River discharge and bathymetry estimations from SWOT altimetry measurements

Kévin Larnier, Jerome Monnier, P.-A. Garambois, J. Verley

ARTICLE TEMPLATE

Rivers discharge and bathymetry estimations from SWOT altimetry measurements

K. Larnier[(1)(2)(3)(4)] and J. Monnier[(1)(3)] and P.-A. Garambois[(4)(5)] and J. Verley[(1)(3)(4)]

[(1)] Institut de Mathématiques de Toulouse (IMT), France; [(2)] CS corporation, Business Unit Espace, Toulouse, France; [(3)] INSA Toulouse, France; [(4)] ICUBE, Strasbourg, France; [(5)] INSA Strasbourg, France.

ARTICLE HISTORY
Compiled February 6, 2019

ABSTRACT
An inversion method to estimate the discharge of rivers observed by the forthcoming SWOT mission (wide swath altimetry) is developed and assessed in detail. The method relies on a variational data assimilation formulation and the Saint-Venant equations (1D shallow-water) combined with dedicated algebraic systems. This hierarchical modeling approach enables to estimate from altimetry measurements the three key flow features: the discharge $Q(t)$ associated with an (effective) bathymetry $b(x)$ and a friction coefficient K. Cross sections river geometry are built up from trapezium superimposition deduced from the altimetry measurements. Extensive numerical results are analyzed for three rivers portions presenting rapid flow variations compared to the observation frequency. Two scenarios of observation are considered: frequent satellite overpasses corresponding to the SWOT Cal-Val orbit ($\sim$1 day period) and SWOT like data ($\sim$ 5-21 days period depending on the latitude). It is shown that the space-time variations of the river discharge $Q(t)$ and the bathymetry profile $b(x)$ are accurately infered; however the infered values of Q may be obtained up to a multiplicative factor depending on the first guesses (prior values). This bias vanishes as soon as an accurate mean value or one reference value of one of the three infered field is provided. Various prior information sources are investigated in view of worldwide applications. Once the assimilation of the satellite measurements is done during a complete hydrological cycle, a dedicated low complexity (algebraic) system may be correctly calibrated, next providing accurate discharge estimations in real-time from newly acquired measurements.

KEYWORDS
Hydrology; inference; data assimilation; altimetry; SWOT; discharge; bathymetry.

1. Introduction

While in-situ observability of the continental water cycle and river flows is declining, a myriad of satellites for earth observations provide increasingly accurate measurements. The future Surface Water and Ocean Topography (SWOT) mission (CNES-NASA, planned to be launched in 2021) equipped with a wide swath radar interferometer will provide river surfaces mapping at a global scale with an unprecedented spatial and temporal resolution. Its measurements will be Water Surface (WS) height, width and

Corresponding author: Jerome.Monnier@insa-toulouse.fr

slope, with a decimetric accuracy on WS height averaged over 1 km [1]. SWOT will cover a great majority of the globe with relatively frequent revisits (1 to 4 revisits per 21 days repeat cycle). By complementing decades of nadir altimetry in inland waters [2], SWOT should offer the opportunity to increase our knowledge of the spatial and temporal distribution of hydrological fluxes.

Thanks to this increased observability of WS worldwide, it will be possible to address a variety of inverse problems in surface hydrology and related fields. Given these WS measurements (elevation, water mask extents), the challenging inverse problem(s) consists to infer: the discharge, the bathymetry (unobservable part of the river cross sections), the friction law parametrization and any lateral contributions. The estimation of the discharge is more or less challenging depending on the space-time WS observation density and the prior information quality (also potentially the measurement errors). Recent literature addresses some aspects of these inverse questions in a purely remote sensing context, see e.g. [3] for a review. Relatively basic inverse methods have been developed; they are either based on the algebraic Manning-Strickler's law (see e.g. [4]) or empirical explicit hydraulic geometries power-laws[5–8]. In [9], numerous approaches are compared on 19 rivers with artificially densified daily observables; the results fluctuate depending on the algorithm tested. No approach turned out to be accurate or robust in all configurations. In this study the potential benefit of having a correct a-priori estimation of the bathymetry was highlighted. In the river hydraulics community, the most employed data assimilation studies are based on sequential algorithms, the Kalman Filter and its variants, see e.g. [10,11] based on the 1D Saint-Venant model and e.g. [12] based on the diffusive wave model.

