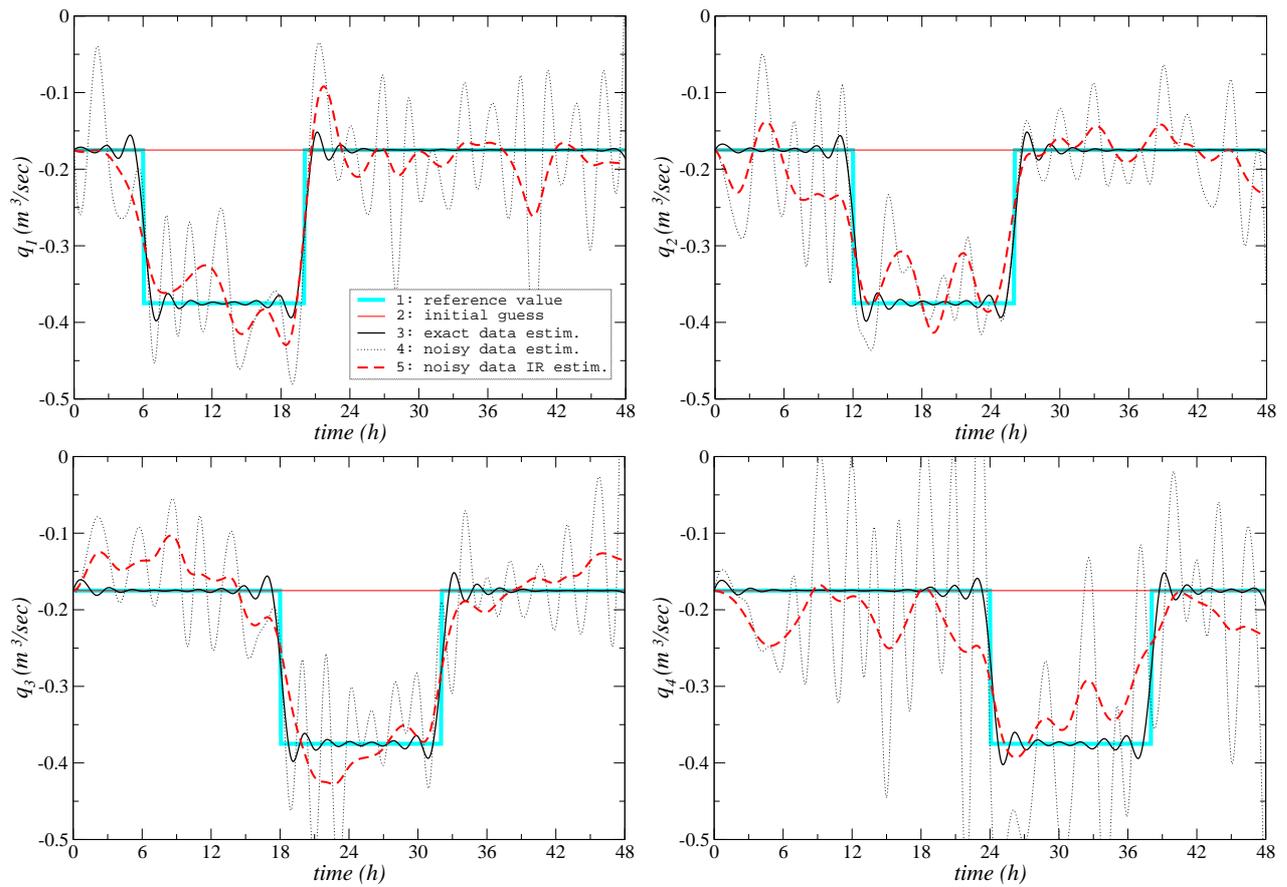


Figure 4: Offtake estimation problem, no uncertainty.

in Fig.3(A). Observations are collected during the time period  $T = 48h$ , the observation time step is  $dt_o = 5min$ . The total number of observations is  $m = 4.609 \cdot 10^3$  (to be used in (4.32)). The observation error mean deviation is  $\sigma_O = 0.01m$ , which is a usual level of the measurement accuracy by a gauge station. The background error mean deviations are  $\sigma_B[q_i] = 0.1m^3/s$  and  $\sigma_B[C_d] = 0.06$ .

In the first series of numerical tests the control vector  $V = (q_1(t), \dots, q_4(t))^T$  is considered, whereas no uncertainty in other model inputs is assumed. The goal here is to investigate the performance of the variational DA algorithm in general, and to compare different regularization techniques. The results are presented in Fig.4 and the relevant auxiliary information in Tab.2.<sup>5</sup> Any subplot in Fig.4 shows: the reference offtake function  $q_k(t)$  in l.1, the initial guess (background)  $q_{b,k}(t)$  in l.2, the estimate  $q_k(t)$  obtained by assimilating **exact** data without regularization in l.3/A1 and the estimate  $q_k(t)$  obtained by assimilating **noisy** data without regularization in l.4/A2. In the last two cases the minimization process has been stopped upon reaching the numerical convergence limit. As expected, the estimate that corresponds to the 'exact' data closely follows the reference function. Small oscillations at the corners of a 'box'-function are due to the chosen control time step ( $dt_c = 1h$ ) in the spline approximation.

<sup>5</sup>In order to refer to a line in the figure and to the corresponding column in the table simultaneously, we shall combine the references. For example, l.3/A1 means: line 3 and table column A1.

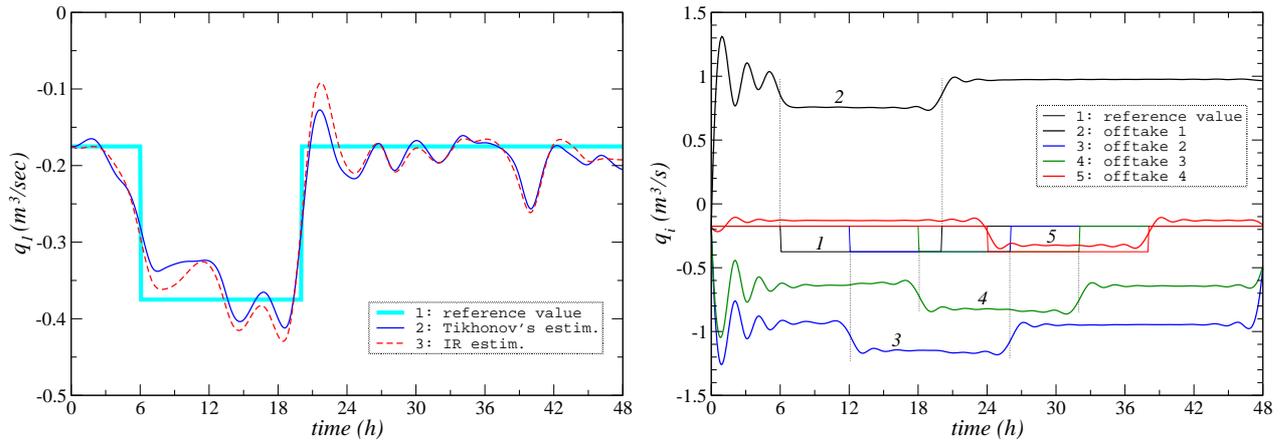


Figure 5: Offtake estimation problem. Left: Tikhonov versus iterative regularization. Right: uncertainty in  $C_d$  ignored.

