



IDENTIFIABILITY AND ASSIMILATION OF SPARSE ALTIMETRIC DATA IN 1D SAINT-VENANT RIVER MODELS

Pierre Brisset, Pierre-André Garambois, Jerome Monnier, Hélène Roux

► To cite this version:

Pierre Brisset, Pierre-André Garambois, Jerome Monnier, Hélène Roux. IDENTIFIABILITY AND ASSIMILATION OF SPARSE ALTIMETRIC DATA IN 1D SAINT-VENANT RIVER MODELS. 2016. hal-02044488v2

HAL Id: hal-02044488

<https://hal.science/hal-02044488v2>

Preprint submitted on 27 Sep 2016 (v2), last revised 18 Feb 2019 (v3)

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.

IDENTIFIABILITY AND ASSIMILATION OF SPARSE ALTIMETRIC DATA IN 1D SAINT-VENANT RIVER MODELS

PIERRE BRISET (1), PIERRE-ANDRÉ GARAMBOIS (2)(3), JÉRÔME MONNIER * (1)(4), HÉLÈNE ROUX (5)(6)

- (1) Institut de Mathématiques de Toulouse (IMT), Toulouse, France.
- (2) INSA Strasbourg, France.
- (3) ICUBE, Strasbourg, France.
- (4) INSA Toulouse, France.
- (5) INPT Toulouse, France.
- (6) Institut de Mécanique des Fluides (IMFT), Toulouse, France.

* Corresponding author: jerome.monnier@insa-toulouse.fr

Abstract. Various altimetry satellite missions provide water level elevation measurements of wide rivers; in particular the future Surface Water and Ocean Topography (SWOT) mission which will provide measurements at unprecedented accuracies for 90% of the globe. The capabilities of these data to infer the inflow discharge and roughness parameters in the 1D Saint-Venant equations are investigated. As a first step, identifiability maps representing the observations and the model waves in the (x, t) plan are analysed. These maps provide a comprehensive overview of the inverse problem difficulties. Next the inflow discharge, and potentially the roughness coefficients, are inferred by variational data assimilation. The inferences are analysed for various observation sparsity degrees. Sensitivity analyses with respect to the observation errors and the first guess values demonstrate the robustness of the VDA inferences. The identifiability maps may make it possible to recover the space-time windows where the inference obtained by VDA are potentially unreliable. Finally this study demonstrates the great potential of these forthcoming measurements but also their limitations in 1D river flow dynamic models if not combined to other data sources.

Keywords. Identifiability, shallow water, variational data assimilation, altimetry, SWOT, river hydraulics, discharge.

1. INTRODUCTION

While the in situ observability of continental water cycle, especially river flows, is declining, satellites provide increasingly accurate measurements. The future Surface Water and Ocean Topography (SWOT) mission, CNES-NASA, equipped with a swath mapping radar interferometer, will provide river surface mapping at a global scale with an unprecedented spatial and temporal resolution [32]. Maps of water elevations are expected at a resolution of approximately 100 m with a vertical accuracy on the order of centimeters if averaged over 1 km² [32]. An other highlight of SWOT will be its global coverage and temporal revisits (1 to 4 revisits per 22 – days repeat cycle). This should offer the opportunity to increase our knowledge of the spatial and temporal distribution of hydrological fluxes including stream and rivers see e.g. [5]. Thanks to this increased observability of water surfaces worldwide, it will be possible to address a variety of inverse problems in surface hydrology. From the river hydraulics point of view, the basic inverse questions are the following: given these surface measurements (elevation, water mask width and potentially the slopes), can we infer the discharge ? the roughness coefficients ? the bathymetry ? the lateral contributions ? given what space-time observations density and at what scale ?

A quite recent but already large literature addresses some of these inverse questions. Since rivers may be correctly modeled by quite simple 1D equations, some inverse problems can be addressed analytically, see e.g. [12, 13]. On similar mathematical approaches but in the SWOT context (hence sparse measurements both in

space and time), [11] demonstrates that the bathymetry and/or the roughness may be successfully inferred, depending on the regime variations of the observed flow, see also [40, 10]. In order to better constrain these underdetermined inverse problems, prior hydraulic information or empirical laws may be required.

Sophisticated data assimilation techniques, both sequential and variational, are more and more commonly employed in the river hydraulic modeling community. Numerous studies employ the sequential approach (Kalman’s filter and variants). Let us cite for example [31] which consider a diffusive wave model with the river bathymetry and roughness coefficients given, and it corrects the upstream discharge via the assimilation of downstream water depth measurements. [39] assimilates synthetic observations corresponding to 8 cycles of SWOT in a 1D simplified shallow water model which improves the discharge estimation (with the roughness coefficient given). [2] demonstrate that the assimilation of SWOT like data makes it possible to improve discharge forecasts for different lead times.

The Variational Data Assimilation (VDA) approach, based on optimal control, consists in minimizing a cost function measuring the discrepancy between the model outputs and the observations. This approach aims at optimally combining the model, the observations and prior statistical information. In some circumstances, it is possible to infer unknown “parameters” such as the boundary conditions (e.g. inflow discharge), model parameters (e.g. roughness) and/or forcing terms. (In meteorology, the historical application of VDA, the main unknown “parameter” is the initial state). Inferring the river discharge and/or hydraulic parameters from the surface measurements is not straightforward and may be even impossible, depending on the flow dynamics, the observations sparsity etc. Several studies tackle the identification by VDA of the river bathymetry and/or discharge in various observational contexts. In [33, 34], the direct model is the 1D Saint-Venant model; the authors manage to infer flood hydrographs from dense water surface width measurements if the bathymetry and the roughness parameters are given. Based on the 2D shallow water river equations, VDA experiments are presented in [3, 20, 19]. For example, upstream, downstream and lateral conditions are identified on the Pearl River (China) from water levels measured at in situ gauging stations [20]; again the river bathymetry and roughness are given. In [22, 21], different densities of spatially distributed water level observations (corresponding to partial snapshot images potentially acquired by SAR and partial time series acquired at a gauging station) are investigated, potentially making possible the estimation of the inflow discharge of the observed flooding event. In other respects the VDA approach provide useful sensitivity maps to better understand the flow and the model, see e.g. [27].

Despite the huge improvement of the forthcoming altimetric data, data remain extremely sparse for river applications, both in space and time, hence making inferences very challenging in river hydraulic models from these data only. Typically as demonstrated in [11], for a steady-state low-Froude model with reach averaged variables (called 0.5D model), the inference of the (bathymetry, roughness) pair may be impossible depending on the observed flow variations. In the SWOT mission context (planned to be launched in 2021), the next step is to address similar inverse questions but for the complete 1D Saint-Venant equations (hence valid for unsteady and non low Froude flows). These equations are adequate to model dynamic river flows with quite fast discharge variations (otherwise the 0.5D model or the Manning-Strickler balance law per reach is generally sufficient). This is the goal of the present study: to investigate the identifiability of input parameters in the Saint-Venant equations (1D shallow water) from the SWOT-like altimetric data, and perform the unknown parameters by VDA. The unknown “parameters” are the inflow discharge $Q_{in}(t)$ and potentially a non uniform roughness coefficient $K(h)$, h denoting the water depth. The bathymetry z_b is given (it is an effective bathymetry defined as a superimposition of trapeziums to be compatible with the assimilation of altimetric data).

In this context, before performing any “blind” VDA process, so-called *identifiability maps* are presented and analysed. The latter are simply the representation in the (x, t) plane of the observations and the model waves (without the topography and friction source term) arising from (or propagating to) the observed space time “windows”. This preliminary reading of the inverse problem presents a few advantages. It may make it possible to anticipate the potential reliability of the VDA process, it helps to define an a-priori adequate identification frequency and it may make possible to determine a-posteriori the reliability of the identification (i.e. identified values not arising from “blind windows”). Also these preliminary analyses led

us to define the unknown inflow discharge as a Fourier series (hence a global function in time) easily readable in terms of frequencies. Various scenarios depending on the observation density / sparsity, on the inflow discharge definitions and on the unknown parameters (discharge only or the (discharge, roughness) pair) are investigated. It turns out that if considering SWOT like data only, then the discharge inference cannot be accurate due to the data sampling sparsity; to be accurate, the altimetric data has to be combined with other data (e.g. in situ or other sensors measurements). In other respect, several sensitivity analyses (with respect to the observation sparsity or the observation errors) strengthen the numerical results and the analyses.

The computational code developed for the present inverse analyses is original since it considers a complete 1D shallow water model dedicated to the altimetric data (specific bathymetry geometry), with all realistic boundary conditions, a roughness coefficient $K(h)$ depending on the water depth h (hence modeling regime flow variations), plus a complete VDA process. The adjoint equations are obtained by automatic differentiation [17] within the computational software DassFlow [6, 26]. The test cases are quite complete since the academic ones include all the targeted difficulties; also a real case is considered (a portion of Garonne river around Toulouse, France, [35, 23]).

The outline of the article is as follows. In Section 2, the Saint-Venant equations with the bathymetry specifically adapted to the altimetric observations are presented; next the inverse based on VDA approach is presented. The cost function is defined from realistic SWOT data (generated following a SWOT simulator). Section 3 presents the academic test cases and the Garonne river case. In Section 4 the inference of the inflow discharge only is addressed. First the identifiability maps provide a comprehensive overview of the inverse problem difficulty; next the discharge is inferred by VDA for various observation sampling sparsity. In Section 5, similar experiments are conducted but aiming at inferring the (discharge, roughness) pair. Sensitivities of the inferred quantities are analysed with respect to the first guesses and the observation errors. In Section 6, the real Garonne test case is investigated for 3 scenarios: the real SWOT temporal sampling, a slightly densified data sampling (by a factor 1.5 or 2) and finally a data sampling densified by a factor 100. A conclusion and some perspectives are proposed in Section 7. The two appendices present details of the numerical scheme in this context of altimetry-compatible river bathymetries.

2. DIRECT AND INVERSE MODELS

In this part, the forward model (1D Saint-Venant equations) and the inverse model, Variational Data Assimilation (VDA), are described. In particular the model geometry (effective river bathymetry), the observation operator and the minimized cost function are detailed.

2.1. Forward model. Open channel flows are commonly described with the 1D Saint Venant equations, which stems from the Navier Stokes equations under several simplifying hypothesis [8, 7] including averaging of flow variables in the vertical directions. The 1D Saint Venant equations in (S, Q) variables write:

$$(2.1) \quad \begin{cases} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} &= 0 & (2.1.1) \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} + P \right) &= g \int_0^h (h - z) \frac{\partial \tilde{w}}{\partial x} dz - gS \left[\frac{\partial z_b}{\partial x} + S_f \right] & (2.1.2) \end{cases}$$

where S is the wet-cross section (m^2), Q is the discharge ($m^3.s^{-1}$), P is a pressure term, \tilde{w} is the water surface top width (m), g is the gravity magnitude ($m.s^{-2}$), H is the water surface elevation (m), $H = (z_b + h)$ where z_b is the bathymetry (m) and h is the water depth (m). S_f denotes the basal friction coefficient (dimensionless), $S_f = \frac{|Q|Q}{K^2 S^2 R_h^{4/3}}$ with K the Manning-Strickler roughness coefficient ($m^{1/3}.s^{-1}$) and R_h the hydraulic radius (m). The hydraulic radius corresponds to the ratio between the wetted cross section S and the wetted perimeter Pe (m). Note that the discharge Q is related to the average cross sectional velocity u ($m.s^{-1}$) by: $Q = uS$.

Let us point out that the left-hand side of the momentum equation is written in its conservative form (hyperbolic part of the model), while the right-hand side, or source term, can be viewed as a “nudging” term, pulling the model to the basic equilibrium: the gravitational force vs the friction forces; it is the classical Manning-Strickler equation. In terms of energy, this source term constitutes the dissipative term.

This classical model is considered with a specific bathymetry geometry since it is designed for altimetry assimilation. This model is implemented into the DassFlow software [1] which is a data assimilation computational platform which include river hydraulic models adapted for the forthcoming SWOT data. The discrete cross sections are asymmetrical trapezium layers; each layer is defined by one triplet (H_i, w_i, Y_i) corresponding respectively to the water elevation, the water surface width associated to H_i and a centering parameter. In a SWOT context, each layer corresponds to a satellite pass.

Remark 1. If the Froude number, $Fr = \frac{u}{c}$, tends to 0, then the 1D St Venant model can be written as a lubrication like model (a depth-averaged scalar equation), so-called diffusive wave model, see e.g. [36, 9, 29].

In the case of a wide channel (the hydraulic radius $R_h \approx h$), the advective term of the equation corresponds to the velocity $\frac{5}{3}u$. In the identifiability maps presented in next sections, this wave velocity $\frac{5}{3}u$ will be compared to the Saint-Venant model wave velocities $(u - c)$ and $(u + c)$ (gravity waves model). Let us point out that whatever the model considered (1D Saint-Venant or diffusive wave equation), those wave velocities do not transport all the information since the presence of the additional dissipation term (e.g. the right hand side in (eq 2.1) and the diffusive term in the diffusive wave equation).

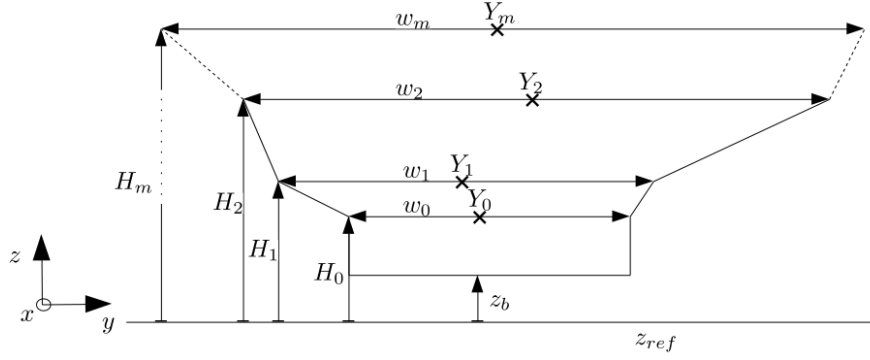


FIGURE 2.1. Effective geometry considered for each cross section: superimposition of m trapeziums (yz -view).

The Manning-Strickler roughness coefficient K_i of a cross section i is defined by a power law in function of the water depth h , hence modeling finely different flow regimes.

Its expression reads:

$$(2.2) \quad K(h) = \alpha h^\beta$$

where α and β are two constants to be determined.

The discharge at upstream boundary $Q_{in}(t)$ will be considered as an unknown parameter of the model (it will be a control parameter of the forward model). Then $Q_{in}(t)$ will be defined by one of these two methods:

IDbasic.: At each identification time t_j , $t_j \in [t_1..t_p]$, a value of $Q_{in}(t_j)$ is computed by the VDA process. Next the identified inflow discharge is continuously constructed by simple linear interpolation.

IDFourier.: The inflow discharge is defined as Fourier series:

$$(2.3) \quad Q_{in}(t) = \frac{a_0}{2} + \sum_{n=1}^{N_{FS}} \left(a_n \cos\left(nt \frac{2\pi}{T}\right) + b_n \sin\left(nt \frac{2\pi}{T}\right) \right)$$

where $\{a_0; a_n, b_n\}$, $n \in [1..N_{FS}]$, are the Fourier coefficients and T is the total simulation time. The lower frequency represented by the Fourier serie is $1/T$ and the highest one is N_{FS}/T . Then this way to identify $Q_{in}(t)$ is global in time (on the contrary to punctual basic approach above). Obviously, the hydrograph must be periodic. In practice this is not a real issue since either the true hydrograph can be truncated or extended it

to make it T-periodic.

Finally the numerical scheme used to solve system of equations (2.1) is the finite volume scheme HLL [16] with Euler integration in time. This numerical scheme with the specificities due to the particular geometrical transformations are presented in Appendix 8.1 and Appendix 8.2

2.2. Inverse model. The inference of the unknown parameters are performed by a VDA approach. It consists to minimize a cost function $J(\mathbf{k})$ measuring the discrepancy between the model output (state variables) and the available measurements (which are sparse and uncertain): $\min_{\mathbf{k}} J(\mathbf{k})$.

Since J depends on k through the model solution (S, Q) , it is an optimal control problem. It is classically solved by introducing the adjoint model and by computing iteratively a “better” control vector \mathbf{k} . The latter contains the inflow discharge $Q_{in}(t)$ and the channel roughness coefficient K .

In the case the unknown parameters are computed at given times $[t_1..t_p]$ (it is the identification time grid), \mathbf{k} is defined by:

$$\mathbf{k} = (Q_{in,1}, \dots, Q_{in,p}, \alpha, \beta)^T$$

In the case the inflow discharge is decomposed as a Fourier series, see 2.3, \mathbf{k} is defined by:

$$\mathbf{k} = (a_0, a_1, b_1, \dots, a_{N_{FS}}, b_{N_{FS}}, \dots, \alpha, \beta)^T$$

The VDA process requires the computation of the gradient of the cost function ∇J with respect to \mathbf{k} . The computation of ∇J is done with DassFlow software which has been originally designed to generate automatically the discrete adjoint model using the source to source differentiation tool Tapenade [18]. The cost function expression J depends on the observations; the latter are presented below while the expression of J is detailed in Section 2.4.