None of these aforementioned studies address the real inverse problem encountered in the satellite context at global scale: the inference of the triplet (discharge $Q(t)$, bathymetry $b(x)$ and friction coefficient $K(x,h)$, h being the water depth).

Variational Data Assimilation (VDA) approaches (optimal control of the dynamic flow models [13–15]) have already been employed to address the present inverse problem. Recall that VDA approaches consists in minimizing a cost function measuring the discrepancy between the model outputs and the observations; somehow combining at best the model, the observations and prior information (fields values, characteristic wave length of variations and fields regularity). In some circumstances (depending on the a-priori available information), it is possible to infer the key unknown "parameters" of the river flow model: the inflow discharge $Q_{in}(t)$, the bathymetry $b(x)$, the friction K and/or forcing terms (e.g. lateral fluxes). Among the pioneer VDA studies dedicated to hydraulic models, let us cite [16–18]; next [19,20] have infered the inflow discharge in 2D shallow water river models. Infering the discharge and complete set of the hydraulic parameters from WS measurements only may lead to "equifinality issues" (ill-posed inverse problems) depending on the flow regime, the adequacy between the observations frequency and the flow dynamics, and the available prior information. The inference of the key triplet (inflow discharge, effective bathymetry and friction coefficient) is investigated in [20,21] but from surface Lagrangian drifting markers providing constraining observations. The upstream, downstream and a few lateral fluxes are identified from water levels measured at in-situ gauging stations (Pearl River, China) in [22]; however again the bathymetry and friction are given. The assimilation of spatially distributed water level observations in a flood plain (a single image acquired by SAR) and a partial in-situ time series (gauging station) are investigated in [23,24], see also [25]. In [26,27] the inference of inflow discharge and lateral fluxes are identified by VDA by superposing a 2D local "zoom model" over the 1D Saint-Venant model; these studies are not conducted in a sparse altimetry

measurement context.

The altimetry measurements of WS are generally sparse compared to the flow dynamics, both in space and time. This important feature of the inverse problem is analyzed in detail in [28] through the original and instructive "identifiability map". The latter is the (x,t)-plane representation of all the available information, that are: the satellite measurements, the model response in terms of wave propagation and the misfit with respect to the equilibrium state.

Concerning the bathymetry inference, it is shown in [7,28,29] that given a single in-situ measurement of bathymetry, the complete bathymetry profile can be reconstructed. [28,30,31] present accurate inferences of $Q(t)$ by a similar VDA process as the present one but based on the 1D Saint-Venant flow model "only" (it does not include a hierarchical modeling strategy as in the present original approach). In the (single) river case considered in [31], the priors are computed from light Gaussian perturbations of the "true" values of K and b and define a (highly controlling) rating curve $Q(Z)$ is applied at downstream. This boundary condition is a strong prior information making converge the minimization process to the correct discharge $Q(t)$ and the bathymetry $b(x)$: the values corresponding to the nearly exact rating curve.

In summary these recently developed VDA algorithms are accurate to infer the inflow discharge $Q_{in}(t)$ but from accurate prior information. The latter may be a controlling downstream flow condition [30,31] or a single depth measurement in the river portion [28].

In view to apply the algorithms to worldwide ungauged rivers, no prior information may be introduced neither in the direct model nor in the inverse method. This was not the case up to now; this is the challenging scenario considered in the present study. To do so, crucial improvements of the aforementioned VDA based inversions are proposed. Firstly the hierarchical modeling strategy based on original dedicated algebraic systems strengthens the robustness of the estimations in particular if a mean value of Q is provided by a database or a large scale hydrological model.

It is formally demonstrated and numerically confirmed that such a mean value of one field (typically Q or b) i.e. a quite rough but accurate information, improves the accuracy of the estimations of all the space-time dependent unknowns $(Q(t), b(x); K(x;h))$.