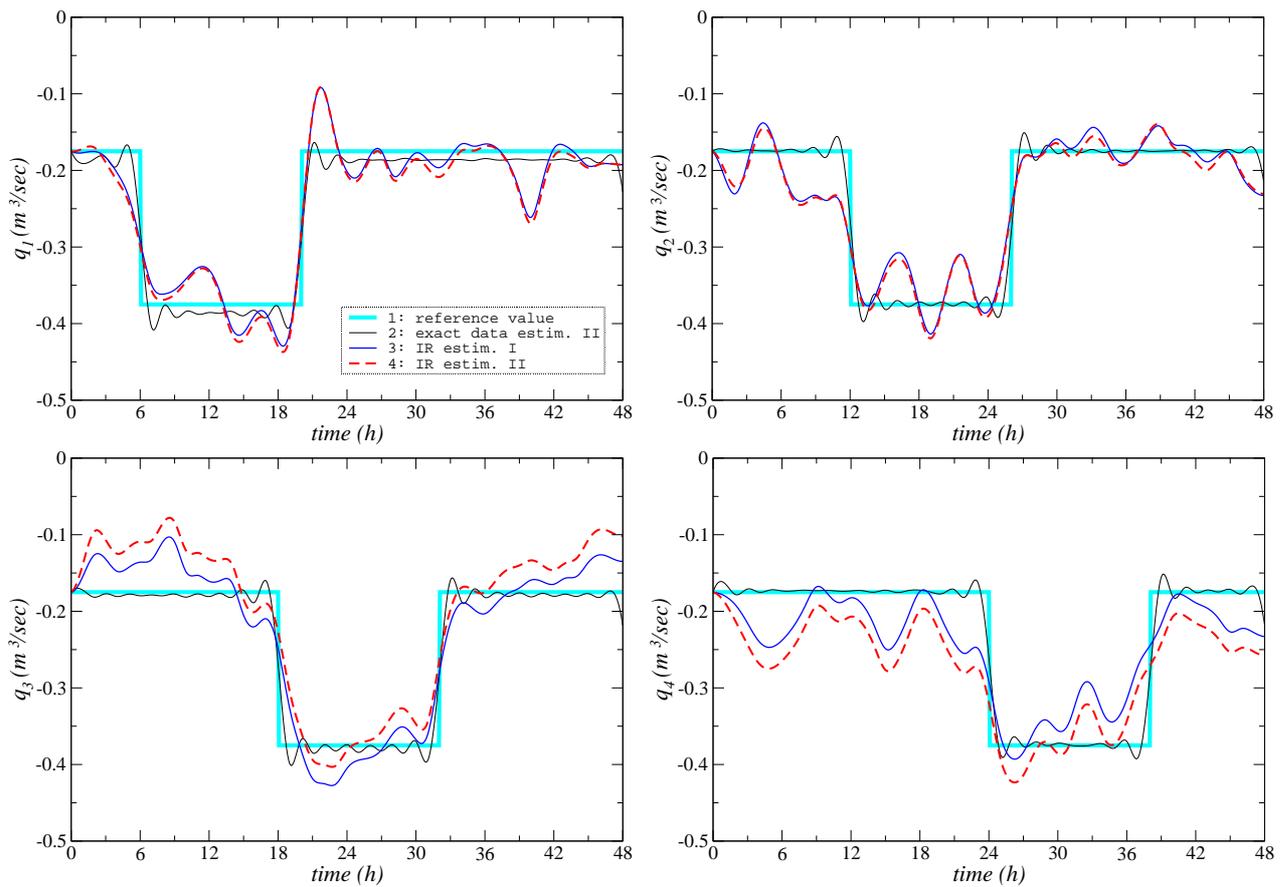


Figure 6: Offtake estimation problem: uncertainty in  $C_d$  taken into account.

However, the estimate that corresponds to the 'noisy' data exhibits quite a strong oscillatory behavior, in particular  $q_4(t)$  in the lower-right corner of Fig.4. This is despite the fact that the spline approximation of the offtakes is used with  $dt_c \gg dt$ , i.e. the solutions in *l.4/A2* are implicitly regularized ('natural' or step regularization). The estimates obtained by assimilating **noisy** data using iterative regularization are presented in *l.5/A4*. Here, the effect of regularization can be clearly seen.

In Fig.5(left) the estimates obtained using the Tikhonov regularization and the iterative regularization methods are displayed. To compare the two methods the values of  $\epsilon$  (5.38) presented in Tab.2 for cases *A3* and *A4* must be compared. One can see that these values are not very different. The same can be seen for the graphs presented in *l.2* and *l.3*, which proves a practical equivalency between the TR and IR in terms of the final result. However, IR has the advantage of the regular convergence. This found to be critically important when dealing with the composite control vector, see Sect.6.2.

In the second series of numerical tests we investigate the offtake estimation problem under uncertainty in the cross-device discharge coefficients  $C_d$ . This type of uncertainty is quite relevant to anthropogenic water distribution systems. Let us assume that the expected value for the properly functioning gate is  $C_d = 0.82$ , but it could take smaller values, should the gate malfunction or be jammed by an external object. For example, we consider the following set of coefficients:  $C_d = 0.82, 0.76, 0.68, 0.82$ , i.e. there is a problem with the second and third gates. However, when solving the DA problem, we assume that all gates are functioning properly. The vector  $V = (q_1(t), \dots, q_4(t))^T$  is estimated by assimilating 'exact' data, without regularization. The results are presented in Fig.5(right). Here we notice that the shape of the 'box'-function is recovered well, but the magnitudes are wrong.

The way to deal with the systematic uncertainties is to include the corresponding variables into the control vector  $V$  (and to remove them from  $U^0$ ). Thus, we consider now  $V = (q_k(t), C_{d,l})^T$ ,  $k, l = 1, \dots, 4$ . The results of data assimilation are presented in Fig.6 and the relevant auxiliary information in Tab.3. Any subplot of Fig.6 shows: the reference offtake function  $q_k(t)$  in *l.1*, and the offtake estimate obtained by assimilating **exact** data without regularization in *l.2/A5*. We notice that the magnitude error which appears in Fig.5(right) has been largely removed. The estimates of  $q_k(t)$  obtained by assimilating **noisy** data using the IR are presented in *l.4/A7*. For comparison, the corresponding estimates from the first test series are presented in *l.3/A6*. The results for  $q_1(t)$  and  $q_2(t)$  look nearly identical, whereas for  $q_3(t)$  and  $q_4(t)$  a relatively small bias can be noticed. This bias exists despite a very small estimation error in  $C_d$ : the largest one is  $\delta C_d \approx 0.4\%$ , see case *A6*. This fact points out a significant sensitivity of the offtake estimates to the errors in  $C_d$ . We also note that the iterative process has converged more slowly: it takes  $N = 63$  against  $N = 22$  iterations to reach about the same level of the residual norm  $\|r(V_N)\|^2 = 4.610 \cdot 10^3$ , i.e. extending the control vector slows down the solution process. Overall, it can be concluded that for the given set of sensors and for the chosen  $\sigma_O$ , resolving several time-dependent controls of a relatively small magnitude (about 3% of the total discharge each) is feasible, even in the presence of uncertainty in  $C_d$ .