The employed optimization algorithm is a the L-BFGS algorithm, more precisely those implemented in the M1QN3 routine, see [14]; it has been designed for large variable numbers. The VDA process is sketched in Fig. (2.2). Given a first guess on parameters \mathbf{k}_0 , the iterates \mathbf{k}_i are searched with the descent algorithm such as the cost function J decreases. For each iteration of the minimization:

- (1) The cost function $J(\mathbf{k}_i)$ and its gradient $\nabla J(\mathbf{k}_i)$ are computed by performing the forward model (from 0 to T) and its adjoint (from T to 0).
- (2) Given \mathbf{k}_i , $J(\mathbf{k}_i)$ and $\nabla J(\mathbf{k}_i)$, the M1QN3 routine is invoked to compute a new iterate such that: $J(\mathbf{k}_{i+1}) < J(\mathbf{k}_i)$.
- (3) The few convergence criteria are tested: either $|J| \leq 10^{-7}$, or $|J(\mathbf{k}_{i+1}) - J(\mathbf{k}_i)| \leq 10^{-5}$ or $i > 100$.

In order to assess the present hydrological model, the classical Nash-Sutcliffe criteria E is computed [30]. This metric is used to evaluate the accuracy of the identified discharge $Q_{in}^{\text{ident}} = (Q_{in,1}^{\text{ident}}, Q_{in,2}^{\text{ident}}, \dots, Q_{in,p}^{\text{ident}})^T$ as follows:

$$(2.4) \quad E(Q_{in}^{\text{ident}}) = 1 - \frac{\sum_{i=1}^p (Q_{in,i}^{\text{real}} - Q_{in,i}^{\text{ident}})^2}{\sum_{i=1}^p (Q_{in,i}^{\text{real}} - \bar{Q}_{in}^{\text{real}})^2}, \text{ with } \bar{Q}_{in}^{\text{real}} = \sum_{i=1}^p \frac{Q_{in,i}^{\text{real}}}{p}$$

where $Q_{in}^{\text{real}} = (Q_{in,1}^{\text{real}}, Q_{in,2}^{\text{real}}, \dots, Q_{in,p}^{\text{real}})^T$ is the true value. The Nash-Sutcliffe value E is close to 1 for values of Q_{in}^{ident} close to Q_{in}^{real} ; it is close to 0 for values of Q_{in}^{ident} is close to $\bar{Q}_{in}^{\text{real}}$; finally it is close to $-\infty$ for values of Q_{in}^{ident} none correlated to the true value Q_{in}^{real} .

For a given quantity u , $e_2(u)$ denotes the 2-norm relative error:

$$(2.5) \quad e_2(u) = \frac{\|u^{\text{ident}} - u^{\text{real}}\|_2}{\|u^{\text{real}}\|_2}$$

In the inverse process, the variable u above will denote Q_{in} , α or β .

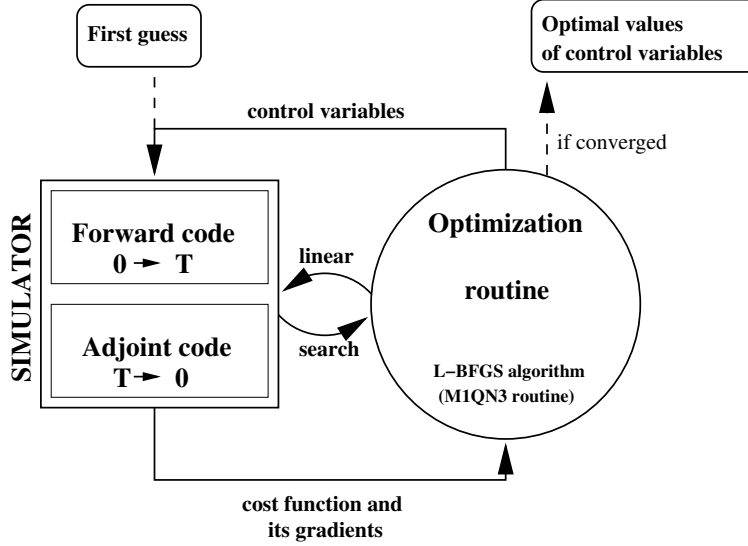


FIGURE 2.2. VDA process, image extracted from [25] .

2.3. The altimetric data (SWOT simulator like). The inverse problem complexity depends greatly on the spatial and temporal density of the water surface measurements. The interferometer of the future SWOT satellite will provide measurements of water surface height, width and slope on two tracks of 50 km width separated by a nadir gap of 20 km. The repeat period is 21 days for a global coverage with 1 to 4 passes above the same location depending on the pattern of ascending and descending tracks, see Fig. 2.3 Left. The accuracy of the water surface elevation measurement should be of: 45 cm for a water mask of 100 m²; 20 cm for 250 m²; 10 cm for 1000 m², see [32]. The expected error on the free surface slope is 1.7 cm/1 km, and better for rivers widths greater than 100 m. The relative error on the total water body area should be smaller than 15 %.

In the present study, synthetic SWOT observations are generated from the 1D forward model described previously; its outputs are averaged in space in order to reproduce SWOT like observations at the reach scale; finally a random noise can then be added. Reaches are defined from the intersection of river centerline and SWOT swath coordinates divided in stripes, see Fig. 2.3. For each group of forward model cells forming a reach r and for each observation time t , the spatial averages $\bar{H}_r^{\mathbf{k}}(t)$ (resp. $\bar{w}_r^{\mathbf{k}}(t)$) of the mesh cell water elevation $H^{\mathbf{k}}(x, t)$ (resp. water surface width $w^{\mathbf{k}}(x, t)$) are computed; \mathbf{k} is the parameter vector corresponding these model outputs.

These observations, which can be perturbed or not depending on the test case, are then used for parameters identification; it is twin experiments.

Let us remark that if the bathymetry is known (or equivalently after many satellite passes and given the set of triplets (H_i, w_i, Y_i)), the water surface width w does not provide any additional information since given the geometry and the water elevation H , the width w can be straightforwardly deduced.

From Section 3.2, SWOT like observations are generated according to the expected ground track, temporal overpasses and SWOT satellite period (21 days). In the case of Garonne River, which is one of the two present test cases, SWOT satellite will pass 3 times over. Then each swath (50 km wide) defined by the SWOT ascending and descending tracks are splitted into 1 km stripes, where the river portions define the reaches for SWOT observables generation. Only 25 stripes contain the considered Garonne river portions, hence as many observed river reaches; this is the swath portions represented in Fig. 2.3. These 25 observed reaches can be classified in 3 groups; each of them is observed at different times (see Eqn (2.9)):

- Garonne reach Group 1: 6 reaches corresponding to $\delta T^1 = 12.58$ days.
- Garonne reach Group 2: 1 reach corresponding to $\delta T^2 = 14.11$ days.
- Garonne reach Group 3: 18 reaches corresponding to $\delta T^3 = 1.51$ days.

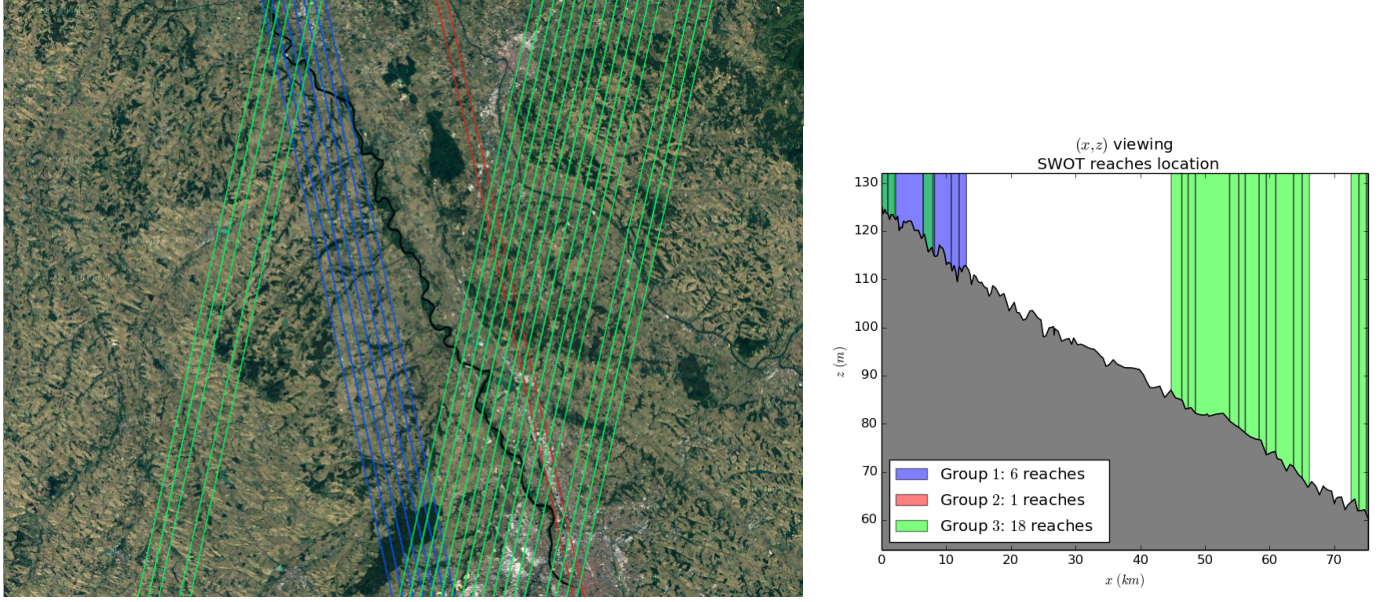


FIGURE 2.3. Location of SWOT reaches on the Garonne case.

(Left) Aerial view of Garonne (black line) and SWOT reaches location. Group of reaches #3 (green line) corresponds to ascending SWOT reaches; groups of reaches #1 and #2 (respectively blue and red lines) corresponds to descending SWOT reaches. The red line is not visible since located at the downstream boundary. Only reaches which cross over the river once have been selected.

(Right) SWOT reaches location in 1D longitudinal view (the only reach of group #2 is located at the extreme right of the domain).

A SWOT High Resolution (HR) simulator have been developed according to the expected measurement errors of SWOT, [32]. The following simulator inputs are considered to generate the synthetic SWOT measurements: a 2D hydrological raster scene (description of water elevation and topography), satellite characteristics (size of swath, satellite period, ...) and physical phenomenon (backscatter coefficient, landtypes, ...). The output data are the Single Look Complex (SLC) data and cloud point data. SLC data consists in a complex raster picture: each point of the raster gives an amplitude of the reflected signal (water or not) and a phase information (from this information it is possible to get elevation). The cloud point data corresponds to the result after processing the SLC picture; each point describes a water area with its corresponding water elevation.

For fully realistic test cases, the SWOT HR simulator can consider errors coming from the instrumental noise, the satellite rolling, the wet/dry troposphere, the ionosphere, the satellite position, layerover, see [4, 24]. The present study aims at analyzing the identifiability of the system given the extremely sparse spatio-temporal sampling of SWOT data, with either perfect data or perturbed by a random Gaussian noise only.

Let us point out that the geometry is supposed to be given, see Section 8.1 for its description, then given H^{obs} (at reach scale) the reach width w^{obs} follows straightforwardly. In the case of uncertain measurements (data are perturbed by a Gaussian noise), the noise is added to the water surface elevation only.

2.4. Cost function. The cost function J to be minimized is defined from the available measurements as follows:

$$(2.6) \quad J(\mathbf{k}) = j^{obs}(\mathbf{k}) + \gamma j^{reg}(\mathbf{k})$$

where $j^{\text{reg}}(\mathbf{k})$ is a regularizing term defined later, and $j^{\text{obs}}(\mathbf{k})$ is defined by:

$$(2.7) \quad j^{\text{obs}}(\mathbf{k}) = \frac{1}{2} \int_0^T \|\bar{H}^{\mathbf{k}}(t) - H^{\text{obs}}(t)\|_W^2 dt$$

where $\bar{H}^{\mathbf{k}}(t)$ and $H^{\text{obs}}(t)$ are defined by:

- $\bar{H}^{\mathbf{k}}(t) = (\bar{H}_0^{\mathbf{k}}(t), \bar{H}_1^{\mathbf{k}}(t), \bar{H}_2^{\mathbf{k}}(t), \dots, \bar{H}_{N_r-2}^{\mathbf{k}}(t), \bar{H}_{N_r-1}^{\mathbf{k}}(t))^T$
- $H^{\text{obs}}(t) = (H_0^{\text{obs}}(t), H_1^{\text{obs}}(t), H_2^{\text{obs}}(t), \dots, H_{N_r-2}^{\text{obs}}(t), H_{N_r-1}^{\text{obs}}(t))^T$

The norm W , W a symmetrical positive semi-definite matrix $N_r \times N_r$, N_r the number of observed reaches, corresponds to an error covariance matrix. Its extra diagonal terms $w_{i,j}$, $i \neq j$, represent the correlation of error observations between reach i and reach j ; its diagonal terms $w_{i,i}$ are the a-priori confidence on the observation of reach i .

In a real measurements case, reaches close to the satellite nadir should give observations more accurate than reaches far of the nadir. Hence, the diagonal coefficient values should depend on the distance between the reach r and the nadir. Extra-diagonal terms are extremely difficult to estimate; they are considering to be vanishing.

In all the following, the matrix W is the identity matrix of \mathbb{R}^{N_r} .

The regularization term $j^{\text{reg}}(\mathbf{k})$ is defined by:

$$j^{\text{reg}}(\mathbf{k}) = j_Q^{\text{reg}}(\mathbf{k}) + j_K^{\text{reg}}(\mathbf{k})$$

where $j_Q^{\text{reg}}(\mathbf{k})$ (respectively $j_K^{\text{reg}}(\mathbf{k})$) is the regularization term on the discharge (respectively the Manning). Typically, a low band filter can be imposed to avoid high frequency in the identified value of $Q_{in}(t)$.

The balance coefficient γ between $j^{\text{reg}}(\mathbf{k})$ and $j^{\text{obs}}(\mathbf{k})$ can be classically set following the empirical Morozov's discrepancy principle [28]. However, in the present context, the observations are so sparse that no regularization term is required.

Let $N_{t,r}$ denotes the number of SWOT observation of the reach r . Then the discrete form of the cost function J reads:

$$(2.8) \quad J(\mathbf{k}) = \frac{1}{2} \sum_{r=1, N_r} \sum_{j=1, N_{t,r}} (\bar{H}_{r,j}^{\mathbf{k}} - \bar{H}_{r,j}^{\text{obs}})^2$$

with $\bar{H}_{r,j}^{\mathbf{k}} = \frac{1}{|\Omega_r|} \sum_{i=1, N_r} H_{i,j}^{\mathbf{k}} dx$.

Let us remark that in an altimetry context, the i^{th} observation time of reach group g , t_i^g satisfies:

$$(2.9) \quad t_i^g = i\Delta T + \delta T^g$$

where ΔT is the satellite period and δT^g is the time lap of the first observation of the reach group g . Thus if a river is observed by 3 satellite passes during 1 repeat period (like it is the case for the Garonne river, see Fig. 2.3), then there are 3 different δT^g (i.e. $g = 1, 2$ or 3).

3. TEST CASES

The identifiability of the river flow model and the altimetry data assimilation are studied on a quite complete academic test case, then on a portion of Garonne river (south-west of France). The twin experiments are set as follows:

- Realistic true values of the parameters (roughness uniform in space and discharge hydrographs) are fixed. Then the forward model is run, which allows to compute the SWOT like data as described previously.
- Given these synthetic data (perturbed or not), the parameter identifiability is investigated for various temporal samplings of observations and hypothesis on the unknowns; also VDA processes are performed.