The complete inversion strategy presented in this article may provide rivers discharge estimations at global scale from the forthcoming SWOT satellite data (NASA-CNES et al. 2021) and considering additional prior information or not. This is the reason why the presented numerical results are relatively extensive. These results are analyzed in detail on three rivers portions, $\sim$ 100 km long each. Each case presents a "low identifiability index" following the definition introduced in [28]: they present a quite high frequency of hydrograph variations compared to the observation frequency; therefore challenging inverse problems.

Two scenarios of observation are considered: 1) a SWOT like Cal-Val orbit with $\sim$1 day period; 2) a SWOT like data with 21 days period (with 1 to 4 passes at mid-latitudes). For each river case and observation scenario, three inverse problems are considered: the most challenging one, ungauged rivers observed by SWOT only; the second one where multi-temporal priors on discharge are available; and finally the third one where in-situ measurements of bathymetry are available. Suitable methods for each case were developed and are presented in this article. All the algorithms and methods developed here are available in the open-source computational software DassFlow [32].

The outline of the article is as follows. Section 2 presents the river flow models: the classical Saint-Venant equations but with the particular cross section shapes and the original low Froude - low complexity systems. These systems are either based on the classical Manning-Strickler's law or on explicit "low Froude bathymetry" expressions; they are useful for different purposes (direct and inverse ones) depending on the scenario. Section 3 presents the advanced VDA method taking into account prior hydraulic scale and error measurement amplitudes. The description of the three test rivers and scenarios (20 numerical experiments in total) are presented in Section 4. The first guess computation is important; it depends on the available prior information and may be estimated by solving the aforementioned low-complexity systems. Section 5 focuses on hydraulic inferences on ungauged rivers. It is formally demonstrated and numerically shown that the space-time variations of the infered discharge $Q_{in}(t)$ are (very) accurate, however up to a potential bias. In Section 6, in-situ measurements are supposed to be available (e.g. from a database), these may be a single bathymetry value or a discharge time series at a gauging station; the estimations are very accurate. Section 7 presents how to compute real-time estimations of discharge using the original algebraic - low complexity system, past the calibration (or "learning") period. A conclusion is proposed in Section 8.

2. River flow hydrodynamic models

The primary river flow model is the classical Saint-Venant equations in $(A, Q)(x, t)$ variables, see e.g. [4,33], with the Manning-Strickler friction coefficient K defined as a power law in water depth h. It is worth to recall that at the current hydrology satellites measurement scale, the river flows present a low Froude number F_r; typically F_r ranges approximatively within $\approx [0.1, 0.3]$. A first consequence is that the imposed downstream condition controls the flow; then in lack of prior information, it is important that this outflow condition is as "transparent" as possible. To do so, a condition related to the normal depth (derived from the equilibrium Manning-Strickler's equation, see e.g. [4,33]) is imposed. A second consequence is the relatively good accuracy of the simple Manning-Strickler's equation to model the flow (see e.g. [6,7,9] for studies in the present SWOT context). However the Manning-Strickler's equation is a scalar algebraic relation highly sensitive to uncertainties; moreover as shown in [7] the actual independent variable of this relation is Q/K and not simply Q. Then below are derived "low complexity (algebraic) systems" which are more robust to the uncertainties of the numerous measurements. These algebraic equations are based on the classical hydraulic assumption "low Froude" hence similar in this sense to the Manning-Strickler's equation, but they natively integrate the numerous WS measurements. The resulting systems are original; they are employed both for direct modeling (real-time estimations past the calibration period) and inverse modeling (in particular the estimation of first guesses in the VDA inverse method). The inversions of these low complexity systems based on the numerous observations are much robust (less sensitive to the errors, uncertainties) than the usual Manning-Strickler scalar equation.

2.1. The 1D Saint-Venant model

The 1D Saint-Venant equations below are extremely classical in river hydraulics. It is depth-integrated equations relying on the long-wave assumption that is the geometrical ratio of the flow $\varepsilon = h^*/L^*$ small, h^* being a characteristic water depth and L^* a

characteristic length scale (shallow water assumption). In their non conservative form in (A, Q) variables, A the wetted-cross section $\left[m^2\right]$, Q the discharge $\left[m^3.s^{-1}\right]$, the equations read as follows, see e.g. [4]:

$$\begin{cases} \partial_t A + \partial_x Q & = 0 \\ \partial_t Q + \partial_x \left(\frac{Q^2}{A}\right) + gA\,\partial_x Z & = -g\,A\,S_f(A, Q; K) \end{cases} \tag{1}$$

where g is the gravity magnitude $\left[m.s^{-2}\right]$, Z is the WS elevation $[m]$, $Z = (b + h)$ where b is the lowest bed level (bathymetry elevation) $[m]$ and h is the water depth $[m]$.