## 6.2 Case B

In this case the inflow discharge estimation problem under uncertainties in bathymetry and friction parameters is considered. This is an important problem in studying and monitoring

N	Case	<i>noise</i>	<i>uncertainties</i>	<i>controls</i>
1	B1.1	<i>none</i>	<i>none</i>	$Q(t)$
2	B1.2	$\sigma = 10cm$	<i>none</i>	$Q(t)$
3	B2.1	<i>none</i>	$z(x)$	$Q(t)$
4	B2.2	<i>none</i>	$z(x)$	$Q(t), z(x)$
5	B2.3	$\sigma = 10cm$	$z(x)$	$Q(t), z(x)$
6	B3.1	<i>none</i>	$C_s(x)$	$Q(t)$
7	B3.2	<i>none</i>	$C_s(x)$	$Q(t), C_s(x)$
8	B5.1	<i>none</i>	$b(x)$	$Q(t)$
9	B5.2	<i>none</i>	$b(x)$	$Q(t), b(x)$
10	B4.1	<i>none</i>	$z(x), C_s(x)$	$Q(t)$
11	B4.2	<i>none</i>	$z(x), C_s(x)$	$Q(t), z(x), C_s(x)$
12	B4.3	<i>none</i>	$z(x), C_s(x)$	$Q(t), z(x)$
13	B6.1	<i>none</i>	$z(x), C_s(x), b(x)$	$Q(t)$
14	B6.2	<i>none</i>	$z(x), C_s(x), b(x)$	$Q(t), z(x), C_s(x), b(x)$
15	B6.3	<i>none</i>	$z(x), C_s(x), b(x)$	$Q(t), z(x)$

Table 1: Subcases of case *B*.

the behavior of natural river systems. The canal computational scheme is presented in Fig.3(B). The canal consists of four 5km-long reaches, each being represented by 26 sections. The distance between sections is 200m. The inflow discharge  $Q(t)$  is specified at the canal inlet node, and the rating curve  $Q(Z, p_{rc})$  is given at the outlet node. As before,  $Q(t)$  is approximated by cubic splines constructed on a set of evenly distributed time points with  $dt_c = 1h$ . The reference values of the spatially distributed parameters (as functions of the section number) are presented in *l.1*, Fig.8, right subplots.

The task is to estimate the function  $Q(t)$  using observations of water levels  $Z$ . The observation points are located just in front and behind the internal nodes as presented in the scheme in Fig.3(B). Observations are collected during the time period  $T = 24h$ , the observation time step is  $dt_o = 5min$ . The observation error mean deviation is  $\sigma_O = 0.1m$ , which is an expected error level in measurements recorded by a remote sensing tool. The number of observations is  $m = 2.302 \cdot 10^3$  (to be used in (4.32)). The background error mean deviations are  $\sigma_B[Q(t)] = 2.5m^3/s$ ,  $\sigma_B[z(x)] = 0.15m$ ,  $\sigma_B[C_s(x)] = 2.8m^{-1/3}s^{-1}$  and  $\sigma_B[b(x)] = 0.1$ .

The list of subcases of case *B* is presented in Tab.1, graphical results - in Fig.7-9, and some auxiliary information in Tab.4. It must be emphasized that we have never been successful in solving the DA problems in the classical formulation (3.18)-(3.19) with the composite control vector. Hence, in case *B* all problems have been solved using the modified formulation (4.27)-(4.28). When assimilating the 'exact' data, the minimization process has been stopped upon reaching the numerical convergence limit, whereas for the 'noisy' data - by the criterion (4.33) (in Tab.4 this is referred as  $IR^*$  and  $IR$ , correspondingly).

In the first two tests the inflow discharge is estimated assuming no uncertainty in other parameters. The results are presented in Fig.7. The left subplot shows  $Q_{ref}$  in *l.1*, the initial guess on  $Q(t)$  in *l.2* and its estimated value based on the 'noisy' data - in *l.3*. The one based on the 'exact' data replicates the reference value almost perfectly and it is not displayed. On the

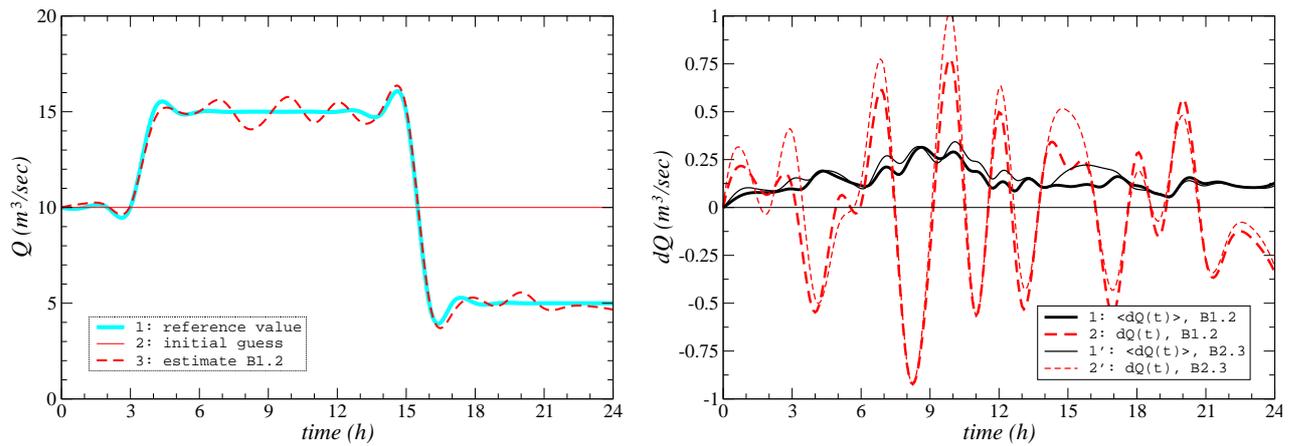


Figure 7: Inflow discharge estimation problem: left - estimate; right - estimation errors.

right subplot the estimation errors  $\langle dQ(t) \rangle$  and  $dQ(t)$  are presented in *l.1* and *l.2*. Assessing these results one may conclude that the inflow discharge has been estimated fairly well (less than 7% relative error for  $dQ(t)$  and 2.5% for  $\langle dQ(t) \rangle$ ).

The inflow discharge is estimated under uncertainty in one spatially distributed function ( $z(x)$  or  $C_s(x)$  or  $b(x)$ ) in tests n.3 – 9. The uncertainty is ignored in cases *B2.1*, *B3.1* and *B5.1*, but it has been taken into account by including the corresponding function into the control vector in cases *B2.2*, *B3.2* and *B5.2*. The results are presented in Fig.8.

Let us consider the upper subplots in Fig.8, for example. Here, the inflow discharge is estimated using an erroneous guess (right subplot, *l.2*) on the actual bed elevation  $z(x)$  (right subplot, *l.1*). While ignoring the uncertainty, one gets the estimation error presented in *l.1* and *l.2*, left subplot. These errors are largely removed if the uncertainty is taken into account, see *l.1'*, *2'*. The corresponding estimate of  $z(x)$  is shown in *l.3*, right subplot. One can see that  $z(x)$  is well estimated. The results relevant to the Strickler coefficient  $C_s(x)$  and to the dilation coefficient  $b(x)$  are presented in the mid and lower subplots, respectively. The comments made previously on cases involving  $z(x)$  are equally valid for  $C_s(x)$  and  $b(x)$ . Assessing these results one may conclude that the composite control vector, which consists of one time-dependent control and one spatially distributed control, can be successfully estimated.

Let us mention that the above results have been obtained using the 'exact' data. It has been noticed that the observation noise (at least uncorrelated) is largely absorbed into the time-dependent control estimates, whereas it has little effect on the spatially distributed or lumped control estimates. To illustrate this statement we consider case *B2.3*, which is equivalent to case *B2.2*, but the 'noisy' data has been assimilated. The corresponding errors  $\langle dQ(t) \rangle$  and  $dQ(t)$  are presented in *l.1'* and *l.2'* in Fig.7, right subplot. One can notice that these errors look very similar to the estimation errors being encountered in case *B1.2*, i.e. without considering uncertainty. At the same time, the estimate of  $z(x)$  in case *B2.3* can hardly be distinguished from the one in case *B2.2*. Therefore, the results in Fig.8 can be extrapolated to the case of 'noisy' data: the errors  $\langle dQ(t) \rangle$  and  $dQ(t)$  should contain additional components similar to those presented in Fig.7, right subplot.