3.1. Academic test case. The aim of the test case below is to investigate the identifiability of several discharge hydrographs and/or roughness on a simple, low cost in terms of CPU time, test case. Small computational times make possible to investigate the response surface of the model. The test case geometry consists in a 1000 *m* lenght channel. Each cross-section is defined as a superposition of 5 trapeziums. The river bed elevation z_b and water surface width w are not constant; they are defined as follows:

$$z_b(x) = z_{\Delta}(x) + z_{\delta}(x)$$

with mean slopes defined by:

$$z_{\Delta}(x) = \begin{cases} 10 - 0.001x & \text{if } 0 \leq x \leq 300 \\ 9.7 - 0.004(x - 300) & \text{if } 300 < x \leq 700 \\ 8.1 - 0.002(x - 700) & \text{else} \end{cases}$$

and local bathymetry oscillations as follows:

$$z_{\delta}(x) = \begin{cases} \sum_{i=0}^4 a_n \sin(b_n(x - 50) \frac{2\pi}{T}) & \text{if } 50 \leq x \leq 950 \\ 0 & \text{else} \end{cases}$$

with $a_n = \{0.01, 0.01, 0.015, 0.02, 0.02\}$ and $b_n = \{1, 2, 4, 8, 16\}$. The triplets $(H_{i,j}, w_{i,j}, Y_{i,j})$ defined in Section 2.1 read:

$$H_{i,j} = H'_i + z_b(x_j) \text{ with } H'_i = \{1, 2, 3, 4, 5\}$$

$$(3.1) \quad w_{i,j} = \begin{cases} w'_{i,j} + \sin\left(\frac{\pi(x_j - 50)}{900}\right) & \text{if } 50 \leq x \leq 950 \\ w'_{i,j} & \text{else} \end{cases} \text{ with } w'_{i,j} = \{3, 4.9, 5.1, 6.4, 7.3\}$$

$$y_{i,j} = \{0, 0, 0, 0, 0\}$$

The academic test case geometry is shown in Fig. 3.1.

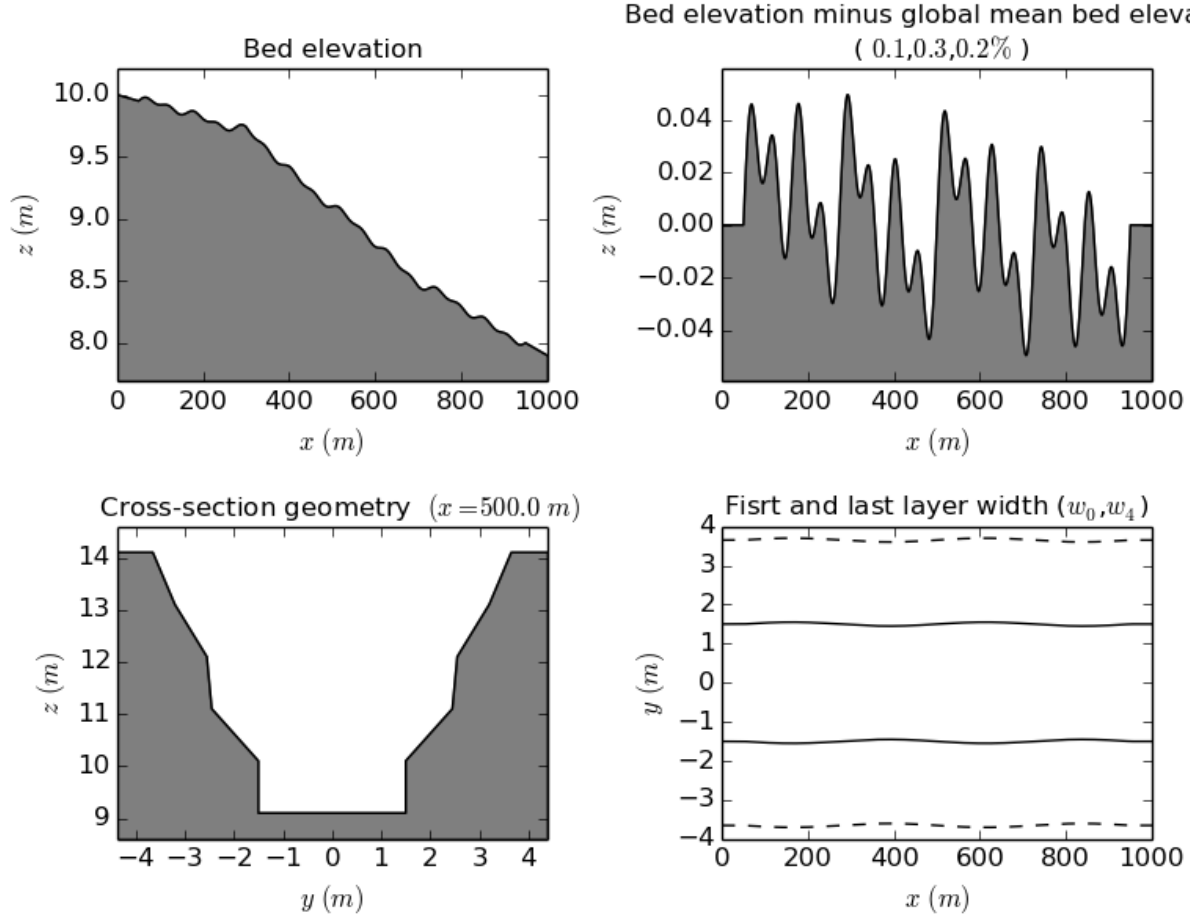


FIGURE 3.1. Geometry of the academic test case: (Left, top) Channel elevation of river vs curvilinear abscissa. (Right, top) Channel bottom elevation anomaly ($z_b - \text{mean slope}$). (Left, bottom) Cross-section geometry for $x = 500.0$ m. (Right, bottom) First and top cross sectional layer width (w_0 and w_4), see 3.1

The roughness coefficient is $K = 25 \text{ m}^{1/3} \cdot \text{s}^{-1}$. The steady flow lines, velocities and Froude numbers ($F = U/\sqrt{gh}$) are presented in Fig. 3.2 for $Q_{in} = 10 \text{ m}^3 \cdot \text{s}^{-1}$ and in Fig. 3.3 for $Q_{in} = 37 \text{ m}^3 \cdot \text{s}^{-1}$.

The downstream boundary condition consists in a power law rating curve defined by: $h_{out} = 0.45 Q_{out}^{0.6} \text{ m}$, see Fig. 3.4 (Left). This downstream law is, with the inflow discharge, a dominant control of the flow. Throughout the study, this downstream condition is given.

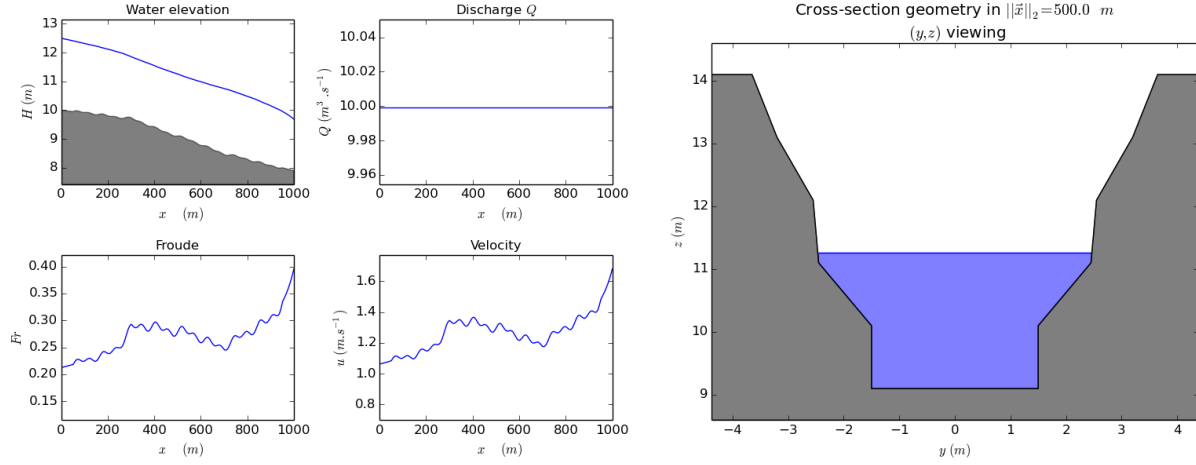


FIGURE 3.2. Academic test case. (Left) Steady state flow for $Q_{in} = 10 \text{ m}^3 \cdot \text{s}^{-1}$: (Left, top) Water elevation H (Right, top) Discharge Q . (Left, Bottom) Froude F and (Right, Bottom) Velocity U vs river curvilinear abscissa. (Right) Cross-section example (for $x = 500 \text{ m}$).

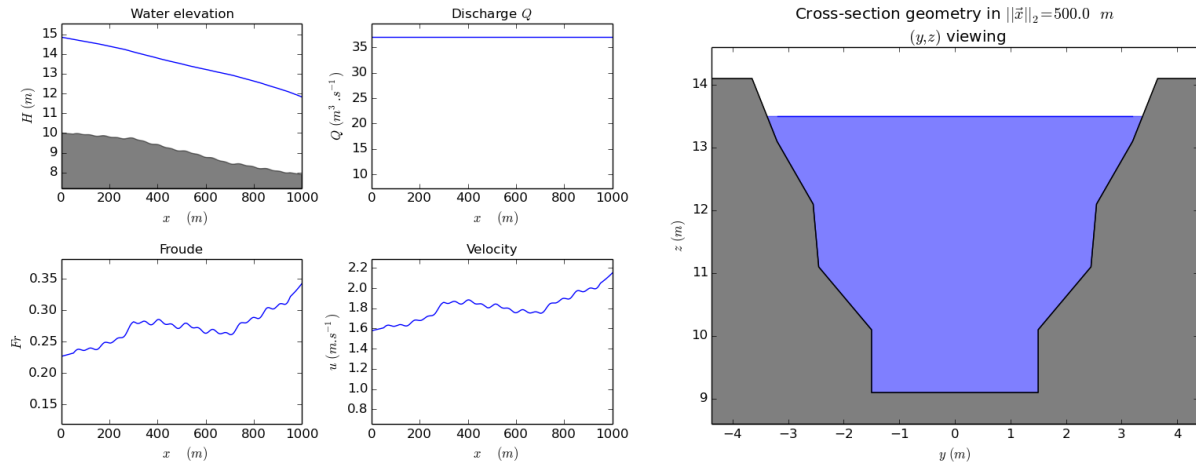


FIGURE 3.3. Academic test case. (Left) Steady state flow for $Q_{in} = 37 \text{ m}^3 \cdot \text{s}^{-1}$: (Left, top) Water elevation H (Right, top) Discharge Q . (Left, Bottom) Froude F and (Right, Bottom) Velocity U vs river curvilinear abscissa. (Right) Cross-section example (for $x = 500 \text{ m}$).

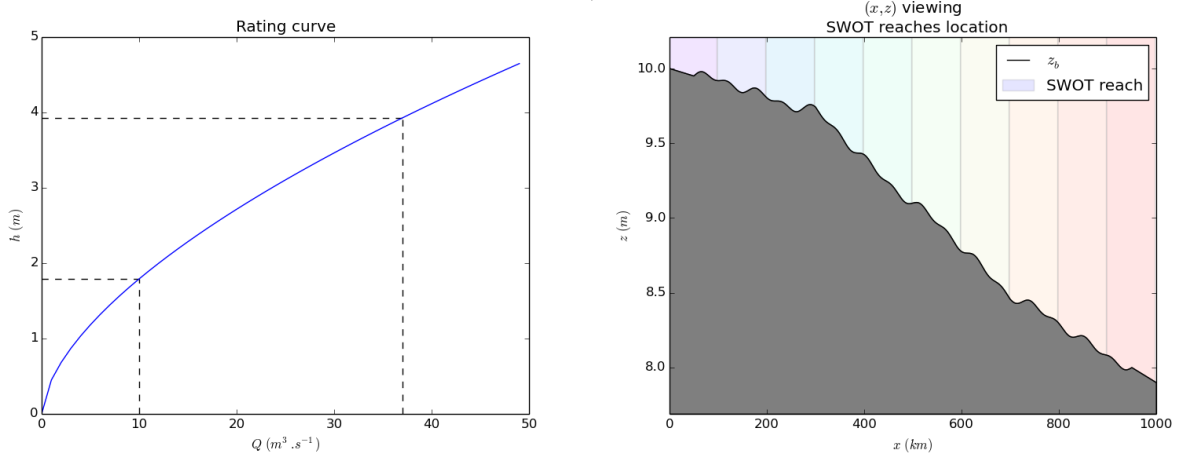


FIGURE 3.4. Academic test case. (Left) Downstream power law rating curve. (Right) Location of observation reaches: 10 reaches (Observation 1).

3.2. Garonne river test case. The zone of interest is located in between Toulouse and Malause (south west of France). The Garonne river drains an area of about 55000 km^2 from its headwaters in the Spanish Pyrenees near the Pic d'Aneto to the Gironde estuary. The 1D Garonne data considered in [35, 23, 11], consists in digital elevation model of river bathymetry as follows:

- 173 cross sections measurements from the field, distant of 56 to 2200 meters with a median value of 438 m,
- a mesh containing 1158 cross sections; they result of linear interpolations of the original 173 cross sections,
- the cross sections are merged into lidar data of banks and floodplain elevations (5 m horizontal accuracy).

The mean slope of Garonne data is -0.0866% (86.6 cm/km or $8.66 \cdot 10^{-4} \text{ rad}$). The final mesh size, i.e. the spacing between interpolated cross sections extended on banks, is between 37.26 m and 70.0 m at maximum (the average spacing being 65.034 m). The characteristics of the resulting mesh are detailed in Fig. 3.5.

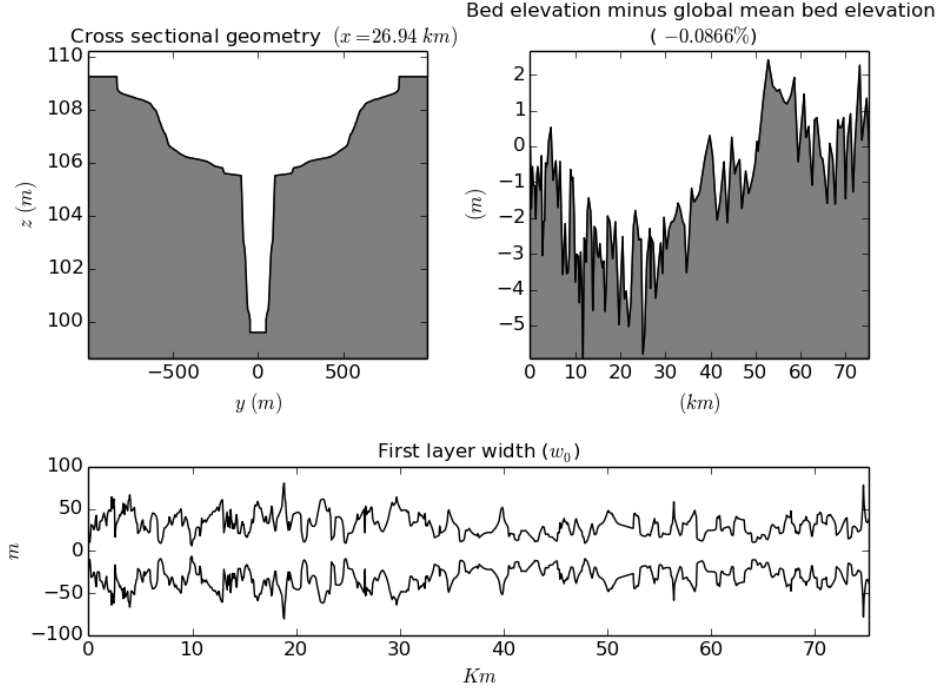


FIGURE 3.5. Display of Garonne data: (Left, top) Location of the study area. (Right, top) Normalized river bottom elevation (z_b —mean slope) in function of the river curvilinear abscissa. (Bottom) First cross sectional layer width w_0 on 151 layers (see Fig. 2.1).

4. DISCHARGE IDENTIFICATION IN THE ACADEMIC TEST CASE

This section aims at analysing the inference capabilities of the altimetric data described previously in the 1D river Saint-Venant model. As a first step, the unknown parameter is the inflow discharge $Q_{in}(t)$ only. The test case is the academic one, see Section 3.1. The different steps of the investigation are the following. From the available observations, (x, t) -identifiability maps are produced; these maps provide an excellent overview of the inference capabilities of the forthcoming VDA process. These maps are analysed in three contexts depending on the observation sparsity; they are denoted as follows:

- OD1:** (Observation Distribution #1), the whole domain is observed as indicated in Fig. 3.4 Left.
- OD2:** the observations are available at upstream and downstream only (6 reaches), see Fig. 4.5.
- OD3:** the observations are available in the middle only (4 reaches), see Fig. 4.5.

Then the inference of $Q_{in}(t)$ is performed either as usual (i.e. by identifying its values on a fixed identification grid). It is the so-called IDbasic case. The other approach is to define $Q_{in}(t)$ as a Fourier series, hence leading to a more global identification process. It is the so-called IDFourier case. In both cases, an analysis of the influence of the identification time grid is investigated.

4.1. Identifiability maps. The observations are generated from a hydrograph adapted to the geometry respecting realistic discharge magnitudes and time scales, see Q_{in}^{real} in Fig. 4.3 Top-Left. The hydrograph indeed creates a similar flooding with a time scale in adequation with the total simulation time (in other words the flood wave has time enough to propagate throughout the domain).

The roughness coefficient is constant, $K = 25 \text{ m}^{1/3} \cdot \text{s}^{-1}$ ($\alpha = 25$ and $\beta = 0$ in (2.2)), the observation time step is $dtobs = 100 \text{ s}$. The discharge first guess is chosen constant in time: $Q = 10 \text{ m}^3/\text{s}$.