S_f is the classical Manning-Strickler friction term:

$$S_f(A, Q; K) = \frac{|Q|Q}{K^2 A^2 R_h^{4/3}} \tag{2}$$

with K the Strickler friction coefficient $\left[m^{1/3}.s^{-1}\right]$, $R_h = {}^{A}/_{P_h}$ the hydraulic radius $[m]$, P_h the wetted perimeter. It will be noticed in next paragraph that R_h may be approximated by the water depth h for large rivers. The discharge Q is related to the average cross sectional velocity u $\left[m.s^{-1}\right]$: $Q = uA$. The Strickler friction coefficient K is defined as an non usual power law in h:

$$K(h) = \alpha\ h^{\beta} \tag{3}$$

where α and β are two constants. This original way to model the friction turns out to be adequate since simply representing the flow friction in function of h hence somehow in function of the flow regime; it is richer than a constant uniform value as it is often set in the literature from a-priori tables of friction in function of river types for instance, see e.g. [4]. This law may depend on the space variable x too. For a sake of simplicity, here the law is supposed to be uniform.

Similar approaches based on hydraulic geometry or power law resistance equations are developed in the literature for predicting mean flow velocity for example on a wide range of river measurements in [34].

At upstream the discharge $Q_{in}(t)$ is imposed. At downstream the Manning-Strickler equation depending on the unknowns $(A, Q; K)_{out}$ is imposed (it is classically integrated in the Preissmann scheme equations). The initial condition is set as the steady state backwater curve profile: $Z_0(x) = Z(Q_{in}(t_0))$. This 1D Saint-Venant model is discretized using the classical implicit Preissmann scheme.

2.2. River description from SWOT measurements

The forthcoming SWOT measurements will provide spatially distributed measurements of river surface elevation Z and width W with temporal revisits [1]. The measurements will be provided at two different scales, the node scale (200m) and the reach scale (a few kilometers). The signals at node scale will allow to capture the high

frequencies at the cost of imporant uncertainties. On the contrary the signals at reach scale will have less uncertainties but will only allow to capture low frequencies. The WS elevations and width will be calculated similarly at reach and node scale. The WS slopes will only be given at the reach scale, but they can also be computed at the node scale using finite differences on Z). All the methods presented after can be applied on both scales (and can even be extended to other scales that are compatible with the hydraulics, e.g. the cross-section scale for the Saint-Venant model). For the sake of clarity, the methods are hereafter presented at the reach scale only.

We consider R reaches, $r = 1, .., R$, and $(P + 1)$ overpasses, $p = 0, .., P$. The overpasses are ordered by increasing flow height. The case $p = 0$ denotes the lowest water level. The SWOT data set on a river domain is: $\{Z_{r,p}, W_{r,p}\}_{R,P+1}$. The direct model (1) is considered with the specific cross-sectional geometry shape described in Fig. 1. It consists in discrete cross sections formed by asymmetrical trapezium layers $(Z_{r,p}, W_{r,p})$, the center ox each in the cross-flow direction being denoted by $Y_{r,p}$. In the following we assume that the WS slope is positive: $S_{r,p} = -\partial_r Z_{r,p} \geq 0$.

The cross-sectional areas $A_{r,p}$ are defined as follows: $A_{r,p} = A_{r,0} + \delta A_{r,p} = A_{r,0} + \int_{h_0}^{h_p} W_r(h)dh \quad \forall r \, \forall p$. The variations $\delta A_{r,p}$ are approximated by the trapeziums $\delta A_{r,p} \approx \sum_{q=1}^{p} \frac{1}{2}(W_r^q + W_r^{q-1})(h_r^q - h_r^{q-1})$. The lowest cross-sectional areas denoted by $A_{r,0} \, \forall r$ are unobserved; they are key unknowns of the flow model. They are represented by rectangles, see Fig. 1, or any other fixed shape (e.g. a parabola); all the other cross-sectional areas are trapezoidal.