The inflow discharge is estimated under uncertainty in more than one spatially distributed

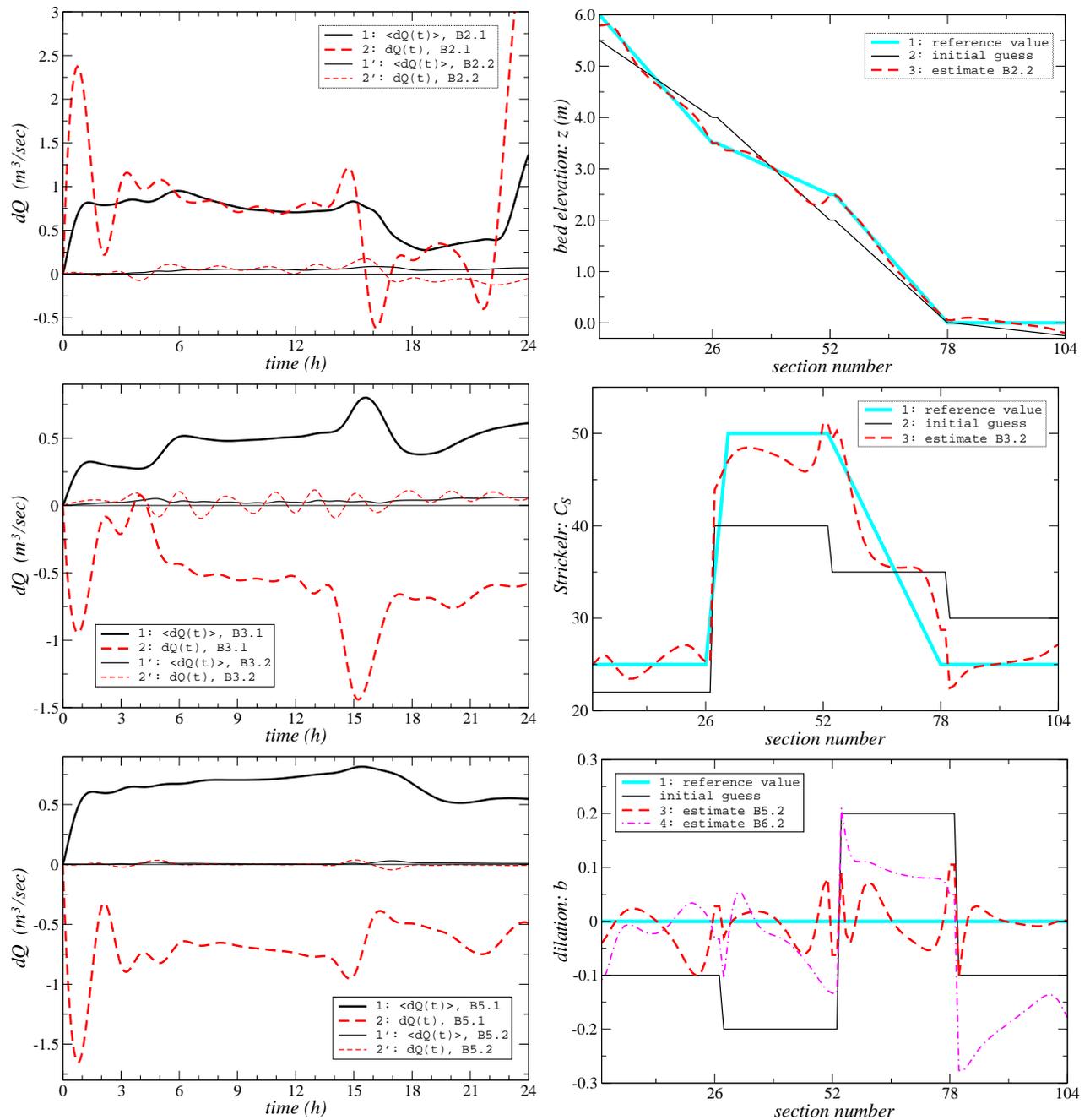


Figure 8: Inflow discharge estimation problem under uncertainty in one spatially distributed function: upper -  $z(x)$ , mid -  $C_s(x)$ , lower -  $b(x)$ .

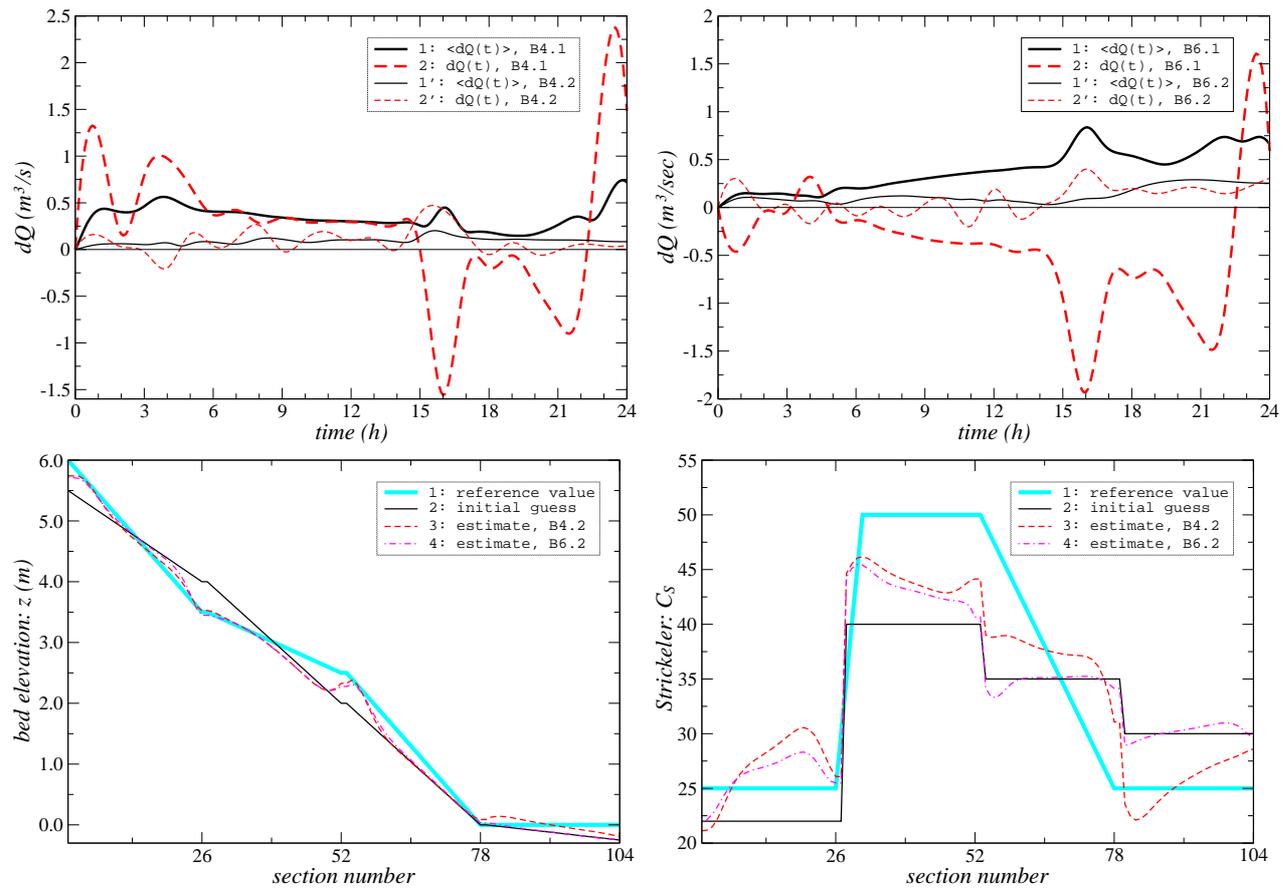


Figure 9: Inflow discharge estimation problem under uncertainty in more than one spatially distributed function.