An original reading of the inverse problem consists to plot so-called identifiability maps in the plan (x, t) . Since the inflow discharge ($Q_{in}(t)$ defined at $x = 0$) is a sought input parameter, the important wave velocity

is the positive one i.e. $(u + c)$. Recall that without the source terms (gravity waves model), the 1D Saint-Venant model wave velocities are $(u - c)$ and $(u + c)$, while the diffusive wave velocity is $\frac{5}{3}u$ (in the case of a wide channel), see e.g. [36, 9, 29] and Remark 1.

Then for each reach r ($N_r = 10$ in the case OD1) and for each time observation t_i^r (11 in the case OD1), the velocity waves of the 1D Saint-Venant model and the diffusive wave model are estimated, then plotted, Fig. 4.1. To do so, \bar{u}_i^r and \bar{c}_i^r corresponding to the reach r at time i are estimated, where \bar{u} denotes the mean velocity value and $\bar{c} = (g\bar{h})^{1/2}$ with \bar{h} the mean water depth (deduced from the surface elevation measurements and the given bathymetry).

Let us point out that in the present synthetic data experiments (twin experiments), \bar{u} is known. In a realistic context, \bar{u} can be estimated from a 0.5D model (based on the Manning-Strickler equation written for each reach) as it is done for example in [11].

The (r, i) observation time interval is defined as follows: $T_{r,i}^w = [t_i^r - L_r/(\bar{u} + \bar{c})_i^r, t_i^r]$ with L_r the reach length and t_i^r the observation time.

Then in Fig. 4.1 each observation space time window $T_{r,i}^w$ is plotted (in color). Each rectangle diagonal corresponds to the local $(\bar{u} + \bar{c})$ line; indeed the height of $T_{r,i}^w$ corresponds to $(\bar{u}_i^r + \bar{c}_i^r)$ magnitude. One can notice that the variation in space time of $(\bar{u} + \bar{c})$ is not significant, see the rectangle height variations and Table 1.

Let us point out that all the domain is observed at $t = 0$ hence the wave velocity $(\bar{u} + \bar{c})$ at $t = 0$ can be estimated accurately.

The identifiability map in (x, t) is plotted for the three cases depending on the observation sparsity: cases OD1, OD2 and OD3, see Fig. 4.1. The rectangle colors represent the misfit to the equilibrium between gravity and friction forces (in norm 1); indeed it is the right-hand side (or source term) magnitude of the momentum equation, see (2.1):

$$(4.1) \quad \text{"equilibrium misfit"} = \text{norm}_1 \left[g \int_0^h (h - z) \frac{\partial \tilde{w}}{\partial x} dz - gS \left(\frac{\partial z_b}{\partial x} + S_f \right) \right]$$

If this source term vanishes (blue colors), it means that the flow is locally at equilibrium, uniform; the Manning-Strickler equation would be sufficient to model locally the flow. On the contrary if the misfit term becomes important (e.g. orange - red colors) then the hyperbolic feature of the model is relatively important.

In terms of energy, this source term representing the misfit to the equilibrium is a dissipative term; while the left-hand side of the 1D Saint-Venant model is conservative, hyperbolic, see (2.1). Then in Fig. 4.1 the colors provide a rough estimation of the propagation and dispersion features of the flow model.

Typically, the peak time at inflow is represented by the rectangle $(r, i) = (1, 6)$; then the corresponding wave velocity $(\bar{u} + \bar{c})$ is faster than those arising from the middle of the domain, rectangle $(6, 6)$ (hence both lines are not parallel). To illustrate the transport-diffusion phenomena corresponding to Fig. 4.1, the discharge throughout the domain is plotted at the three observations times 400s, 500s (peak time at inflow) and 600s in Fig. 4.2.

All these represented information in the (x, t) plan represents the so-called *identifiability map*, in particular with respect to the inflow discharge $Q_{in}(t)$ (i.e the control of $Q(t)$ at $x = 0$).

Typically, in the present case, the identifiability maps show that any change on Q_{in} is observed at least few times at some locations; and this is true for the three cases OD1, OD2 and OD3, see Fig. 4.1. In other words, there is no blind time-space window, indeed the same hydraulic information is even observed few times. Then the forthcoming identification experiments based on VDA should be robust and accurate for the complete simulation time range $[0, T]$. This a-priori analysis are confirmed by the VDA experiments presented below.

Let us remark that the present source term estimation provides an a-posteriori model error if employing the usual Manning-Strickler equation to model the flow.

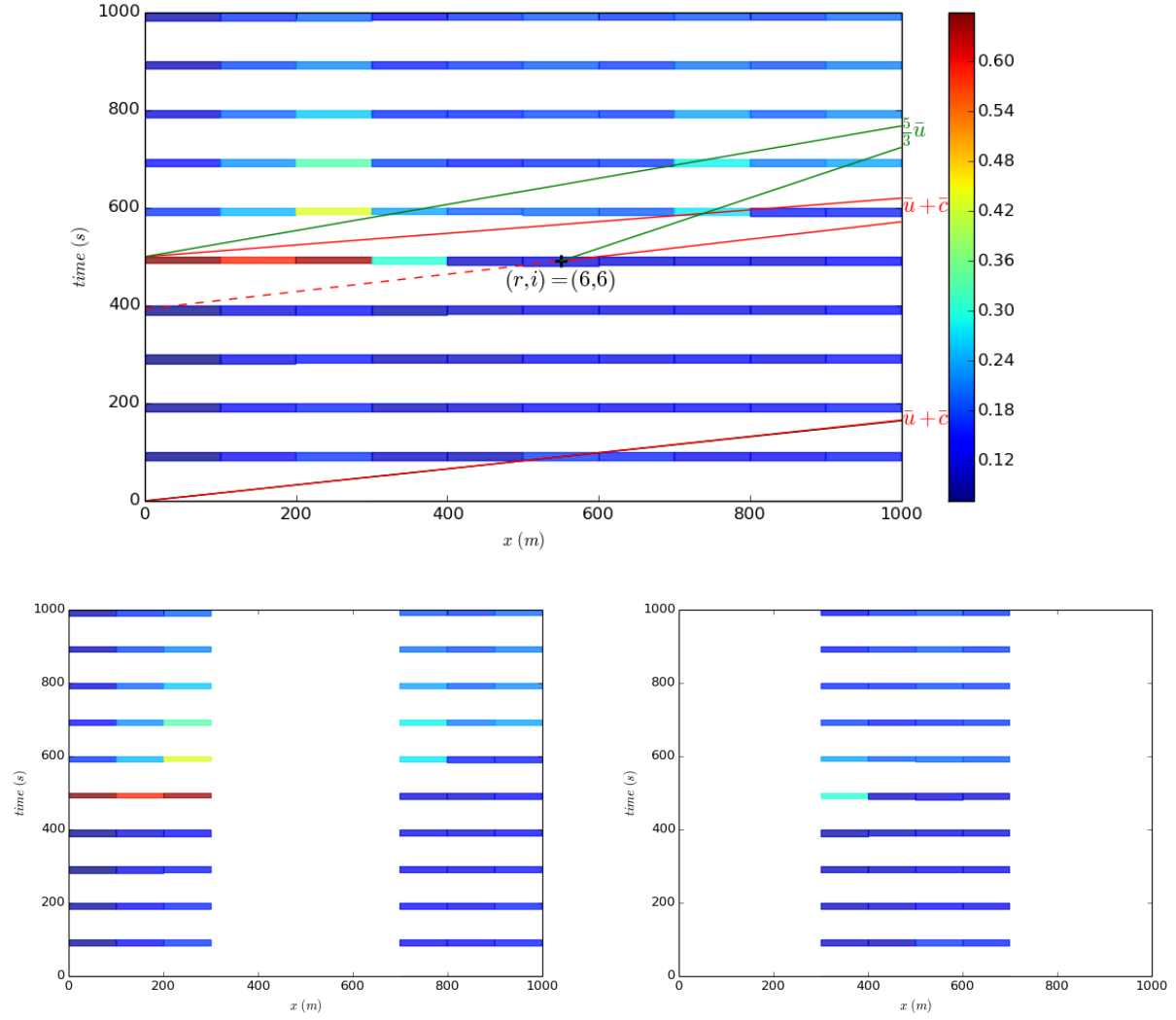


FIGURE 4.1. Up. The identifiability map in (x, t) in the case: (Top) OD1 (full observations); (Bottom, Left) OD2; (Bottom, Right) OD3

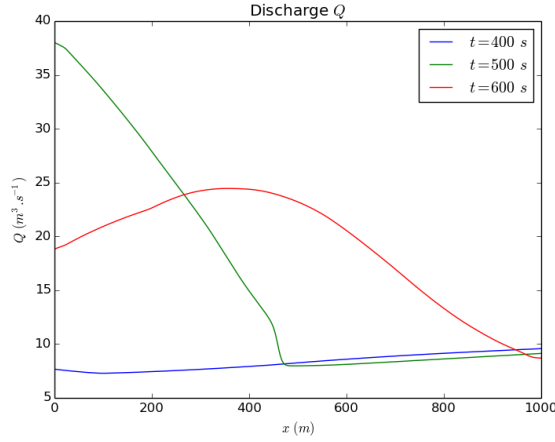
The estimated wave velocities for $OTSD(r, i)$ are plotted in black lines. For the same $OTSD(r, i)$, the dashed lines denote the incoming velocity waves.

The red lines correspond to the first (resp. the last) reach at the first (resp. the last) observation time.

The misfit term magnitude is represented in each rectangle by the colors

Statistics	mean μ	standard deviation σ
$\min(u + c)$	6.3027	0.6063
$\text{mean}(u + c)$	6.3521	0.6224
$\max(u + c)$	6.4028	0.6394
$\min(u)$	1.3018	0.2916
$\text{mean}(u)$	1.376	0.2972
$\max(u)$	1.4361	0.3069

TABLE 1. Academic test case. Statistics on the wave velocities (in space-time); Gaussian distribution.

FIGURE 4.2. Discharge $Q(x, \cdot)$ vs x at three observations times: 400s, 500s (peak time at inflow), 600s.

4.2. IDbasic identification for various dta . In this section few identifications of inflow discharge are performed with a fixed observation time step $dt_{obs} = 100$ s and various assimilation time steps dta , ranging from $1/10$ to 1 dt_{obs} . The inflow discharge value is identified at each assimilation time step following the inflow discharge definition IDbasic, see Section 2.1. The parameter vector is $\mathbf{k} = (Q_1, \dots, Q_p)^T$ with $dta = (t_{i+1} - t_i) = \forall i \in [1..p - 1]$.

The inflow discharge and the gradient value are plotted in Fig. 4.3 Top, for $dta = dt_{obs}/10$ and $dta = dt_{obs}$. In the case $dta = dt_{obs}/10$ the result is excellent, while if $dta = dt_{obs}$, the accuracy remains good (excepted at peak time, $t \approx 500$ s). The convergence curve of the cost function J is similar than those presented in Fig. 5.1 Right-top. These two experiments have converged in 45 and 17 iterations respectively.

The errors on the identified inflow discharge are plotted in Fig. 4.3 Bottom. Both the 2-norm error and $(1 - E)$, E the Nash–Sutcliffe criteria, are the lowest for dta between 20 s and 50 s = $dt_{obs}/2$ (with $(1 - E) \sim 0.0077$).

Roughly, the error is improved if $dta < dt_{obs}$ but not too small. Indeed, for $dta \ll dt_{obs}$, typically $dta = dt_{obs}/10$, over and under estimations of the discharge appear. Indeed, if dta is small, any change of Q_{in} between two identification times, is not observed, Fig. 4.3 Left-top.

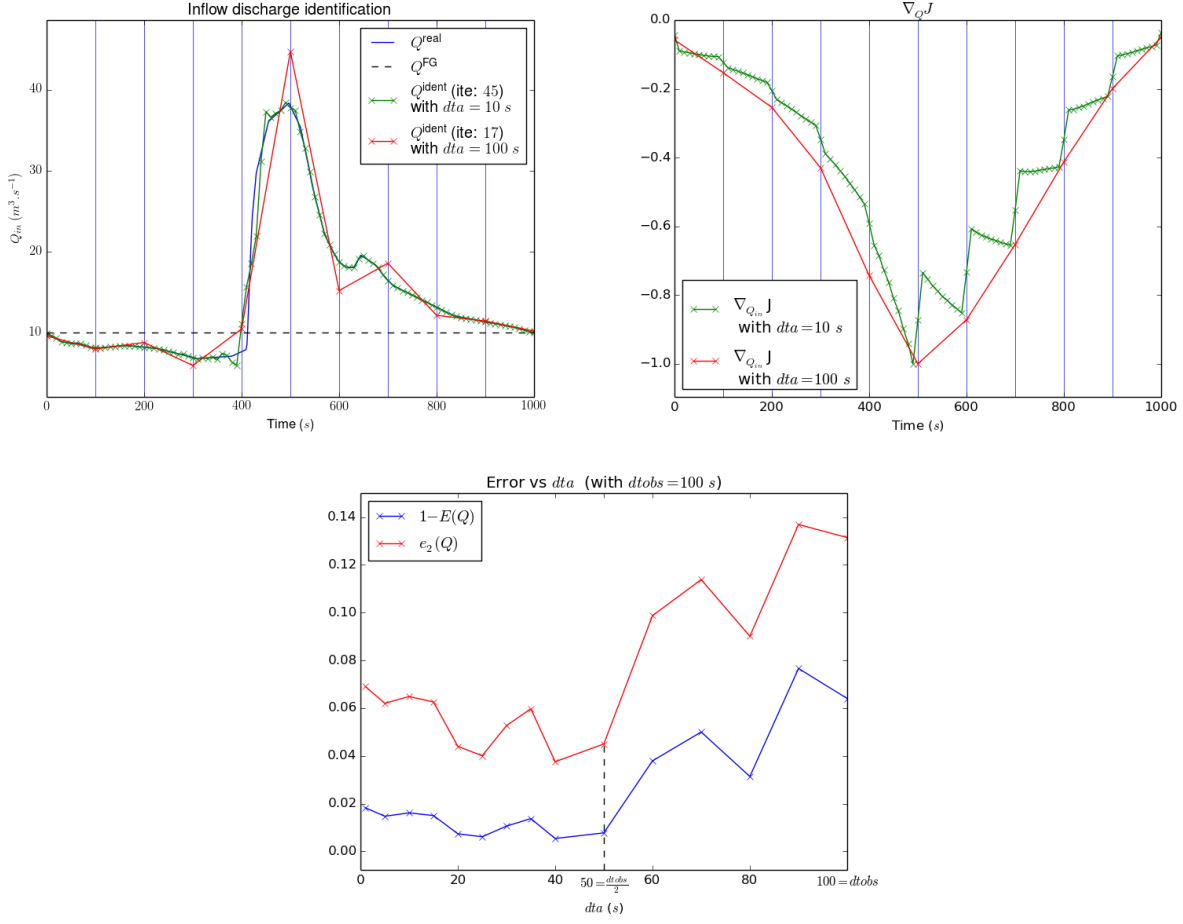


FIGURE 4.3. Discharge identification: IDbasic approach with $dtobs = 100$ s. (Left, top) Discharge identification with $dta = dtobs = 100$ s and $dta = dtobs/10 = 10$ s. (Right, top) Normalized gradient $\nabla_Q J$ with $dta = dtobs$ and $dta = dtobs/10 = 10$ s. (Bottom) Errors vs dta .

4.3. IDFourier identification for various dta . The same discharge identification experiments as previously are performed but with inflow discharge represented in a reduced basis. Here $Q_{in}(t)$ is approximated by a Fourier series, see IDFourier in Section 2.1 and (2.3). The parameter vector reads: $\mathbf{k} = (a_0, a_1, b_1, \dots, a_{N_{FS}}, b_{N_{FS}})^T$. Then the VDA process consists in retrieving these Fourier coefficients; it is done for $N_{FS} = 7$ and $N_{FS} = 25$. Roughly the same accuracy and behaviors as in the previous case are obtained, see Fig. 4.4.

The minimal errors (e.g. $(1 - E) \sim 0.005$) occur for $T/N_{FS} = 40$ and 60 , see Fig. 4.4 right.

An advantage of identifying $Q_{in}(t)$ as a Fourier series is that the optimal solution is always smooth; hence this circumvents the potential oscillations obtained in the case IDbasic with $dta \ll dtobs$.