We have: $h_{r,p} \approx R_{r,p}^h \approx \frac{A_{r,p}}{W_{r,0}+2h_{r,p}+W_{r,p}} \quad \forall r$. It is worth pointing out that this approximation has been numerically verified on all considered rivers. Since $W >> h$, it follows the effective depth expression: $h_{r,p} = (A_{r,0} + \delta A_{r,p})(W_{r,0} + W_{r,p})^{-1}$. The flows are supposed to be gradually varied that is: $\partial_x W_{r,p}/W_{r,p} \approx 0$.

The unobserved wetted cross section A_0 can be represented by a rectangle, a trapezium, a parabola or even a triangle, by setting: $A_0 = \alpha_A h_{0,\alpha} W$ with $\alpha_A \in [\frac{1}{2}, 1]$. However it will be shown that the a-priori effective cross-section shape does not influence the low Froude ("low complexity") flow relations, see Section 2.3. We define the hydraulic mean depth h_0 by: $h_0 = \frac{A_0}{W}$; it is the depth corresponding to an effective rectangular cross-section.

2.3. "Low complexity" algebraic models

In this paragraph, so-called "low complexity systems" are derived; this terminology "low complexity" being in comparison with the space-time dependent Saint-Venant model which requires higher CPU time computations. Indeed these low complexity systems are algebraic and they are solvable in real-time (Section 7.2). Their equations are based on classical hydraulic assumptions but they are derived in the particular present context: the assimilation of SWOT altimetry data. They natively integrate the measured fields and they may be differently formulated depending on which field is given or unknown. Note that another formulation of the low complexity system is derived in Section 6.2 too.

To derive these equations, the basic assumptions made on the flows and data are the following.

A1) The altimetry measurements are relatively "large scale" compared to the river flows dynamics. In terms of temporal dynamics, since the satellite cannot make infer dynamics much faster than its own frequency revisit (which is $\approx$ 10 days), at each observation instant p, the flow may be described as a *steady state flow*. In terms of

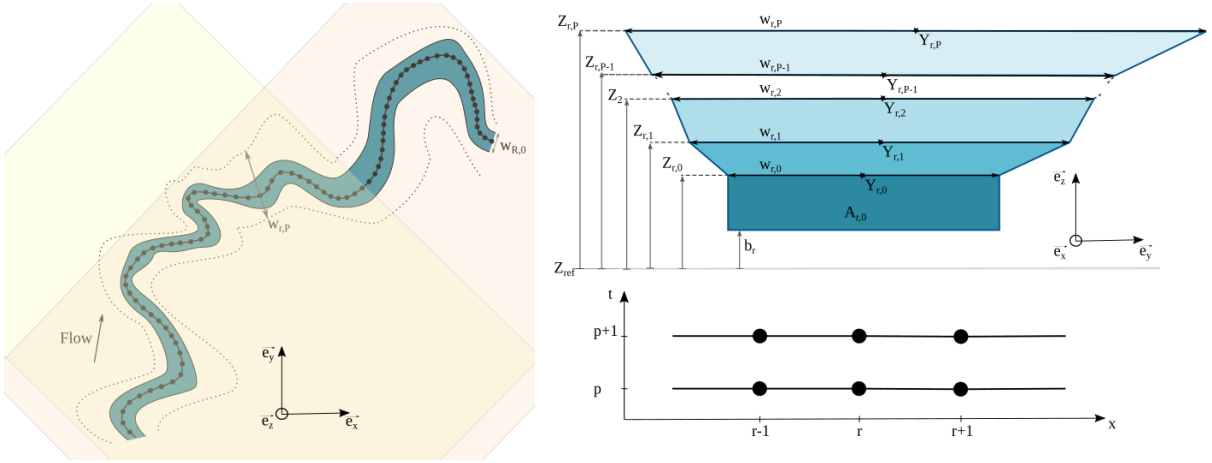

Figure 1. (Left) Schematic river plane view with satellite swaths (large colored rectangles), longitudinal grid with 1D averaged satellite measurements (black dots) along the river centerline (black line). (Right Top) Effective river cross section at reach r defined from SWOT data set $\{Z_{r,p}, W_{r,p}\}_{R,P+1}$. (Right Bottom) Space - time stencil (r, p). x denotes the curvilinear abscissa along the river centerline defined at low flow by $Y_{r,0}$ with $Y_{r,p}$ the middle of the cross sectional width.