functions in tests 10 – 15:  $z(x)$  and  $C_s(x)$  in cases  $B4.1$ - $B4.3$  and  $z(x)$ ,  $b(x)$  and  $C_s(x)$  in cases  $B6.1$ - $B6.3$ . The testing principle is the same as before: first, the uncertainty is ignored (cases  $B4.1$  and  $B6.1$ ), then the uncertainty is taken into account by including the corresponding variables into the control vector (cases  $B4.2$  and  $B6.2$ ). The results are presented in Fig.9. The upper subplots show the errors  $\langle dQ(t) \rangle$  and  $dQ(t)$ . As before,  $l.1$  and  $l.2$  show the estimation error when the uncertainty is ignored, and  $l.1'$  and  $l.2'$  - when the uncertainty is taken into account. The lower/left subplot shows the graphs related to  $z(x)$ , and the lower/right subplot - those related to  $C_s(x)$ . The estimate of  $b(x)$  in case  $B6.2$  is presented in  $l.4$  in Fig.8(lower/right).

We notice that taking uncertainty into account allows the errors  $\langle dQ(t) \rangle$  and  $dQ(t)$  to be reduced significantly, though not as significantly as in tests n.3 – 9. At the same time the improvement over the prior in  $z(x)$ ,  $C_s(x)$  and  $b(x)$  is far less evident. Assessing these results one may conclude that the time-dependent control can be well resolved against spatially distributed controls, however it may be difficult to resolve simultaneously different spatially distributed controls, for example,  $z(x)$  and  $C_s(x)$ . This may depend, of course, on the flow dynamics and on the observations available.

Let us also comment on tests n.12 (case  $B4.3$ ) and n.15 (case  $B6.3$ ). The motivation to perform these tests is as follows. We assume that the estimate of  $Q(t)$  is the main objective

of data assimilation, whereas all other functions are included into the control vector just to mitigate the influence of the associated uncertainties on  $Q(t)$ . Therefore, one may try to deal with all parameter uncertainties by introducing a single spatially-dependent model error. For example, along with  $Q(t)$  only the 'effective' value of  $z(x)$  can be sought. Some results on tests n.12 and n.15 are presented in Tab.4. Let us consider case *B4.3*. Here, the achieved value of the residual norm is  $1.080 \cdot 10^2$ . This is not too different from the corresponding value in case *B4.2*, which is  $0.640 \cdot 10^2$ , given that the residual threshold value is  $\Delta = 2.302 \cdot 10^3$ . While the estimation errors  $\langle dQ(t) \rangle$  and  $dQ(t)$  have been only slightly larger than in case *B4.2*, the solution process has required  $N = 151$  iterations against  $N = 61$  in case *B4.2*. Similar results have been obtained in case *B6.3*. Thus, in terms of accuracy the 'effective bathymetry' approach gives comparable results, however more iterations may be required.

## Conclusions

1. A novel version of the variational DA method which is suitable for solving estimation-under-uncertainty problems is suggested. The method includes several components described in Sections 4, 5.1 and 5.4. The idea of regularizing iterative process (rather than the optimal solution) is acquired from the IR concept. An approach utilized to implement this idea (change of variables) leads us to the iterative process (4.30). The key property of this iterative process is its regular convergence, which helps (albeit not guaranties) keeping the iterates within the domain where the model solution exists. Using the suggested iterative process the solution has been successfully obtained for all subcases of test case 'B', for example. This has not been achieved while using the conventional iterative process 3.20.

The method can be used with 2D and 3D models if the operator-vector product  $B^{1/2}v$  (or  $Bv$ ) is available. However, using the quasi-Newton method with the 3D models may not be feasible because of computational limitations.

2. Automatic differentiation engine TAPENADE has been successfully applied to the SIC<sup>2</sup> model to generate its tangent linear and adjoint models. While SIC<sup>2</sup> was not originally designed for this type of processing, only a few minor changes to the original code have been introduced. The adjoint model has been extensively tested in different configurations (beyond those presented in this paper) to check the gradient. It must be acknowledged that TAPENADE has greatly evolved since one of the authors used it for the first time in 2005 [18], and has become a powerful tool for processing the progressively complex models, such as SIC<sup>2</sup> or the full physics ocean model ORCA [24]. It is hardly a secret that difficulty with constructing the adjoint model often prevents researches and practitioners from using the variational DA method (or any other gradient based optimization). These examples should facilitate application of this method in relevant circumstances.

3. Several data assimilation problems have been tried within the identical twin experiment framework. Attention has mainly been paid to inflow discharge and offtake estimation problems under uncertainty in basic model parameters. Having the minimization problem solved to a certain level of accuracy does not necessarily imply that all components of the composite control vector are properly estimated. Due to the problem's nonlinearity, there may exist many local minima in which the minimization process is likely to be trapped. The most difficult case is when the difference between the estimates at these (stationary) points is essential, while the

difference between the corresponding cost function values remains negligible as compared to the observation noise.

The numerical results show that, having no uncertainty in model parameters, one can well estimate the inflow discharge or resolve several offtakes or tributaries. Under uncertainty, one can successfully resolve the composite control which includes time-dependent controls and lumped parameters (case A), or one time-dependent control and one spatially distributed control (case B). However, it is difficult to resolve several spatially distributed controls, if more than one have been included into the control vector (e.g. bed elevation and Strickler coefficient) alongside the time-dependent control. The affirmative result is that even in the latter case the time-dependent control itself (which is often the main object of interest) is estimated sufficiently well.

4. The identifiability issue has to be investigated in a systematic way, rather than empirically by considering the estimation error for a set of test cases. This can be achieved by analyzing the eigenvalues of the Hessian of an auxiliary control problem, see e.g. [19], which is an immediate future task. Let us note that such analysis is possible thanks to availability of the tangent linear and adjoint counterparts of SIC<sup>2</sup>. Another nearest methodological task is to account for 'stochastic' uncertainty components using the 'nuisance' parameter concept.

5. It has been mentioned that development of methodology and software for discharge estimation under uncertainty using SIC<sup>2</sup> is partly motivated by the future SWOT satellite mission. This mission is going to provide new types of data, namely, water surface width and slope, which requires an appropriate observation operator to be defined. Another feature of the SWOT-type data is its sparse-in-time, but dense-in-space nature. It is interesting to investigate how this relatively crude data interacts with dense-in-time, but sparse-in-space accurate *in-situ* data. We are also working on application of the developed tools to Garonne river benchmark [38].

## Acknowledgments

The authors thank Jean-Pierre Baume and David Dorchies, the members of the G-EAU team, IRSTEA-Montpellier, for their help with the SIC<sup>2</sup> software acquaintance, and Prof. Victor Shutyaev, Institute of Numerical Mathematics, Russian Academy of Science, Moscow, for useful discussions on the methodological issue.

## A

Appendix 1: Short description of the SIC<sup>2</sup> model SIC<sup>2</sup> is a hydrodynamic model that has been developed at IRSTEA (CEMAGREF) for more than 30 years. It is an industrial software distributed to different type of users, including consultant companies, irrigation canal managers, engineering schools and universities all over the world (France, Spain, Italy, Portugal, Netherlands, England, Germany, Morocco, Tunisia, Egypt, Senegal, USA, Mexico, Pakistan, Iraq, Sri Lanka, Vietnam, China, etc). It has many innovative features, that make it the leader among this type of software, for some specific applications including irrigation canal design, irrigation canal manual or automatic control, and data assimilation.