4.4. Identification robustness vs observation sparsity. A VDA process is global in time. The previous numerical experiments demonstrate that refining too much the identification time grid dta compared to $dtobs$ (typically $dta = dtobs/10$) deteriorates the identification accuracy. In other words, given an observation time grid, the identification of the time dependent inflow discharge cannot be obtained at much finer time scale. All these previous experiments have been performed with observations available in the whole domain (case

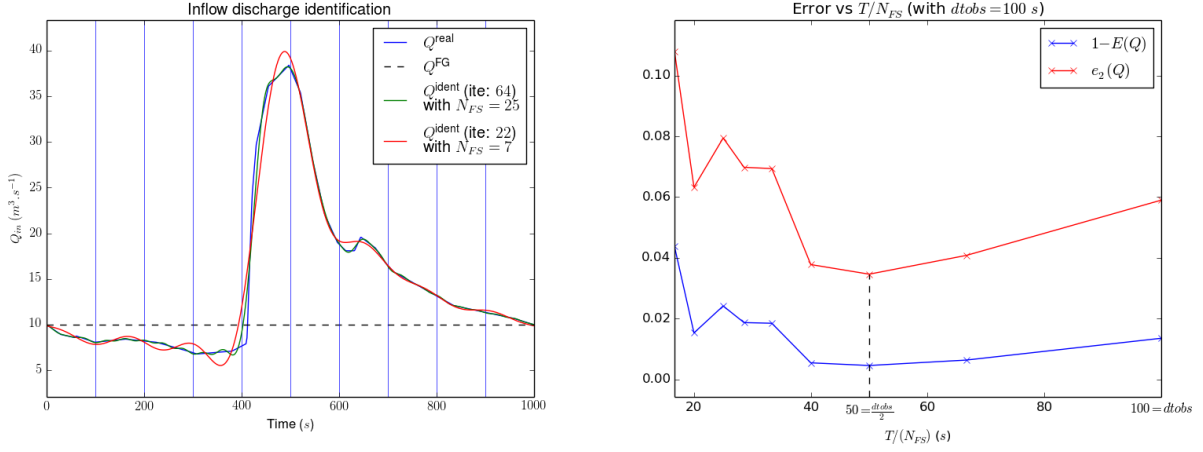


FIGURE 4.4. Discharge identification: IDFourier approach (Fourier series reconstruction) with $dt_{obs} = 100$ s: (Left) Discharge identification with $N_{FS} = 7$ and $N_{FS} = 25$. (Right) Errors vs T/N_{FS} .

OD1, see the introduction of Section 4). In a real case (e.g. SWOT data of Garonne river test, see Section 2.3) the observations are not available for the whole domain at the same time. In the present experiments, the robustness and accuracy of the discharge identification is investigated if considering real-like SWOT data i.e. less sparse observations.

The inflow identification are performed with a pseudo-optimal assimilation time step $dta = 25$ s (still with $dt_{obs} = 100$ s) for the three cases OD1, OD2, OD3, see Fig. 3.4 Left and Fig. 4.5.

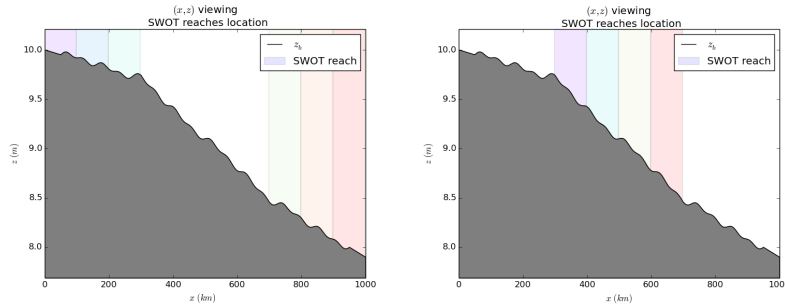


FIGURE 4.5. Location of the observation reaches. (Left) Case OD2: observations are located at upstream and downstream (6 reaches). (Right) Case OD3: observations are located in the middle (4 reaches).

As discussed in Section 4.1, the identifiability maps (see Fig. 4.1) indicate that for all the three cases the identification should be accurate. Indeed, the numerical results obtained by VDA confirm this a-priori analysis since the error is extremely low, typically $(1 - E) > 0.99$, see Tab. 2. Since the results are similar to the previous case, the identified inflows are not plotted.

	OD 1	OD 2	OD 3
Nash-Sutcliffe coefficient (E)	0.993	0.994	0.991

TABLE 2. Nash-Sutcliffe coefficient (E) for $dta = 25$ s in function of the observations availability; cases OD1, OD2, OD3.

5. DISCHARGE AND ROUGHNESS IDENTIFICATION IN THE ACADEMIC TEST CASE

In the previous section, the inflow discharge $Q_{in}(t)$ only was inferred. In the present section both the time-dependent inflow discharge and the roughness coefficient K (time-independent) are inferred. Let us recall that K is defined by $K = \alpha h^\beta$. Then the control vector reads: $\mathbf{k} = (Q_{in,1}, Q_{in,2}, \dots, Q_{in,p}, \alpha, \beta)^T$ in the IDbasic case and $\mathbf{k} = (a_0, a_1, b_1, \dots, a_{N_{FS}}, b_{N_{FS}}, \alpha, \beta)^T$ in the IDFourier case (Fourier series reconstruction). The synthetic observations are generated from the same hydrograph (inflow discharge) as previously and an uniform roughness coefficient $K = 25$ i.e. $\alpha = 25$ and $\beta = 0$ in Eqn (2.2). Let us point out that similar results have been obtained for a non vanishing value of β (i.e. for K depending on the water depth h).

First guesses are respectively chosen equal to $Q_{in}(t) = 100 \text{ m}^3 \cdot \text{s}^{-1}$ for all t , and to $(\alpha, \beta) = (23.5, 0.1)$ (hence considering K depending on h). The observations are available in the whole domain: case OD1.

5.1. Identifications in the IDbasic and IDFourier cases. The identified inflow discharge with a basic linear reconstruction (case IDbasic) is as accurate as in the previous case (i.e. while identifying Q_{in} only). The identified discharge are plotted in Fig. 5.1 Left top in the case $dta = 10$ s and $dta = 100$ s.

For $dta = 10$ s, the identification of the roughness parameters α and β is accurate, see Fig. 5.1 Bottom; the minimization algorithm has converged in 64 iterations. On the contrary, for $dta = 100 = dtobs$, the minimization algorithm has difficulties to converge, see Fig. 5.1 Top right.

In the case IDFourier (Fourier series reconstruction of $Q_{in}(t)$), the results are quite similar, see Fig. 5.2.

In both cases (IDbasic and IDFourier), the identified quantities are accurate if the identification time step dta is small enough (compared to $dtobs$), or if the Fourier mode number N_{FS} is large enough. In such cases, the identification of $Q_{in}(t)$ is as accurate and as robust as in the previous case (where Q_{in} only was identified).

But if $dta = dtobs$ or equivalently if N_{FS} is small enough, then the minimization algorithm has difficulties to converge, hence the VDA process provides quite unaccurate quantities.

The errors on the roughness coefficients are plotted in Fig. 5.3. Again the best accuracy are obtained for dta (resp. N_{FS}) small enough but not much (resp. large enough). Since the error behaviors made on the identified discharge are very similar than in the previous case, see Fig. 4.3 Bottom and Fig. 4.4 Right, they are not plotted.

5.2. Sensitivity with respect to the first guesses. The sensitivity of the identified quantities (Q_{in} and (α, β)) with respect to the first guess values $Q_{in,FG}$ and $(\alpha, \beta)_{FG}$ is investigated. To do so, the response surfaces of the identified values are examined in terms of Nash-Sutcliffe criteria E for the discharge and in terms of relative errors for the roughness parameters, see Fig. 5.4. All the computations are made in the OD1 case (complete observations) and in the IDFourier case with $N_{FS} = 20$.

Three sensitivity maps are presented, see Fig. 5.4. For each of them the parameter β is fixed ($\beta = 0$); then each figure corresponds to the error of one of the three identified parameter vs the first guess values of α and Q_{in} .

The three sensitivity maps show that the identification of inflow discharge $Q_{in}(t)$ and the roughness coefficients are accurate for a large value range of $Q_{in,FG}$. However the accuracy is less high for low values of $Q_{in,FG}$.

The results are very similar if the fixed parameter is the discharge Q_{in} or the roughness parameter α , that is the reason why the corresponding figures are not presented.

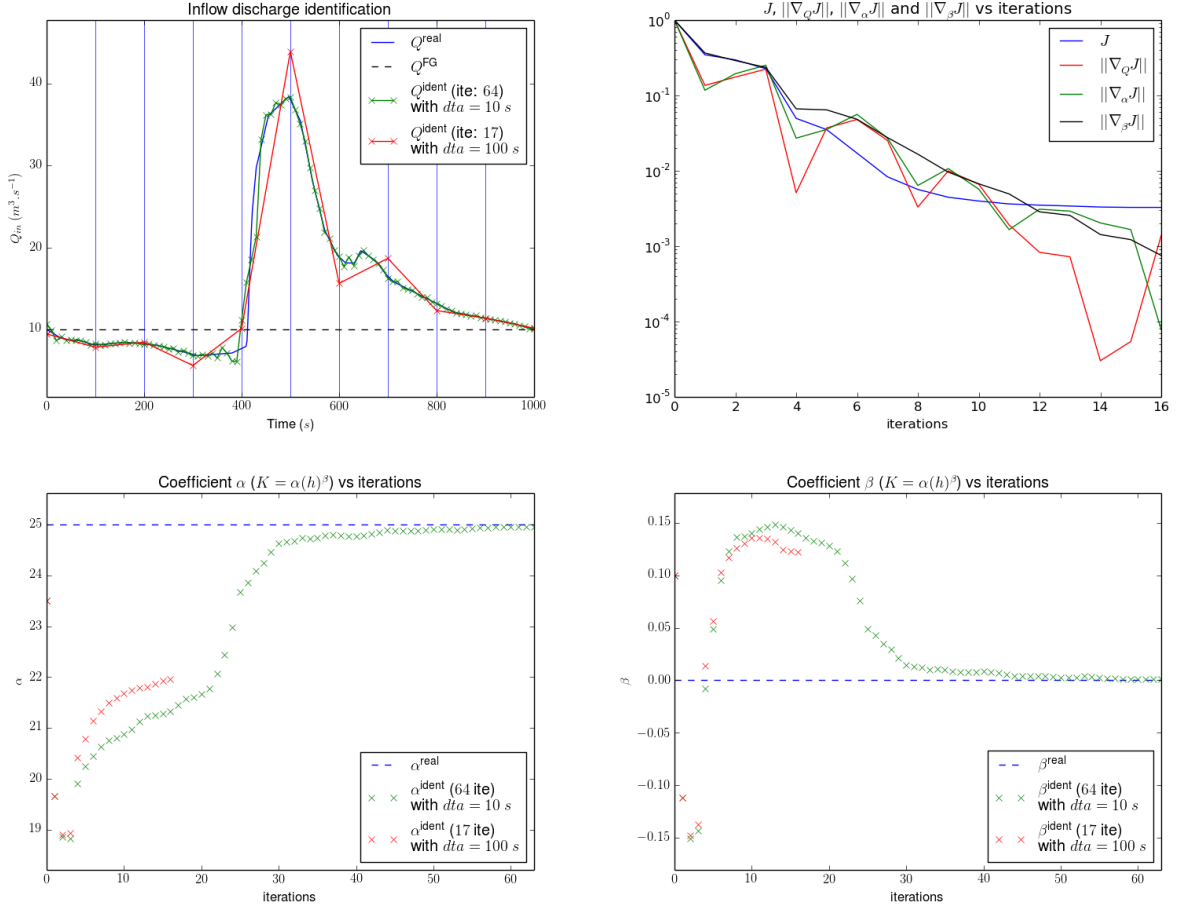


FIGURE 5.1. Discharge and roughness identification in the academic test case (IDbasic case). (Left, top) Discharge identification with $dta = dtobs = 100$ s and $dta = dtobs/10 = 10$ s. (Right, top) Function cost $J, ||\nabla_Q J||, ||\nabla_\alpha J||$ and $||\nabla_\beta J||$ vs minimization iterations. (Left, bottom) Roughness coefficient α vs minimization iterations. (Right, bottom) Roughness coefficient β vs minimization iterations.

5.3. Sensitivity with respect to the observation errors. Finally the impact of observation errors on the three identified quantities are presented in Fig. 5.5. A Gaussian noise $\mathcal{N}(0, \sigma)$ is added to the water elevation data H^{obs} . In the case $\sigma = 0.1$ m (which corresponds to the expected error of the forthcoming SWOT instrument), the error on the roughness parameters (α, β) equals approximately 5% and the Nash–Sutcliffe criteria $E \approx 0.5$. In a bad observational context with $\sigma = 0.5$ m, the error on the roughness parameters (α, β) equals approximately 10 – 25% and the Nash–Sutcliffe criteria $E(Q)$ becomes negative.

Therefore, the identification of the composite control parameter $(Q_{in}(t); \alpha, \beta)$ turns out to be quite sensitive to the observation errors but its inference remains accurate in the case of the forthcoming SWOT instrument accuracy ($\sigma = 0.1$ m).

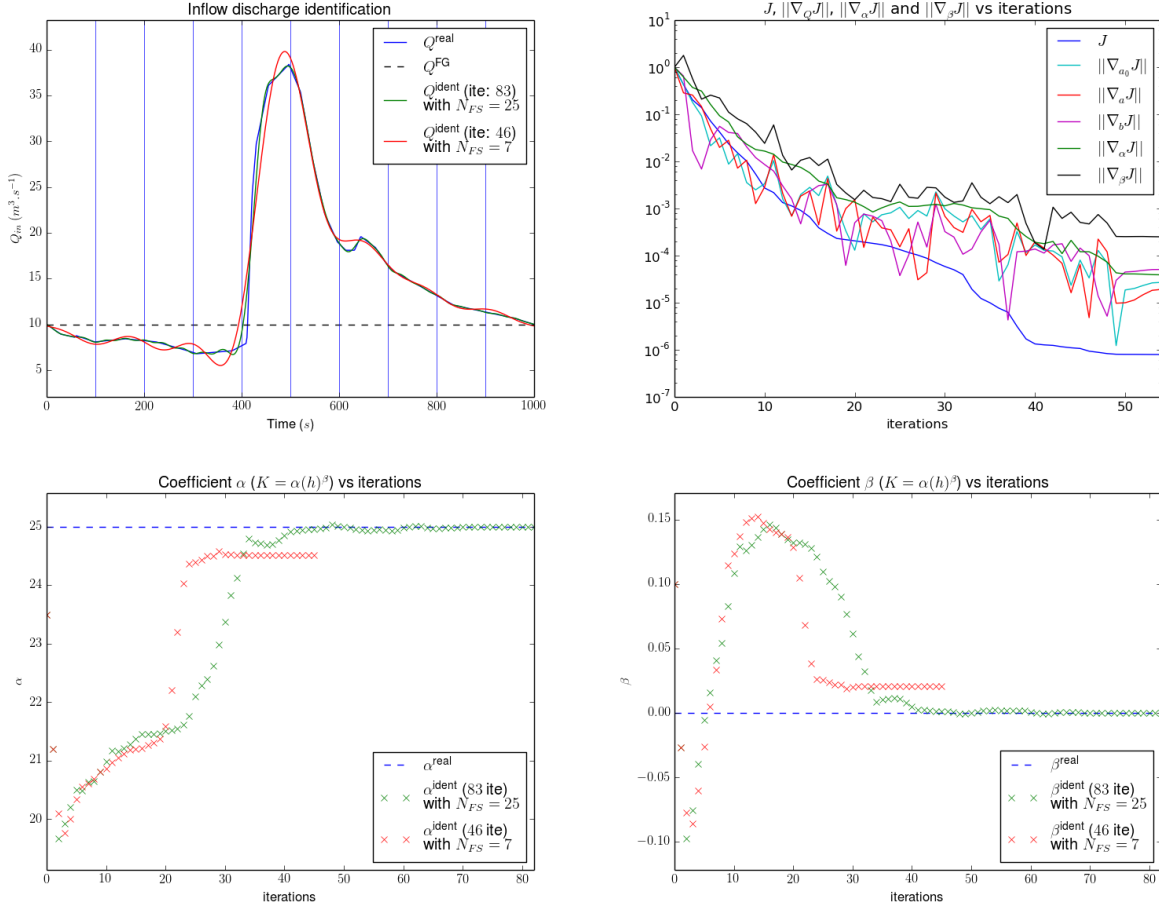


FIGURE 5.2. Discharge and roughness identification in the academic test case (IDFourier case). (Left, top) Discharge identification with $N_{FS} = 7$ and $N_{FS} = 25$. (Right, top) Function cost $J, ||\nabla_{a_0} J||, ||\nabla_{a_n} J||, ||\nabla_{b_n} J||, ||\nabla_{\alpha} J||$ and $||\nabla_{\beta} J||$ vs minimization iterations. (Left, bottom) Roughness coefficient α vs minimization iterations. (Right, bottom) Roughness coefficient β vs minimization iterations.