spatial variations, the SWOT instrument will provide an average value of the WS measurement at relatively large scale (RiverObs measurements in the SWOT community [35]). Therefore the flow may be described as *locally uniform at the reach length scale*.

A2) At the reach scale defined above, the flow presents low Froude number values; under the *low Froude flow assumption,* the inertial terms in the momentum equations can be neglected.

A3) The considered rivers do not present lateral fluxes.

Moreover the considered rivers are wide enough to assume that the hydraulic radius $R_h \approx h$ in the Manning-Strickler friction term S_f, see (1).

2.3.1. The low Froude flow model: system of local Manning-Strickler's equations

For example for a rectangular cross section, the Froude number Fr satisfies: $Fr^2 = uc_e^{-1}$ with c_e the wave celerity, $c_e = \sqrt{gh}$. Assuming that the flow is permanent and low Froude, $Fr^2 << 1$, the momentum equation simplifies as the Manning-Strickler law: $Q = KAh^{2/3}\sqrt{S}$. Considering this law at each river reach r and each observation time p, this provides the following system:

$$(K_{r,p}^{-1}Q_{r,p})^{3/5} = (A_{0,r} + \delta A_{r,p})W_{r,p}^{-2/5}S_{r,p}^{3/10} \ \forall r, \forall p \tag{4}$$

for each reach $r \in [1...R]$ and each satellite pass $p \in [0...P]$. The terms $\delta A_{r,p}, W_{r,p}, S_{r,p}$ are measured, while the terms $Q_{r,p}, A_{0,r}$ and $K_{r,p}$ are not.

This system (4) can be employed differently depending on the available information and the unknowns. In the inversion computations presented in Section 6.2, (4) is reformulated in (14) to estimate pairs $(K_r^{3/5}, A_{r,0})$ from prior discharge values $Q_{r,p}$ (ancillary data).

2.3.2. The effective low Froude bathymetry

By injecting the expression of Q in (4) (we skip the subscripts $_{r,p}$) into the mass conservation equation ($\partial_x Q = 0$), it follows: $\partial_x \left(KAh^{2/3} S^{1/2} \right) = 0$. Assuming a constant friction coefficient K, this simplifies to: $\frac{\partial_x h}{h}(x) = -\frac{3}{2} \left(\frac{\partial_x (S^{1/2})}{S^{1/2}} + \frac{\partial_x A}{A} \right)(x)$.

Next, following [29] and given a reference depth value h_{ref} (measured at one reference reach ref), the explicit expression of h follows: $h(x) = h_{ref} \left(AS^{1/2} \right)_{ref}^{3/2} \cdot \left(AS^{1/2} \right)^{-3/2}(x)$. This estimation of water depth h is analyzed in detail in [7,29].

Let us extend this estimation of the water depth h (hence the bed elevation b) to the present context, that is from a complete set of WS measurements. If considering an effective cross-sectional area $A = \alpha_A h W$ with $\alpha_A \in [\frac{1}{2}, 1]$ then the bathymetry expression reads: $b(x) = Z(x) - C_{ref} \cdot \alpha_A^{-3/5}(x) \cdot (WS^{1/2})^{-3/5}(x)$ with $C_{ref} = h_{ref} \left(\alpha_A W S^{1/2} \right) |_{ref}^{3/5}$. The three parameters (α_A, W, S) describe the three dimensions of the flow.