The basic features of the SIC<sup>2</sup> model are as follows:

a) the model is based on the full Saint-Venant 1D non-linear partial derivative equations [4];

- b) the model is based on the semi-implicit Preissmann scheme [37];
- c) two-step solution approach is used: the boundary conditions for the reaches are computed first, then the water profiles in the reaches are recovered. The second step can be potentially implemented in a parallel setting;
- d) in the version of SIC<sup>2</sup> used, only subcritical flows are allowed in the unsteady mode, but local critical and supercritical flows can be managed within the cross-devices <sup>6</sup>;
- e) the canal can be composed of a minor, medium (with a different Strickler coefficient) and major bed (can be used as a storage area during the canal overflow events) and ponds at nodes. The minor - medium bed interactions are modelled using the Debord formula, validated on large laboratory experiments, giving better results than the more classical Divided Channel Method [17];
- f) the model allows the pressurized flow conditions using the Preissmann slot approach;
- g) the model has two separate modules: one calculating real steady flow solutions, even in branched and looped networks, without a priori knowledge of flow repartition, and one calculating unsteady flow solutions on the same type of networks. The steady flow module is able to manage any well posed boundary conditions, such as water levels, discharges and rating curves, at either upstream, downstream or intermediate boundary conditions.

One original and unique feature of SIC<sup>2</sup> is to be able to describe any operational rule or algorithm either of feedforward or feedback type, moving any dynamical cross or lateral device (gate, weir, pump, etc) using any measurement over the hydraulic system. This allows to design, test and optimize management rules on irrigation canal, or on rivers having dynamical devices (dams, hydroelectrical power plants, moving weirs, etc). Some predefined algorithms are already available into a library (ex: PID), even some of them with auto-tuning procedures (ex: ATV). More advanced algorithms can be implemented using several programming languages (ex: MatLab, Scilab, Fortran, WDLangage) taking advantage of an embedded interface of these languages into SIC<sup>2</sup>. Using this feature some very advanced MIMO (Multi Input, Multi Output) automatic controllers have been tested such as  $LQG$ ,  $\ell_1$ ,  $H_\infty$  [29, 30].

Another original and unique feature of SIC<sup>2</sup> is its capability to model complex hydraulic structures that are encountered on irrigation canals, such as hydrodynamic gates (AMIL, AVIS, AVIO, Mixte gates). Also, the modelling of more classical devices such as gates and weirs are modelled in such a way that it allows all possible flow conditions and all continuous transitions between these conditions.

SIC<sup>2</sup> has been used to design and test different Data Assimilation algorithms such as Kalman Filter, Monte Carlo particle filters, on both test cases or real systems (Rhône River) [22].

For a detailed description of the model see [4, 31] and the User's Manual at the website: <http://sic.g-eau.net>.

## A

Appendix 2: List of notations a) Latin

$A(Z, p_g)$  - wetted area

ATV - Auto-Tuned Variation method to tune a *PID* controller

---

<sup>6</sup>In the recent versions of SIC<sup>2</sup> the supercritical flow regime is supported in both steady and unsteady flow calculation, implementing ideas developed in [40]

$\mathcal{B}$  - background covariance matrix of  $U$   
 $B$  - background covariance matrix of  $V$  (sub-matrix of  $\mathcal{B}$ )  
 $b(x)$  - dilation coefficient, scales the cross-section profile (part of  $p_g$  vector)  
 $C$  - observation operator  
 $C_d$  - cross-device discharge coefficient  
 $C_L(x)$  - lateral discharge coefficient  
 $C_s(x)$  - Strickler coefficient  
 $dQ(t)$  - inflow discharge estimation error  
 $dt$  - model integration time step  $dt_o$  - observation time step  $dt_c$  - time-dependent control time step  $\langle dQ(t) \rangle$  - spatially averaged discharge estimation error  
 $\mathcal{F}(Z, C_d)$  - cross-device function  
 $g$  - gravity acceleration  
 $h$  - 'head'  
 $H, \tilde{H}$  - Hessian matrices (in different spaces)  
 $H_\infty$  - Automatic controller algorithm based on the  $H_\infty$  norm minimization  
 $J(U)$  - data assimilation cost function  
 $J'_V(U)$  - cost function gradient with respect to  $V$  at point  $U$   
 $\ell_1$  - Automatic controller algorithm based on the  $\ell_1$  norm minimization  
 $LQG$  - Linear Quadratic Gaussian automatic controller algorithm  
 $m$  - dimension of the observation vector  
 $O$  - observation error covariance matrix  
 $p_g$  - general notation for parameters defining the bathymetry  
 $p_{rc}$  - rating curve data  
 $P(Z, p_g)$  - wetted perimeter  
 $PID$  - Proportional Integral Derivative automatic controller algorithm  
 $q_i$  - offtake/tributary discharge at node number  $i$  (lumped source term)  
 $q_{b,i}$  - prior of  $q_i$   
 $Q(x, t)$  - local discharge (state variable)  
 $Q_{ref}$  - 'true' value of  $Q(x, t)$   
 $Q_0(x)$  - discharge initial value (initial condition)  
 $Q(t)$  - discharge at boundary node (boundary conditions)  
 $Q_{ref}$  - 'true' value of  $Q(t)$   
 $Q_L(x, t)$  - lateral discharge (distributed source term)  
 $\mathcal{R}$  - control-to-observation mapping  
 $r(W)$  - residual vector  
 $t$  - time  
 $T$  - final time  
 $\mathcal{U}$  - control space  
 $U$  - full control vector  
 $U^0$  - complement of  $V$  in  $U$   
 $U_b$  - prior (background) of  $U$   
 $v$  - mean velocity  
 $V$  - subset of  $U$   
 $V^{ref}$  - 'true' value of  $V$   
 $\hat{V}$  - estimate of  $V$

$V_b$  - prior (background) of  $V$   
 $W$  - modified control variable  
 $\mathcal{X}$  - state space  
 $x$  - the computational section longitudinal abscissa  
 $\mathcal{Y}$  - observation space  
 $Y$  - observed model predictions  
 $Y^*$  - observations  
 $\mathcal{Z}(t)$  - water level at boundary node (boundary conditions)  
 $Z(x, t)$  - local water level (state variable)  
 $Z_0(x)$  - elevation and discharge initial value (initial condition)  
 $z(x)$  - bed elevation, the lowest point of the cross-section profile (part of  $p_g$  vector)

b) Greek

$\alpha$  - regularization parameter  
 $\beta$  - descent step  
 $\gamma$  - confidence level  
 $\Delta$  - observation uncertainty norm (residual threshold)  
 $\sigma_B$  - background error mean deviation  
 $\sigma_O$  - observation error mean deviation  
 $\chi^2(m, \beta)$  -  $\chi^2$ -probability distribution  
 $\xi_b$  - background error  
 $\xi_o$  - observation error

## References

- [1] Alifanov O.M., Artyukhin E.A., Rumyantsev S.V., *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*, Begel House Publishers, (1996).
- [2] Atanov G.A., Evseeva E.G., Meselhe E.A., Estimation of the roughness profile in trapezoidal open channels. *Journal of Hydraulic Engineering*, 1999, Vol. 125, Issue 3, pp. 309–312.
- [3] Barthelemy S., Ricci S., Pannekoucke O., Thual O., Malaterre P.-O., Emulation of a Kalman filter algorithm on a flood wave propagation model. *Journal of Hydrology*, 2014, in review.
- [4] Baume J.-P., Malaterre P.-O., Belaud G., Le Guennec B., SIC: a 1D Hydrodynamic Model for River and Irrigation Canal Modeling and Regulation, *Métodos Numéricos em Recursos Hídricos 7, ABRH (Associação Brasileira de Recursos Hídricos)*, Coppetec Fundacao, Editor Rui Carlos Vieira da Silva, 2005, pp. 1-81.
- [5] Bocquet M., Sakov P., An iterative ensemble Kalman smoother. *Quart. J. Roy. Meteor. Soc.*, 2013, Vol. 140, Issue 682, pp. 1521–1535.