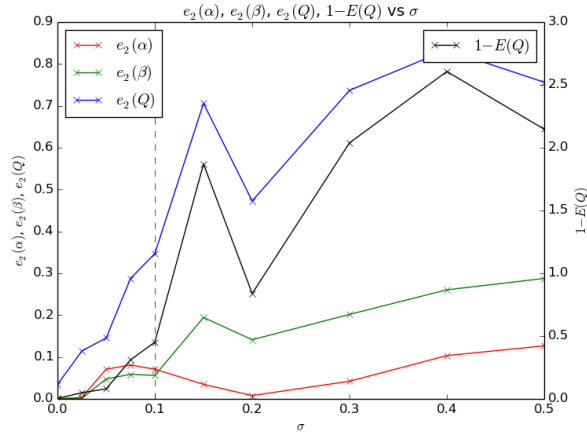


FIGURE 5.5. Error on the identified quantities with $\mathbf{k} = (a_0, a_1, b_1, \dots, a_{N_{FS}}, b_{N_{FS}}, \alpha, \beta)^T$ vs the observation error σ (standard deviation of the Gaussian noise). The vertical dashed line represents the expected error of the SWOT mission, both in norm 2 and Nash-Sutcliffe criteria.

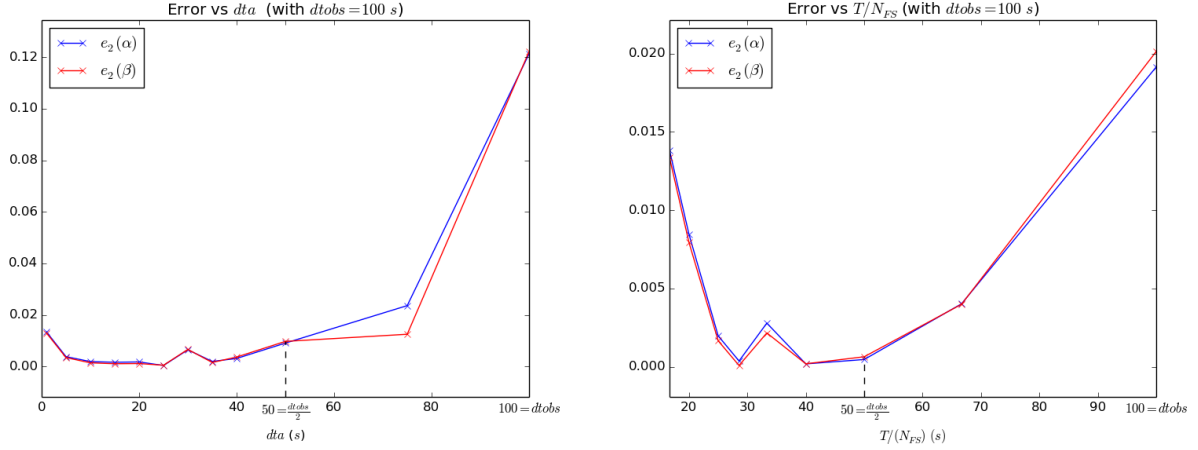


FIGURE 5.3. Discharge and roughness identification in the academic test case: errors e_2 on (α, β) . (Left) IDbasic case: errors vs dta . (Right) IDFourier case: errors vs T/N_{FS} .

6. GARONNE RIVER TEST CASE

The accuracy and the robustness of the same VDA process as before, see sections 2.2 and 2.4, is investigated in a realistic data context. The test case is the Garonne river (portion downstream of Toulouse) presented in Section 3.2. The hydrograph corresponds to the discharge in Toulouse for a 80 days period during the year 2010. The SWOT-like observations are generated by the model following the method presented in Section 2.3; the roughness coefficient is constant: $K = 25 \text{ m}^{1/3} \cdot \text{s}^{-1}$ ($\alpha = 25$ and $\beta = 0$ in (2.2)). For the VDA computations the first guess $Q_{in,FG}$ is chosen constant and equal to $268 \text{ m}^3/\text{s}$ (the mean value of the true hydrograph), see the horizontal dotted lines in the inflow discharge graphs, Fig. 6.3.

As a first step and following Section 4.1, the identifiability maps are computed. These preliminaries analyses show the forthcoming difficulties, or even the impossibility, to infer the inflow discharge from the altimetric SWOT like data only. Then three different scenari are performed and analysed.

Scenario 1 (Section 6.2) is based on a real SWOT temporal sampling as defined in Section 2.3 (for a simulation time $T = 80$ days); Scenario 2 (Section 6.3) is based on a densified SWOT temporal sampling by a factor 1.5 and 2 (i.e. for a simulation time $T = 8$ days); Scenario 3 (Section 6.4) is based on a densified SWOT temporal sampling by a factor 100 i.e. for a simulation time $T = 0.8$ day.

6.1. Identifiability maps. The identifiability maps in (x, t) are computed from the available observations, see Fig. 6.1, following the approach described in Section 4.1. These identifiability maps are presented for the real like case (Scenario 1), see Fig. 6.2 Left) and Scenario 3, see Fig. 6.1.

Let us point out that contrarily to the academic test case, no observation is available at $t = 0$ hence the wave velocity $(\bar{u} + \bar{c})$ propagating from $t = 0$ cannot be estimated.

Fig. 6.2 Left (identifiability map based on the velocity waves without the friction effects) show that in the true sampling case (Scenario 1), the identifiability of $Q_{in}(t)$ is limited to the observation time, hence extremely limited compared to the total simulation intervall $[0, T]$. Then in the forthcoming VDA experiments (next paragraph), a global identification of $Q_{in}(t)$ in a Fourier series basis (IDFourier case) seems preferable compared to a “sequential” one (IDbasic case) although whatever the $Q_{in}(t)$ basis, the VDA process should not be able to fullfill the “blind” space-time windows with physical information.

Next, the 100 times more dense scenario (Scenario 3) is analysed. In this case, see Fig. 6.1, the inflow discharge identifiability is represented by the vertical dashed lines at $x = 0$: in red those coming from the “far” green observed reaches (hence identifiability information should not be really reliable); in black those coming from the close blue observed reaches (hence much more reliable). Recall that this identifiability information is based on the wave velocities estimations only, hence the dissipation due to the source term

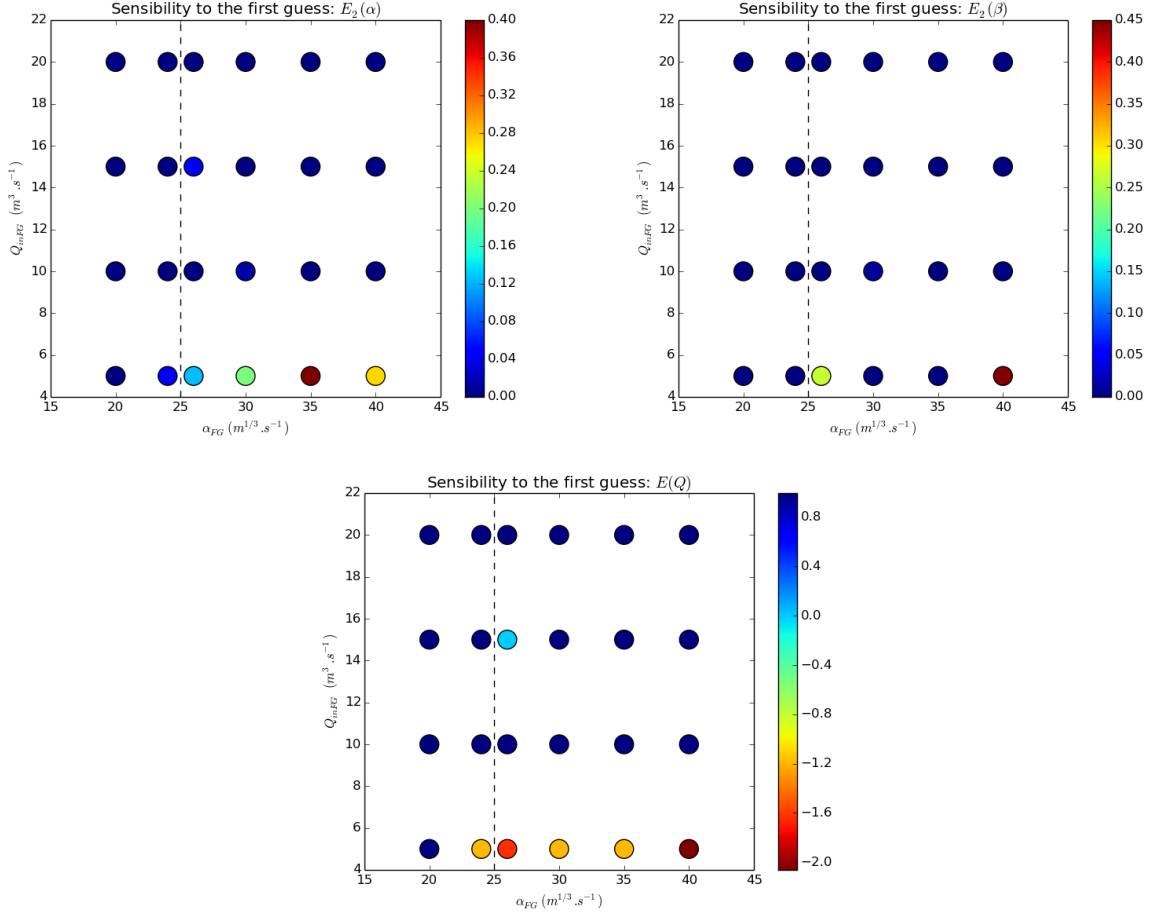


FIGURE 5.4. Sensitivity to the first guess: errors on the identified quantities vs α_{FG} and $Q_{in,FG}$ (β is fixed). (Left, top) Error $e_2(\alpha)$. (Right, top) Error $e_2(\beta)$. (bottom) Error on Q_{in} : Nash-Sutcliffe criteria E .

(roughly the Manning-Strickler equation) is not taken into account. Then this figure indicates that a large proportion of inflow values should be identifiable (see the vertical points at upstream). The forthcoming VDA experiments confirm the present preliminary analysis, see e.g. Fig. 6.5, where the dashed vertical lines (red and black) are taken back on the identified discharge graphs.

The identifiability map of Scenario 2 is similar to the identifiability map of Scenario 1 but with larger identifiable windows (r, i) (respectively 1.5 and 2 times more dense); then it is not plotted here.

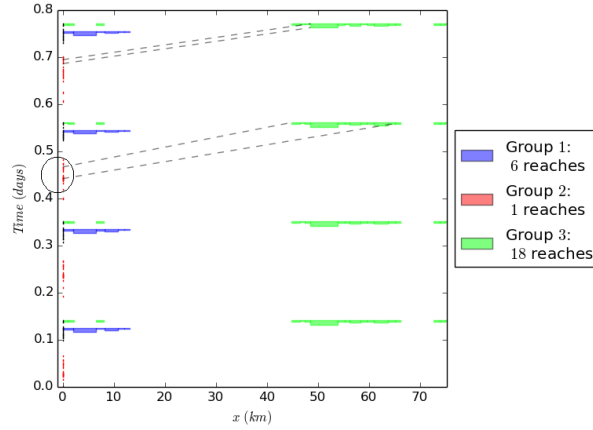


FIGURE 6.1. Identifiability maps in (x, t) for the Garonne river in the Scenario 3 case (temporal sampling densified by a factor 100). The velocity waves $(\bar{u} + \bar{c})$ (dotted lines) are estimated at each reach from the available observations. The rectangle heights are proportional to the local value $(\bar{u} + \bar{c})$. The dashed vertical lines at upstream represent the characteristic feet i.e. the sets of points which can be identified in the model without source term: in red those coming from the “far” green observed reaches (hence identifiability information not fully reliable); in black those coming from the close blue observed reaches (hence much more reliable). The circle centered at $t \approx 0.45$ days corresponds to the inflow peak.

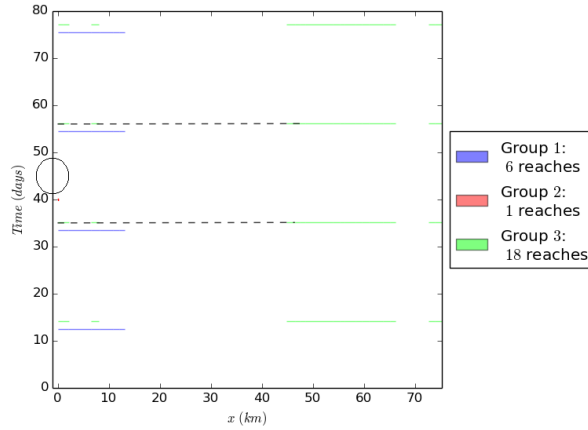


FIGURE 6.2. Identifiability maps in (x, t) in the Garonne river case, Scenario 1: true temporal sampling. The “identifiability lines” (the characteristics feet) are so tiny that they are not visible in the figure. The circle centered at $t \approx 0.45$ days corresponds to the inflow peak.

Statistics	mean μ	standard deviation σ
$\min(u + c)$	5.4502	0.6805
$\text{mean}(u + c)$	6.0739	0.6561
$\max(u + c)$	6.6827	0.8052
$\min(u)$	0.7198	0.2507
$\text{mean}(u)$	1.1023	0.1968
$\max(u)$	1.391	0.2607

TABLE 3. Garonne test case. Statistics on the wave velocities (in space-time); Gaussian distribution.

6.2. Scenario 1: real SWOT temporal sampling. Given the bathymetry z_b and the roughness coefficient K , the inflow discharge is identified by VDA from the real SWOT space time sampling. Following the preliminary study based on the identifiability maps, $Q_{in}(t)$ is decomposed as a Fourier series (IDFourier case) with $N_{FS} = 5$ (Fig. 6.3 Left) and $N_{FS} = 10$ (Fig. 6.3 Right). Then as expected, the identification is accurate in the vicinity of each observation (the vertical colored lines in Fig. 6.3) but unaccurate elsewhere.

The too important sparsity of the SWOT data prevents to constraint sufficiently the inflow discharge. Also as expected, increasing the identification frequency (case $N_{FS} = 10$) does not improve the coarser approximation ($N_{FS} = 5$) since the latter already corresponds to an adequate frequency compared to the observation mean frequency, see Fig. 6.3. In short and following the identifiability map, only the values inferred in the vicinity of the observation times are reliable.

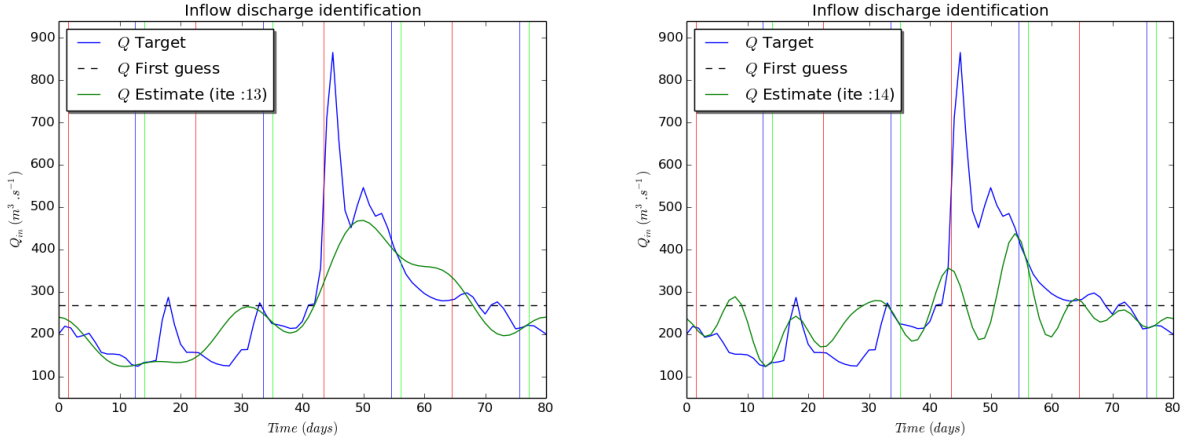


FIGURE 6.3. Garonne river, Scenario 1. Discharge identification with Fourier series with: (Left) $N_{FS} = 5$. (Right) $N_{FS} = 10$.

Vertical lines corresponds to the time observations (blue for Group 1, red for Group 2 and green for Group 3). The horizontal dotted line corresponds to the first guess.

6.3. Scenario 2: densified SWOT temporal sampling by a factor 1.5 and 2. Following the previous numerical experiment, the data sampling is densified in time by a factor 1.5, Fig. 6.4 Left, and by a factor 2, Fig. 6.4 Right; the inverse computations are performed with $N_{FS} = 10$ (corresponding to an identification time interval $\delta t^{ident} = 8$ days).

Like in the previous case and following the identifiability map, the identification is accurate only in the vicinity of the available observations (the vertical colored lines in Fig. 6.4) and is a-priori unaccurate elsewhere because of the large “blind” space time intervalls.

These two experiments show that if the observations are closer to the peak time that the latter is better captured. However the inflow peak is not properly captured in none of these two experiments since no observation is available at the corresponding time.