Given the WS measurements (Z, W, S) and the value S_{ref} at an arbitrary reach ref, this explicit expression of $b(x)$ provides an effective bathymetry. Observe that the effective bathymetry elevation b is equivalent if considering different wetted cross-section shapes determined by α_A. In other words, in the present low Froude relations, the shape of the effective cross-sectional area A may be chosen arbitrarily. In the following we set $\alpha_A = 1$; that is a rectangular effective shape. Finally the effective bathymetry expression under the assumptions A1)-A3) previously mentioned reads:

$$b_r + (Z_{ref,p} - b_{ref}) \cdot \mathcal{O}_{ref,p}^{-1} \cdot \mathcal{O}_{r,p} = Z_{r,p} \ \forall r, \ r \neq ref, \ \forall p \geq 0 \tag{5}$$

with the observational term $\mathcal{O}_{r,p} = \left(W_{r,p} \sqrt{S_{r,p}} \right)^{-3/5} \ \forall r \, \forall p$. (The only observational terms of this estimation are Z and the product $WS^{1/2} \equiv \mathcal{O}$).

In the following the bathymetry $\{b_r\}_r$ solution of (5) is called the "Low Froude bathymetry".

The system (5) contains $(P+1)\times(R-1)$ equations (given 1 reference reach only) with R unknowns: the bathymetry vector $b = (\{b_r\}_r, b_{ref})$ $r \in [1, R], r \neq ref$. The system reads: $(A.b)_p = (F)_p \, \forall p$, with$(F_r)_p = Z_{r,p} - Z_{ref,p} \mathcal{O}_{ref,p}^{-1} \cdot \mathcal{O}_{r,p}$ and the $(R-1)(P+1)\times R$ matrix $(A)_p = \left[I_{(R-1\times R-1)} | \mathcal{O}_{ref,p}^{-1} \cdot \mathcal{O}_{r,p} \right]$. For $P \geq 1$, A is of maximal rank excepted if the observational vectors $(\mathcal{O}_r)_p$ are linearly dependent; that is not the case if the considered water flow lines represent flow variations.

On the coupled influence between b and a non constant friction coefficient K. As shown above and already demonstrated in [7,29], the low Froude assumption enables to separate the bathymetry effect from the friction effect *if* the friction coefficient K is constant, see (5); K does not appear anymore in (5). On the contrary, if K varies in space or depends on the depth h, like in (3) or in the Einstein formula, see e.g. [4], then K and b have a coupled influence even at low Froude. Typically if considering the following power-law: $K(h) = K_0 \cdot (h - h_0)^{\beta}(x)$ with gradually varied coefficients i.e. $\partial_x K_0 \sim 0 \sim \partial_x \beta$, after calculations the depth expression reads: $[h(x) \cdot |h(x) - h_0|^{3\beta/2}] = c_0 \cdot (AS^{1/2})^{-3/2}(x)$ with $c_0 = |h_{ref} - h_0|^{3\beta/2} \cdot h_{ref} \cdot (AS^{1/2})_{ref}^{+3/2}$ (h_{ref} the reference depth value). The case K constant is recovered from the equation above by setting $\beta = 0$.

Relationship with empirical laws. It may be practical to consider empirical laws based on hydraulic geometry such as: $Z_{r,p} = b_r + a_r W_{r,p}^{\beta}$. In other respects, (5) re-writes as follows: $Z_{r,p} = b_r + C_{ref,p} \cdot \alpha_{A,r}^{-3/5} \cdot (WS^{1/2})|_{r,p}^{-3/5} \ \forall r, \ \forall p, \ p = 0,..,P$ with $C_{ref,p} = h_{ref,p} \left(\alpha_A W S^{1/2}\right)_{ref,p}^{3/5}$. We assume constant cross section shapes *($\alpha_{A,r} = \alpha_A \ \forall r$).* By equaling the two estimations it follows: $W_{r,p}^{\beta} \cdot (W_{r,p} S_{r,p}^{1/2})^{3/5} = a_r^{-1} \cdot h_{ref,p} \left(WS^{1/2}\right)_{ref,p}^{3/5} \ \forall r \ \forall p$. Assuming that the flow is uniform in terms of the observational term $(WS^{1/2})$ i.e. this quantity is constant for the considered reaches, we obtain that: $W_p^{\beta} = a^{-1} \ h_p$; equivalently: $Z_p = b + a \ W_p^{\beta}$.Therefore the low Froude estimation (5) contains empirical laws of the form indicated above. Given β and time series of WS elevation and width this relation allows to infer an effective river bed elevation e.g. as in [5]. Such law could be defined in function of hydraulic geometry knowledge.