- [6] Cacuci D.G., Navon I.M., Ionescu-Bugor M. *Computational Methods for Data Evaluation and Assimilation*. Taylor and Francis CRC Press, 2013.
- [7] Courtier P., Thepaut J.N., Hollingsworth A., A strategy for operational implementation of 4D-Var, using an incremental approach. *Q.J.R. Meteorol. Soc.*, 1994, Vol. 120, pp. 1367–1387.
- [8] Davies E.G.R., Simonovic S.P. Global water resources modeling with an integrated model of the social-economic-environmental system. *Advances in Water Resources*, 2011, Vol. 34, Issue 6, pp. 684–700.
- [9] Derber J., Bouttier F.A. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, 1999, Vol. 51, pp. 195–221.
- [10] Derber J., Wu W.-S. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Weather Rev.*, 1998, Vol. 126, pp. 2287–2299.
- [11] Ding Y., Wang S.Y., Optimal control of flood diversion in watershed using nonlinear optimization. *Advances in Water Resources*, 2012, Vol. 44, pp. 30–48.
- [12] El-Hanafy H., Copeland G.J.M., Gejadze I.Yu, Estimation of predictive uncertainties in flood wave propagation in a river channel using adjoint sensitivity analysis. *Int. J. Num. Meth. Fluids*, 2008, Vol. 56, Issue 8, pp. 1201–1207.
- [13] Garcia-Pintado J., Mason D.C., Cloke H.L., Neal J.C., Freer J., Bates P.D., Satellite-supported flood forecasting in river networks: A real case study. *Journal of Hydrology*, 2015, Vol. 523, pp. 706–724.
- [14] Hartnack J.N., Madsen H., Sorensen J.V.T., Data assimilation in a combined 1D-2D flood model using the ensemble Kalman filter. *Proceedings CMWR XVI*, June 19-22, 2006, Copenhagen.
- [15] Evensen G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.*, 1994, Vol. 99 (C5), pp. 10143–10162.
- [16] Evensen G., van Leuwen P.J., An Ensemble Kalman Smoother for Nonlinear Dynamics, *Mon. Wea. Rev.*, 2000, Vol. 128, pp. 1852-1867.
- [17] Fernandes J.N., Leal J.B., Cardoso A.H., Flow structure in a compound channel with smooth and rough floodplains, *EWRA European Water Publications*, 2012, Vol. 38, pp. 3-12.
- [18] Gejadze I., Monnier J., On a 2D zoom for 1D shallow-water model: coupling and data assimilation. *Comput. Meth. Appl. Mech. Eng.*, 2007, Vol. 196 (45-48), pp. 4628–4643.
- [19] Gejadze I., Le Dimet F.-X., Shutyaev V., On optimal solution error covariances in variational data assimilation problems. *J. Comput. Phys.*, 2010, Vol. 229, pp. 2159–2178.

- [20] Hansen P.C., O’Leary D.P., The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.*, 1993, Vol. 14, No. 6, pp. 1487–1503.
- [21] Hascoët L., Pascual V., TAPENADE 2.1 user’s guide. *INRIA Technical Report*, 2004, no.0300, 78 pp.
- [22] Jean-Baptiste N., Malaterre P.-O., Dorée C., Sau, J. Data assimilation for real-time estimation of hydraulic states and unmeasured perturbations in a 1D hydrodynamic model, *Math. Comput. Simul.*, 2011, Vol. 81, pp. 2201-2214
- [23] Kaltenbacher B., Neubauer A., Scherzer O., *Iterative Regularization Methods for Nonlinear Ill-posed Problems*. Radon Series on Computational and Applied Mathematics. de Gruyter, 2008.
- [24] Kazantsev E., Optimal boundary conditions for ORCA-2 model. *Ocean Dynamics*, 2013, Vol .63, Issue 8, pp. 943–959.
- [25] Le Dimet F.X., Talagrand O., Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 1986, Vol. 38, pp. 97–110.
- [26] Lions J.L. *Contrôle Optimal des Systèmes Gouvernés par des Équations aux Dérivées Partielles*. – Paris: Dunod, 1968.
- [27] Liu D., Nocedal J., On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 1989, Vol. 45 B, pp. 503–528.
- [28] Lorenc A.C., Iterative analysis using covariance functions and filters. *Q.J.R. Meteorol. Soc.*, 192, Vol. 118, pp.569–591.
- [29] Malaterre P.-O., Khammash M.,  $\ell_1$  controller design for a high-order 5-pool irrigation canal system. *ASME Journal of Dynamic Systems, Measurement, and Control*, 2003, Vol. 125, pp. 639–645
- [30] Malaterre P.-O., PILOTE: linear quadratic optimal controller for irrigation canals. *ASCE Journal of Irrigation and Drainage Engineering*, 1998, Vol. 124, pp. 187–194
- [31] Malaterre P.-O., Dorchies D., Baume J.-P., Simulation and Integration of Control for Canals software (*SIC*<sup>2</sup>), for the design and verification of manual or automatic controllers for irrigation canals. *USCID Conference on Planning, Operation and Automation of Irrigation Delivery Systems*, Phoenix, Arizona, December 2-5, 2014.
- [32] Marchuk G.I., Agoshkov V.I., Shutyaev V.P., *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*. – New York: CRC Press Inc., 1996.
- [33] Mirouze I., Weaver A.T., Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Q.J.R. Meteorol. Soc.*, 2010, Vol. 136, Part B, pp.1421–1443.
- [34] Morozov V.A., *Methods of solving incorrectly posed problems*, Springer-Verlag, New-York, 1984.

- [35] Munier S., Polebistki A., Brown C., Belaud G., Lettenmaier P.P. SWOT data assimilation for operational reservoir management on the upper Niger River Basin. *Water Resources Research*, 2015, Vol. 51, Issue 1, pp. 554–575.
- [36] Navon I.M., Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dynamics of Atmospheres and Oceans*, 1998, Vol.27, Issue 1-4, pp. 55–79.
- [37] Novak P., Guinot V., Jeffrey A., Reeve D.E., *Hydraulic Modelling - An Introduction: Principles, Methods and Applications*. CRC Press, 2010.
- [38] Oubanas H., Gejadze I., Malaterre P.-O., Mercier F., Estimation of river discharge from in-situ and remote sensing data, using variational data assimilation and a full Saint-Venant hydraulic model. 3rd Space for Hydrology Workshop, Frascati (Rome), Italy, September 15–17, 2015
- [39] Sanders B.F., Katopodes N.D., Control of canal flow by adjoint sensitivity method. *Journal of Irrigation and Drainage Engineering*, 1999, Vol. 125, Issue 5, pp. 287–297.
- [40] Sart C., Baume J.-P., Malaterre P.-O., Guinot V., Adaptation of Preissmann’s scheme for transcritical open channel flows. *Journal of Hydraulic Research*, 2010, Vol. 48, Issue 4, pp. 428–440.
- [41] Tamura H., Bacopoulos P., Wang D., Hagen S.C., Kubatko E.J. State estimation of tidal hydrodynamics using ensemble Kalman filter. *Advances in Water Resources*, 2014, Vol. 63, pp. 45–56.
- [42] Thacker W.C., The role of the Hessian matrix in fitting models to measurements, *J. Geophys. Res.*, 1989, Vol. 94, pp. 6177–6196.
- [43] Tikhonov A.N., Arsenin V.Y., *Solutions of Ill-posed Problems*, Winston-Wiley, New-York, 1977.
- [44] Weerts A.H., El Serafy G.Y., Hummel S., Dhondia J., Gerritsen H., Application of generic data assimilation tools (DATools) for flood forecasting purposes. *Computers and Geosciences*, 2010, Vol. 36, pp. 453–463.
- [45] Wu Q, Tinka A., Weekly K., Beard J., Bayen A.M. Variational Lagrangian data assimilation in open channel networks. *Water Resources Research*, 2014, Vol. 51, pp. 1916–1938.