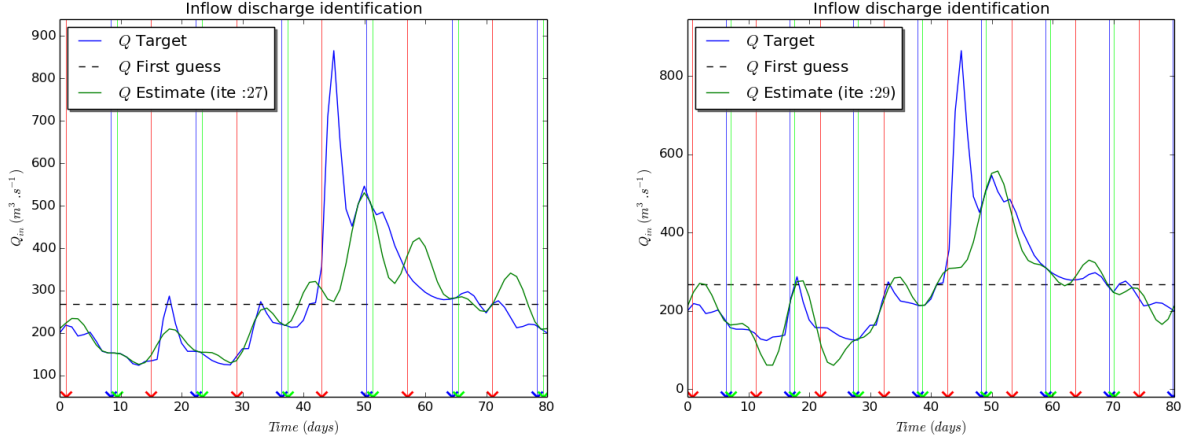


FIGURE 6.4. Garonne river, Scenario 2. Discharge identification with Fourier series, $N_{FS} = 10$ and: (Left) $T_{obs} = T_{obs}^{ori}/1.5$. (Right) $T_{obs} = T_{obs}^{ori}/2$. Vertical lines correspond to the time observations (blue for Group 1, red for Group 2 and green for Group 3). The horizontal dotted line corresponds to the first guess.

6.4. Scenario 3: densified SWOT temporal sampling by a factor 100. In this last test case, the data sampling and the hydrograph are re-scaled (densified) in time by a factor 100. Then the hydrograph remains consistent with the domain length since the peak duration is slightly higher than the response time of the complete domain. As already discussed and as indicated on the identifiability map Fig. 6.1, a majority of the inflow information has time enough to travel throughout the domain and is partially observed. Then many inflow values should be identifiable (although not all).

Before performing the VDA process, the first step is to define empirically an adequate identification frequency (equivalently by setting N_{FS}) from the observation times. In the present case, Fig. 6.5, a good choice for δT^{ident} would be 0.05 day at least and 0.1 day at last. This corresponds to $N_{FS} = 16$ and 8 respectively. The VDA experiments are strictly the same as the previous ones but the time scale; they are performed with N_{FS} equal to 5, 10, 15 and 40, see Fig. 6.5.

To better understand the origin of the identification errors, the approximation of the exact inflow discharge by the Fourier series are plotted, see the 4 curves “Exact FS with $N_{FS}=...$ ” in Fig. 6.5. Also in each figure, the characteristic feet indicated at $t = 0$ as vertical lines in Fig. 6.1 are taken back. This defines the so-called identifiability intervalls.

As expected, the VDA process provides an accurate inflow discharge at the times corresponding to the black identifiability intervalls. On the contrary, the peak is captured partially only since it occurs during a red identifiability intervall; the latter being much less reliable since these identifiability intervalls does not take into account the energy dissipation (the source term in (2.1)). As a result the peak amplitude is partly captured only since not fully observed. However, the identification is globally correct considering the quite sparse observations.

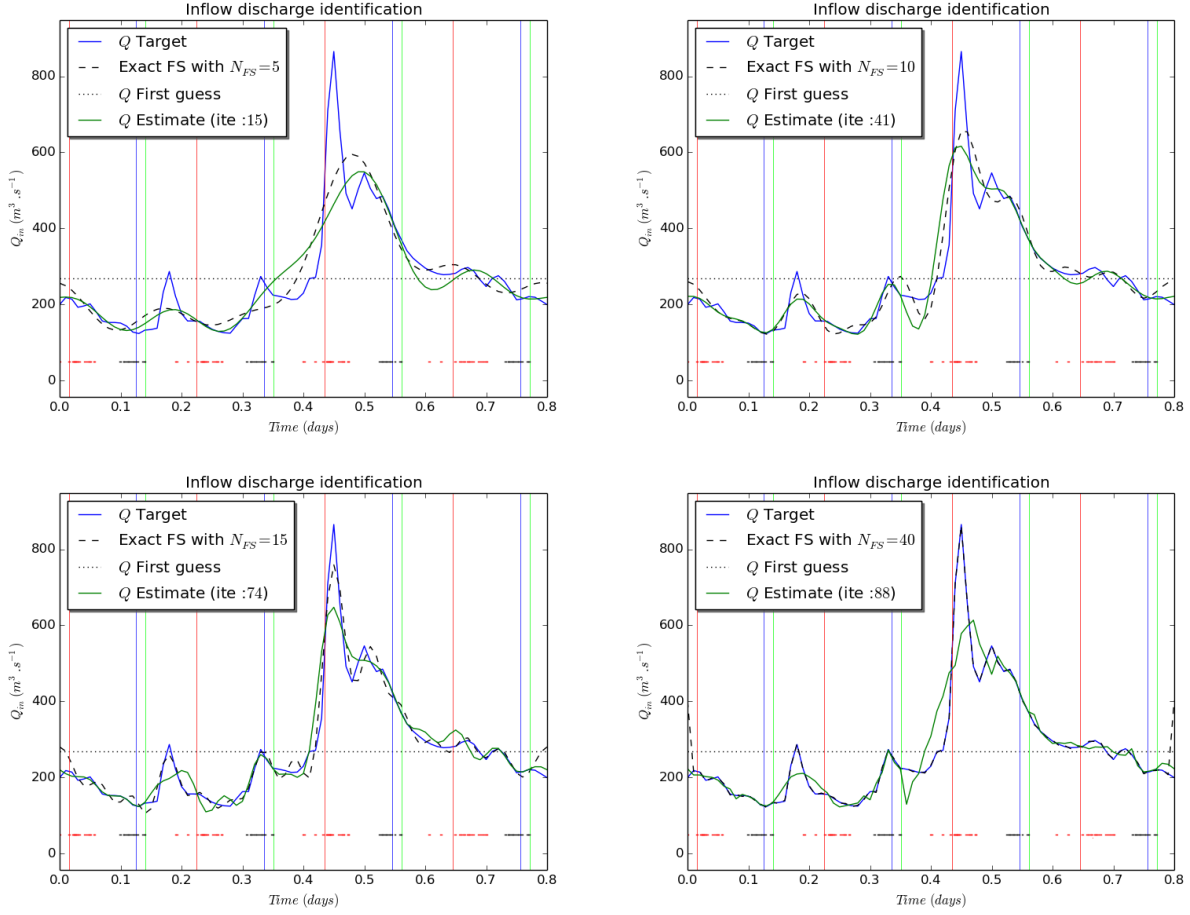


FIGURE 6.5. Garonne river, Scenario 3. Discharge identification with Fourier series with: (Left, top) $N_{FS} = 5$ (Right, top) $N_{FS} = 10$ (Left, bottom) $N_{FS} = 15$. (Right, bottom) $N_{FS} = 40$.

The vertical colored lines correspond to the observation times (blue for Group 1, red for Group 2 and green for Group 3), see Fig. 2.3.

The horizontal colored dashed lines (red and black) at $Q = 50$ correspond to the characteristic feet, lines taken back from the identifiability maps.

Remark 2. It can be noticed that since the wave velocity $\frac{5}{3}\bar{u}$ slower than $(\bar{u} + \bar{c})$, see Fig. 6.6 and 4.1, then the inflow discharge identifiability in the diffusive wave model would be less severe than in the present 1D Saint-Venant model. However, in many hydraulic situations the diffusive wave model is unsatisfying, see e.g. [29] and references therein.

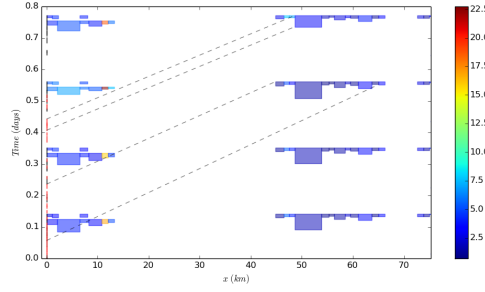


FIGURE 6.6. Identifiability maps in (x, t) in the Garonne river case. Scenario 3 (temporal sampling densified by a factor 100) like Fig. 6.1 but with the velocity waves equal to $\frac{5}{3}u$ (diffusive wave model). The vertical lines at upstream represent the sets of points which can be identified in the model without source term (in red those coming from far, in black coming from close observed reaches).

7. CONCLUSION

The identifiability of inflow discharge, and potentially roughness coefficients, have been investigated in the context of SWOT like data (altimetric data of the river surface). Before performing the Variational Data Assimilation (VDA) processes, inverse problem identifiability maps (aiming at connecting the observations and the model waves in the (x, t) plane) have been analyzed. These maps provide a comprehensive overview of the inverse problem and may give some information in terms of reliability of the VDA results; also this indicates how to define an adequate identification frequency. Typically, in the SWOT case, the identifiability map suggests that the sampling sparsity prevents to infer accurately a complete hydrograph, between sparse observations times, in the Saint-Venant model. Also, these preliminary maps make possible to recover the space-time windows where the inference obtained by VDA should be reliable. The numerous VDA experiments which have been performed (both on academic test cases and on a 80 km portion of the Garonne river) have confirmed these identifiability map analyses. The numerical results have been analysed for various observation sampling densities. Sensitivity analyses with respect to the observation errors and with respect to the first guesses values demonstrate the good robustness of the VDA inferences. Surprisingly if inferring the roughness coefficients (defining a power law $K(h)$) in addition to the inflow discharge $Q_{in}(t)$, then the inference robustness and accuracy remain nearly the same (compared to the discharge inference only). The present precise study related to the only SWOT like data and the Saint-Venant model (1D shallow water model for river flows) demonstrate on one hand the great contribution of these forthcoming measurements, on the other hand their limitations for 1D river flow dynamic models. The present study completes the previous analyses led on this topic and constitute an important stage before addressing the identifiability and inferences in the flow model by VDA but from multi-sensor and complimentary data.

The present identifiability map concept should be investigated more into details in the context of multi-source data (e.g. multi-satellites or combination of in situ data and satellite data), and potentially for the more simple diffusive wave model. In other respect, these identifiability maps may be employed to define the spatio-temporal data assimilation windows, then leading to better constraint of the flow model hence more accurate inflow discharge identifications. Finally similar numerical experiments should be performed for longer time simulations, typically complete multi-year hydrographs, from an already quite accurate first guess. (Good first guesses are generally obtained from simple models like the Manning-Strickler equation, see e.g. [40, 10]).

Acknowledgements. The first author has been funded by CNES TOSCA grant (related to the forthcoming NASA-CNES SWOT mission) during his two years contract of engineer CNES (based at IMT-INSAToulouse). The other authors has been partly supported by a CNES TOSCA grant.

8. APPENDICES

8.1. River model geometry . The resolution of the Saint-Venant equation (1D shallow water) (2.1) requires the computation of wet surface S and perimeter Pe in function of water depth h and geometrical parameters. Then sequences of wet surface $(S_i)_{0 \leq i \leq I}$ and perimeter $(Pe_i)_{0 \leq i \leq I}$ are introduced with $I \in 0, \dots, N_p$ where N_p is the maximal number of triplets $(H_i, w_i, Y_i)_{0 \leq i \leq N_p}$.

For the notations, the reader should refer to Fig. 8.1.

- The wet surfaces $(S_i)_{0 \leq i \leq I}$ are defined by:

$$\begin{cases} S_0 = (H_0 - z_b)w_0 \\ S_i = \frac{1}{2} (w_{i-1} + w_i) (H_i - H_{i-1}) \quad \forall i \in \llbracket 1, N_p \rrbracket \end{cases}$$

- The wet perimeters $(Pe_i)_{0 \leq i \leq I}$ are defined by:

$$\begin{cases} Pe_0 = w_0 + 2(H_0 - z_b) \\ Pe_i = \underbrace{\left(\left(\frac{W_i}{2} - \left(\frac{W_{i-1}}{2} - y_i \right) \right)^2 + (H_i - H_{i-1})^2 \right)^{1/2}}_{=Pe_{1i}} + \underbrace{\left(\left(\frac{W_i}{2} - \left(\frac{W_{i-1}}{2} + y_i \right) \right)^2 + (H_i - H_{i-1})^2 \right)^{1/2}}_{=Pe_{2i}} \end{cases} \quad \forall i \in \llbracket 1, N_p \rrbracket$$

with $y_i = Y_{i-1} - Y_i$, $i \in \llbracket 1, N_p \rrbracket$.

Let $m \in \mathbb{N}$ such that: $H_m < h < H_{m+1}$; or equivalently, $\sum_{i=1}^m S_i < S < \sum_{i=1}^{m+1} S_i$.

Thanks to the sequences $(S_i)_{0 \leq i \leq I}$ and $(Pe_i)_{0 \leq i \leq I}$, it is possible to define the following geometric functions:

- Function $Pe(h)$:

$$Pe(h) = \begin{cases} 0 & \text{if } h = 0 \\ 2h + w_0 & \text{if } 0 < h \leq H_0 - z_b \\ (2h + w_0) + \sum_{i=1}^m Pe_i + Pe'_m & \text{if } h > H_0 - z_b \end{cases}$$

with:

$$Pe'_m = \left(\left(\frac{W_{m+1}}{2} - \left(\frac{W_m}{2} - y_{m+1} \right) \right)^2 + (H_{m+1} - H_m)^2 \right)^{1/2} \left(\frac{(h+z_b)-H_m}{H_{m+1}-H_m} \right) + \left(\left(\frac{W_{m+1}}{2} - \left(\frac{W_m}{2} + y_{m+1} \right) \right)^2 + (H_{m+1} - H_m)^2 \right)^{1/2} \left(\frac{(h+z_b)-H_m}{H_{m+1}-H_m} \right)$$

- Function $S(h)$:

$$S(h) = \begin{cases} 0 & \text{if } h = 0 \\ hw_0 & \text{if } 0 < h \leq H_0 - z_b \\ \sum_{k=0}^m s_k + s'_m & \text{if } h > H_0 - z_b \end{cases}$$

with:

$$s'_m = \frac{1}{2} \left(2w_m + \left((Pe_{1(m+1)}^2 - (H_{m+1} - H_m)^2)^{\frac{1}{2}} + (Pe_{2(m+1)}^2 - (H_{m+1} - H_m)^2)^{\frac{1}{2}} \right) \left(\frac{(h+z_b)-H_m}{H_{m+1}-H_m} \right) \right) ((h+z_b)-H_m)$$

- Function $h(S)$:

$$h(S) = \begin{cases} 0 & \text{if } S = 0 \\ \frac{S}{w_0} & \text{if } S \leq s_0 \\ H_m - z_b + h'_m & \text{if } S > s_0 \end{cases}$$

with:

$$h'_m = - \left(\frac{w_m - X}{w_{m+1} - w_m} \right) (H_{m+1} - H_m), \text{ where } X = \sqrt{w_m^2 + 2 \left(\frac{w_{m+1} - w_m}{H_{m+1} - H_m} \right) (S - s_m)}$$

If m is such that $w_{m+1} = w_m$, so the relation is simplify by: $h'_m = \frac{(s-s_m)}{w_m}$

- Function $w(h)$:

$$w(h) = \begin{cases} 0 & \text{if } h = 0 \\ w_0 & \text{if } 0 < h \leq H_0 - z_b \\ w_m + \alpha_{1(m+1)}((h + z_b) - H_m) + \alpha_{2(m+1)}((h + z_b) - H_m) & \text{if } h > H_0 - z_b \end{cases}$$

With α_{1i} and α_{2i} the slope of trapezium i so:

$$\alpha_{1i}, \alpha_{2i} = \frac{\left(\frac{w_i}{2} - \left(\frac{w_{i-1}}{2} \pm y_i \right) \right)}{H_i - H_{i-1}}$$

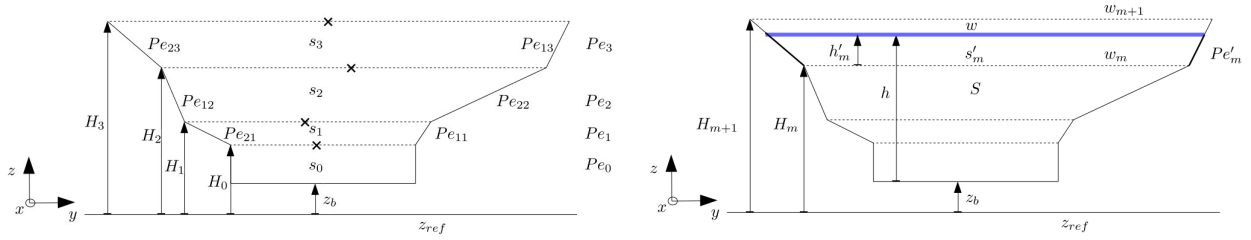


FIGURE 8.1. Notation used for the geometric functions.

8.2. Finite volume scheme. The Saint-Venant equation (1D shallow water) (2.1) are computationally solved by the following first order finite volume scheme.

8.2.1. *First order scheme.* Eqn (2.1) are rewritten in conservative form as follows:

$$(8.1) \quad \begin{cases} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = 0 & (8.1.1) \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} + P \right) = g \int_0^h (h - z) \frac{\partial \tilde{w}}{\partial x} dz - gS \frac{\partial z_b}{\partial x} - gSS_f & (8.1.2) \end{cases}$$

where P is a “pressure term” as proposed by [38], see also e.g [15]. It is defined by:

$$(8.2) \quad P(x, \tilde{S}, t) = g \int_0^{h(x,t)} (h(x, t) - z) \tilde{w}(x, z, t) dz$$

Then (2.1) is re-written as follows:

$$(8.3) \quad \frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = \mathbf{S}(\mathbf{U})$$

with $\mathbf{U} = \begin{bmatrix} S \\ Q \end{bmatrix}$, $\mathbf{F}(\mathbf{U}) = \begin{bmatrix} Q \\ \frac{Q^2}{S} + P \end{bmatrix}$, $\mathbf{S}(\mathbf{U}) = \begin{bmatrix} 0 \\ g \int_0^h (h - z) \frac{\partial \tilde{w}}{\partial x} dz - gS \frac{\partial z_b}{\partial x} - gSS_f \end{bmatrix}$ and $P = g \int_0^h (h - z) \tilde{w} dz$

The Jacobian matrix of \mathbf{F} reads:

$$\mathbf{J}_{\mathbf{F}} = \begin{bmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{bmatrix} \text{ since } c = \sqrt{\frac{\partial P}{\partial S}} = gh \text{ and } u = \frac{Q}{S}$$

The eigenvalues of \mathbf{J}_F are: $\lambda_1 = u + c$ and $\lambda_2 = u - c$; their associated eigenvectors are: $r_1 = (1, u + c)^T$ and $r_2 = (1, u - c)^T$.

To solve the homogeneous form of (8.3), the classical scheme based on the Euler time scheme is used:

$$(8.4) \quad \mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \Delta t^n \frac{\mathbf{F}_{i+1/2}^n - \mathbf{F}_{i-1/2}^n}{\Delta x_i}$$

The numerical flux \mathbf{F}_j^n are computed by the standard HLL formula, such as derived in [16], see also e.g. [37] and references therein.

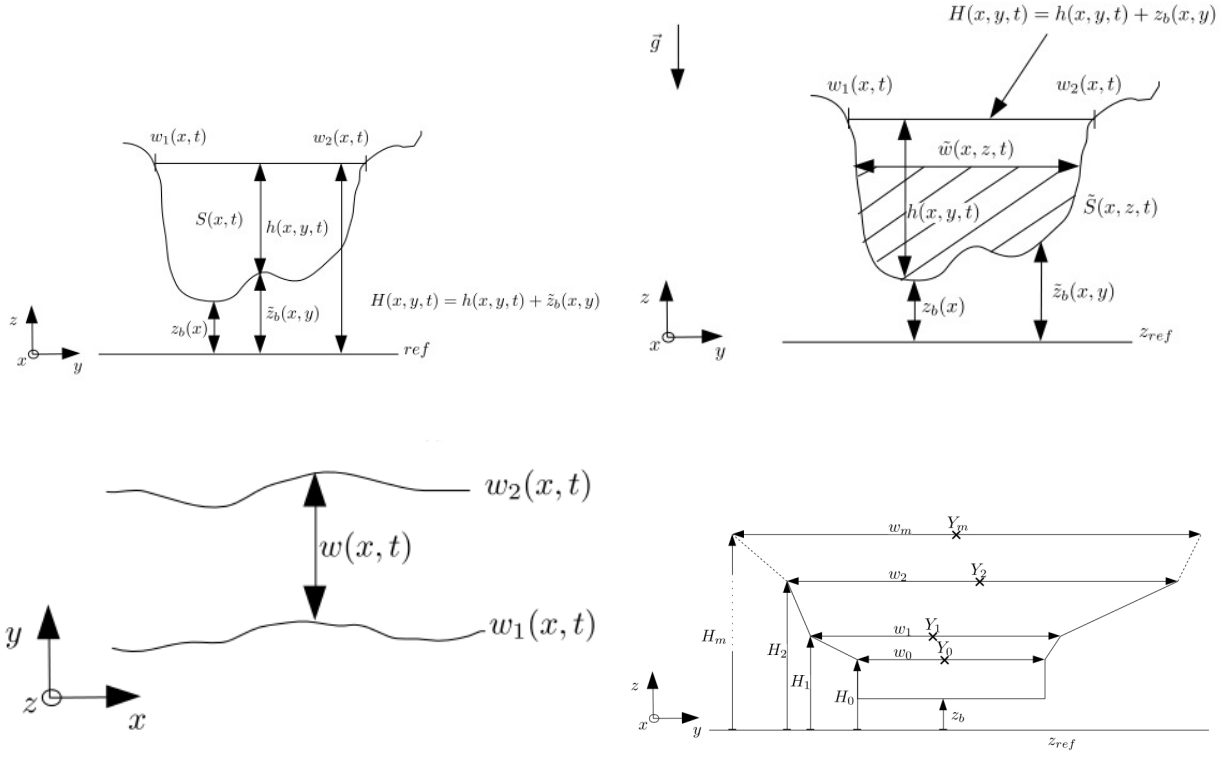


FIGURE 8.2. Notations. (Left, top) Notations for the river cross sections in (yz) -view. (Right, top) Variational notations for the river cross sections in (yz) -view. (Left, bottom) Notations for the river cross sections in (xy) -view. (Right, bottom) Effective geometry considered for each cross section: superimposition of m trapeziums. For the Garonne river case, $m = 150$ (yz -view).

8.2.2. Pressure term discretization. The pressure term $P = g \int_0^h (h-z) \tilde{w} dz$ has to be correctly discretized to obtain a correct convergence of the HLL scheme. Thanks to the particular geometry, it is possible to compute the pressure term piecewise. This computational step is CPU time consuming if the number of trapezium is high (recall 150 for the For the Garonne river case).

Let P_i^n be the discrete pressure term with i the cross section number; let j be the trapezium layer number. Let us denotes: $h_i^n \equiv h$, $H_{j,i} \equiv H_j$, $z_{bi} \equiv z_b$, $\alpha_{1j,i} \equiv \alpha_{1i}$, $\alpha_{2j,i} \equiv \alpha_{2i}$, $w_{j,i} \equiv w_j$ and $h_j = (H_j - z_b)$ with $h_{-1} = 0$. Then,

- If $h(x, t) = 0$, $P_i^n = 0$.

- If $0 < h(x, t) \leq H_0$,

$$P_i^n = \frac{1}{2} g w_{0,i} (h_i^n)^2$$

- Else,

$$\begin{aligned} P_i^n = & g \sum_{j=0}^m \left(w_{j-1} \left(h(h_j - h_{j-1}) + \frac{h_{j-1}^2 - h_j^2}{2} \right) + g w_m \left(h(h - h_m) + \frac{h_m^2 - h^2}{2} \right) \right. \\ & + g(\alpha_{1j} + \alpha_{2j}) \left(\frac{h_{j-1}^3 - h_j^3}{3} + (h + h_{j-1}) \frac{h_j^2 - h_{j-1}^2}{2} + h h_{j-1} (h_{j-1} - h_j) \right) \\ & \left. + g(\alpha_{1(m+1)} + \alpha_{2(m+1)}) \left(\frac{h_m^3 - h^3}{3} + (h + h_m) \frac{h^2 - h_m^2}{2} + h h_m (h_m - h) \right) \right) \end{aligned}$$

8.2.3. Source term discretization. In order to solve the non-homogeneous problem (8.3), a classical splitting method is used, see e.g. [37].

Let us denote $\bar{\mathbf{U}}_i^{n+1} = [\bar{S}_i^{n+1}, \bar{Q}_i^{n+1}]^T$ the solution of the homogeneous problem (8.3) at point x_i and time t^{n+1} ; let us denote $\mathbf{U}_i^{n+1} = [S_i^{n+1}, Q_i^{n+1}]^T$ the solution of the non-homogeneous problem at x_i and t^{n+1} . Then the complete numerical scheme to solve (8.3) reads:

$$(8.5) \quad \begin{cases} \bar{\mathbf{U}}_i^{n+1} = \mathbf{U}_i^n - \Delta t^n \frac{\mathbf{F}_{i+1/2}^n - \mathbf{F}_{i-1/2}^n}{\Delta x_i} \\ \mathbf{U}_i^{n+1} = \bar{\mathbf{U}}_i^{n+1} + \Delta t^n \mathbf{S}(\bar{\mathbf{U}}_i^{n+1}) \end{cases}$$

REFERENCES

- [1] Dassflow software: Data Assimilation Computation for Free Surface Flows. <http://www.math.univ-toulouse.fr/DassFlow/index.html>. CNES/CNRS/IMT/ICUBE/INSA Group/University of Toulouse.
- [2] K. Andreadis and G. Schumann. Estimating the impact of satellite observations on the predictability of large-scale hydraulic models. *Advances in Water Resources*, 73:44–54, 2014.
- [3] E. Bélanger and A. Vincent. Data assimilation (4d-var) to forecast flood in shallow-waters with sediment erosion. *Journal of Hydrology*, 300(174):114 – 125, 2005.
- [4] S. Biancamaria. *Etude du cycle hydrologique des régions boréales et apport de l’altimétrie à large fauchée*. Phd thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier, 2009.
- [5] S. Biancamaria, K. M. Andreadis, M. Durand, E. A. Clark, E. Rodriguez, N. M. Mognard, D. E. Alsdorf, D. P. Lettenmaier, and Y. Oudin. Preliminary characterization of SWOT hydrology error budget and global capabilities. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 3:6–19, 2010. 1.
- [6] P. Brisset, P-A. Garambois, and J. Monnier. Dassflow shallow 1D: Equations and numerical schemes. Technical report, Internal research report. CNES, INSA group. <http://www.math.univ-toulouse.fr/DassFlow>, 2016.
- [7] M. Carlier. *Hydraulique générale et appliquée*. Eyrolles, Paris, France, 1982.
- [8] V.T. Chow. *Open-channel Hydraulics*. Mc Graw-Hill, New-York, USA, 1959.
- [9] V.T. Chow. *Handbook of applied hydrology*. McGraw-Hill Book Co., New-York, 1467 pages, 1964.
- [10] M. Durand, C.J. Gleason, P-A. Garambois, D. Bjerkli, L.C. Smith, H. Roux, E. Rodriguez, P.D. Bates, T.M. Pavelsky, J Monnier, et al. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resources Research*, 2016.
- [11] P-A. Garambois and J. Monnier. Inference of effective river properties from remotely sensed observations of water surface. *Advances in Water Resources*, 79:103–120, 2015.
- [12] A. F. Gessese, M. Sellier, E. Van Houten, and G. Smart. Reconstruction of river bed topography from free surface data using a direct numerical approach in one-dimensional shallow water flow. *Inverse Problems*, 27(2):025001, 2011.
- [13] A. F. Gessese, K. Wa, and M. Sellier. Bathymetry reconstruction based on the zero-inertia shallow water approximation. *Theoretical and Computational Fluid Dynamics*, 27(5):721–732, 2013.
- [14] J. C. Gilbert and C. Lemaréchal. Some numerical experiments with variable-storage quasi-newton algorithms. *Mathematical programming*, 45(1-3):407–435, 1989.
- [15] N. Goutal and F. Maurel. A finite volume solver for 1D shallow-water equations applied to an actual river. *International Journal for Numerical Methods in Fluids*, 38(1):1–19, 2002.
- [16] A. Harten, P.D. Lax, and B. Van Leer. On upstream differencing and godunov-type scheme for hyperbolic conservation laws. *SIAM Review*, 25(1):35–61, 1983.
- [17] L. Hascoët and V. Pascual. *Tapenade 2.1 user’s guide*. PhD thesis, INRIA, 2004.
- [18] L. Hascoët and V. Pascual. The Tapenade Automatic Differentiation tool: Principles, Model, and Specification. *ACM Transactions On Mathematical Software*, 39(3), 2013.
- [19] M. Honnorat. *Assimilation de données lagrangiennes pour la simulation numérique en hydraulique fluviale*. PhD thesis, INPG, Grenoble, France, 2007.

- [20] M. Honnorat, X. Lai, F-X. le Dimet, and J. Monnier. Variational data assimilation for 2D fluvial hydraulics simulation. *CMWR XVI-Computational Methods for Water Ressources. Copenhagen, june 2006.*, 2006.
- [21] R. Hostache, X. Lai, J Monnier, and C. Puech. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. *Journal of Hydrology*, 390:257–268, 2010. 3-4.
- [22] X. Lai and J. Monnier. Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: mathematical method and test case. *Journal of Hydrology*, 377:1–11, 2009. 1-2.
- [23] K. Larnier. *Modélisation thermohydraulique d'un tronçon de Garonne en lien avec l'habitat piscicole : Approches statistique et déterministe*. PhD thesis, Toulouse, 2010.
- [24] C. Lion. *Simulation des données SWOT haute résolution et applications à l'étude de l'estuaire de l'Amazonie*. PhD thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier, 2012.
- [25] J. Monnier. *Variational data assimilation: from optimal control to large scale data assimilation*. 2014.
- [26] J. Monnier, P. Brisset, F. Couderc, and P-A. Garambois. Dassflow: Data assimilation for free surface flows. Technical report, CNES-CNRS-INSa-University of Toulouse. <http://www.math.univ-toulouse.fr/DassFlow>, 2016.
- [27] J. Monnier, F. Couderc, D. Dartus, K. Larnier, R. Madec, and J-P. Vila. Inverse algorithms for 2D shallow water equations in presence of wet dry fronts. application to flood plain dynamics. *Advances in Water Resources*, 2016.
- [28] V.A. Morozov. The error principle in the solution of operational equations by the regularization method. *USSR Computational Mathematics and Mathematical Physics*, 8(2):63–87, 1968.
- [29] R. Moussa and C. Bocquillon. Approximation zones of the Saint-Venant equations for flood routing with overbank flow. *Hydrology and Earth System Sciences*, 4:251–261, 2000.
- [30] J.E. Nash and J.V. Sutcliffe. River flow forecasting through conceptual models. Part I - A discussion of principles. *J. of Hydrology*, 10:282–290, 1970.
- [31] S. Ricci, A. Piacentini, O. Thual, E. Le Pape, and G. Jonville. Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. *Hydrol. Earth Syst. Sci.*, 15:3555–3575, 2011. 11.
- [32] E. Rodríguez. SWOT Science requirements document. JPL document, JPL, 2012.
- [33] H. Roux and D. Dartus. Parameter identification using optimization techniques in open-channel inverse problems. *J. of Hyd. Res.*, 43:311–320, 2005. 3.
- [34] H. Roux and D. Dartus. Use of parameter optimization to estimate a flood wave: Potential applications to remote sensing of rivers. *J. of Hydrology*, 328:258–266, 2006.
- [35] S. Simeoni-Sauvage. *Modélisation hydrobiogéochimique de la Garonne à l'étiage estival : Cas de l'azote entre Toulouse et Agen (120 kilomètres)*. PhD thesis, Toulouse, INPT, 1999.
- [36] O. Thual. *Hydrodynamique de l'environnement*. Ecole polytechnique, 2010.
- [37] EF. Toro. *Shock-capturing methods for free-surface shallow flows*. Wiley, 2001.
- [38] J-P. Vila. *Théorie et approximation numérique de problèmes hyperboliques non linéaires applications aux équations de Saint-Venant et à la modélisation des avalanches de neige dense*. PhD thesis, Université Paris VI, March 1986.
- [39] Y. Yoon, M. Durand, C.J. Merry, E.A. Clark, K.M. Andreadis, and Alsdorf D.E. Estimating river bathymetry from data assimilation of synthetic swot measurements. *Journal of Hydrology*, 464 - 465(0):363 – 375, 2012.
- [40] Y. Yoon, P-A. Garambois, R. Paiva, M. Durand, H. Roux, and E. Beighley. Improved error estimates of a discharge algorithm for remotely sensed river measurements: Test cases on Sacramento and Garonne Rivers. *Water Resources Research*, 52(1):278–294, 2016.