Case	A1	A2	A3	A4
noise	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
regul.	<i>none</i>	<i>none</i>	<i>TR</i>	<i>IR</i>
$N$	114	75	32	22
$\ r(V_0)\ ^2$	$1.094 \cdot 10^4$	$1.567 \cdot 10^4$	$1.567 \cdot 10^4$	$1.567 \cdot 10^4$
$\ r(V_N)\ ^2$	$3.177 \cdot 10^0$	$4.487 \cdot 10^3$	$4.613 \cdot 10^3$	$4.612 \cdot 10^3$
$\ J'_V(V_0)\ ^2$	$1.142 \cdot 10^{10}$	$1.162 \cdot 10^{10}$	$1.162 \cdot 10^{10}$	$6.650 \cdot 10^8$
$\ J'_V(V_N)\ ^2$	$3.243 \cdot 10^1$	$1.671 \cdot 10^3$	$2.954 \cdot 10^5$	$9.561 \cdot 10^3$
$\epsilon$	0.424	27.48	0.165	0.123

Table 2: Case A: auxiliary information

Case	A5	A6	A7
noise	<i>no</i>	<i>no</i>	<i>yes</i>
regul.	<i>none</i>	<i>none</i>	<i>IR</i>
$N$	68	122	63
$\ r(V_0)\ ^2$	$9.357 \cdot 10^4$	$9.357 \cdot 10^4$	$9.813 \cdot 10^4$
$\ r(V_N)\ ^2$	$4.303 \cdot 10^3$	$3.131 \cdot 10^0$	$4.609 \cdot 10^3$
$\ J'_V(V_0)\ ^2$	$1.189 \cdot 10^{11}$	$2.661 \cdot 10^{12}$	$2.582 \cdot 10^{11}$
$\ J'_V(V_N)\ ^2$	$3.679 \cdot 10^1$	$8.044 \cdot 10^1$	$5.719 \cdot 10^4$
$\epsilon$	95.9	0.569	0.126
$\delta C_{d,1}$	0.0	$-2.051 \cdot 10^{-3}$	$-8.308 \cdot 10^{-4}$
$\delta C_{d,2}$	$-6.000 \cdot 10^{-2}$	$-2.378 \cdot 10^{-4}$	$-2.122 \cdot 10^{-4}$
$\delta C_{d,3}$	$-1.200 \cdot 10^{-1}$	$-4.016 \cdot 10^{-4}$	$4.834 \cdot 10^{-5}$
$\delta C_{d,4}$	0.0	$1.137 \cdot 10^{-4}$	$-3.221 \cdot 10^{-3}$

Table 3: Case A (extension): auxiliary information

Case	<i>B1.1</i>	<i>B1.2</i>	<i>B2.1</i>	<i>B2.2</i>	<i>B2.3</i>
noise	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>	<i>yes</i>
regul.	<i>none</i>	<i>IR</i>	<i>IR*</i>	<i>IR*</i>	<i>IR</i>
<i>N</i>	51	33	63	68	39
$\ r(V_0)\ ^2$	$2.279 \cdot 10^5$	$2.504 \cdot 10^5$	$4.308 \cdot 10^5$	$4.093 \cdot 10^5$	$4.308 \cdot 10^5$
$\ r(V_N)\ ^2$	$1.479 \cdot 10^{-3}$	$2.323 \cdot 10^4$	$1.179 \cdot 10^5$	$1.439 \cdot 10^1$	$2.323 \cdot 10^4$
$\ J'_V(V_0)\ ^2$	$9.713 \cdot 10^8$	$3.061 \cdot 10^9$	$5.774 \cdot 10^9$	$5.865 \cdot 10^9$	$5.832 \cdot 10^9$
$\ J'_V(V_N)\ ^2$	$6.355 \cdot 10^{-1}$	$8.957 \cdot 10^2$	$1.959 \cdot 10^1$	$1.853 \cdot 10^3$	$1.328 \cdot 10^3$
$\epsilon$	0.738	0.706	1.475	0.698	0.710
Case	<i>B3.1</i>	<i>B3.2</i>	<i>B4.1</i>	<i>B4.2</i>	<i>B4.3</i>
noise	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
regul.	<i>IR*</i>	<i>IR*</i>	<i>IR*</i>	<i>IR*</i>	<i>IR*</i>
<i>N</i>	61	68	75	61	151
$\ r(V_0)\ ^2$	$2.589 \cdot 10^5$	$2.589 \cdot 10^5$	$4.075 \cdot 10^5$	$4.075 \cdot 10^5$	$4.075 \cdot 10^5$
$\ r(V_N)\ ^2$	$2.706 \cdot 10^4$	$9.225 \cdot 10^0$	$1.292 \cdot 10^5$	$6.400 \cdot 10^1$	$1.080 \cdot 10^2$
$\ J'_V(V_0)\ ^2$	$1.878 \cdot 10^9$	$1.884 \cdot 10^9$	$4.090 \cdot 10^9$	$4.193 \cdot 10^9$	$4.152 \cdot 10^9$
$\ J'_V(V_N)\ ^2$	$1.413 \cdot 10^0$	$1.715 \cdot 10^2$	$5.693 \cdot 10^0$	$1.351 \cdot 10^3$	$1.356 \cdot 10^2$
$\epsilon$	0.871	0.795	1.295	0.721	0.657
Case	<i>B5.1</i>	<i>B5.2</i>	<i>B6.1</i>	<i>B6.2</i>	<i>B6.3</i>
noise	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
regul.	<i>IR*</i>	<i>IR*</i>	<i>IR*</i>	<i>IR*</i>	<i>IR*</i>
<i>N</i>	6	140	80	65	142
$\ r(V_0)\ ^2$	$2.299 \cdot 10^5$	$2.299 \cdot 10^5$	$4.967 \cdot 10^5$	$4.967 \cdot 10^5$	$4.967 \cdot 10^5$
$\ r(V_N)\ ^2$	$2.084 \cdot 10^4$	$2.193 \cdot 10^{-1}$	$2.338 \cdot 10^5$	$4.906 \cdot 10^1$	$9.386 \cdot 10^1$
$\ J'_V(V_0)\ ^2$	$1.418 \cdot 10^9$	$1.432 \cdot 10^9$	$2.128 \cdot 10^9$	$2.449 \cdot 10^9$	$2.235 \cdot 10^9$
$\ J'_V(V_N)\ ^2$	$4.066 \cdot 10^{-1}$	$1.109 \cdot 10^1$	$4.066 \cdot 10^{-1}$	$1.373 \cdot 10^3$	$1.430 \cdot 10^2$
$\epsilon$	1.182	0.742	1.183	0.708	0.582

Table 4: Case B: auxiliary information