On the accuracy of these low Froude - low complexity systems. The two systems (4) and (5) turn out to be reasonably accurate in the present altimetry context (Assumptions A1) -A3)). Employed as direct modeling, their solutions have been numerically assessed in [7]; it is done in the present context too, see Section 6.

3. The Variational Data Assimilation (VDA) formulation

VDA method principles. Given the WS measurements, the VDA method aims at estimating the "input parameters" of the Saint-Venant flow model that is: the inflow discharge $Q_{in}(t)$ of the Saint-Venant model, the bathymetry $b(x)$ and the friction coefficient $K(h)$ defined by (3). In discrete form, this unknown "parameter" reads:

$$c = (Q_{in,0}, \ ..., \ Q_{in,P}; \ b_1, \ ..., \ b_R; \ \alpha, \ \beta)^T \tag{6}$$

Here the subscript p denotes the instant, $p \in [0..P]$, r denotes the node number, $r \in [1..R]$, see Fig. 1 and α and β are the law parameters defined by (3). When using the VDA method with the full Saint-Venant model, it is recommended to consider the node scale as it allows to capture the high frequencies in the observations; this is what is done in the present study.

Since the relation between the elevation Z and the cross-sectional area A (see the river description section 2.2) defines a bijection function, measuring Z is equivalent to measuring A. Of course the parameters used for imposing a normal depth at downstream, see Section 2.1, are considered as unknown otherwise it would be equivalent to impose an exact condition highly controlling the solution.

The cost function and optimization problem. The considered cost function j is defined by:

$$j(c) = j_{obs}(c) + \gamma \, j_{reg}(c) \tag{7}$$

The term $j_{obs}(c)$ measures the misfit between the observations and the model output:

$$j_{obs}(c) = \frac{1}{2} \left\| (Z(\mathrm{c}) - Z_{obs}) \right\|_N^2 \tag{8}$$

The norm N is defined from the a-priori covariance operator N (a positive definite matrix): $\|\cdot\|_N = \|N^{1/2}\cdot\|_2$. The regularization term $j_{reg}(c)$ is detailed below; the weighting coefficient $\gamma > 0$ is set following an iterative regularization strategy detailed below.

The WS elevation Z depends on c through the flow model (1). The inverse problem reads as: $c^* = \text{argmin j}(c)$.

This minimization problem (actually an optimal control problem of (1) is numerically solved by a Quasi-Newton descent algorithm (here the classical L-BFGS algorithm presented in [36]). This first order method requires the computation of the cost gradient $\nabla j(c)$. The gradient is computed by introducing the adjoint model (enabling to consider large control vector dimensions). The adjoint code is obtained by employing the automatic differentiation tool Tapenade [37]. We refer to the pioneer studies [14,15,38,39] for VDA concepts (and e.g. [38,39] for online detailed courses containing know-hows including implementation ones).

The unknown parameter c contains three variables of different physical nature which are space and/or time dependent. Moreover the bathymetry $b(x)$ and the friction coefficient $K(h)$ are correlated and they may have a similar influence in terms of WS signature therefore leading to an ill-posed inverse problem, see e.g. [7] for such a discussion in the present inversion context. Then the inverse problem is regularized in two ways following relatively classical techniques.

Firstly, the regularization term j_{reg} , see (7) is simply set as: $j_{reg}(c) = \frac{1}{2}\|b"(x)\|_2^2$. j_{reg} imposes a smoothing effect on the infered bathymetry profiles $b(x)$). Secondly the following metrics based on (classical) covariance operators are introduced.

Covariance operators and change of control variable. The following natural change of variable is made, see e.g. [40]:

$$k = B^{-1/2}(\text{c} - \text{c}_{prior}) \tag{9}$$

where B is a covariance matrix. Recall that the unknown parameter (the control variable) c is defined by (6); c_{prior} is a prior value (also called "background" or "first-guess" value). The value of c_{prior} depends on the prior information.

The choice of B may be viewed as an important prior information too since the optimal solution k^* (strongly) depends on B. Indeed, after this change of variable, the optimality condition reads: $B^{1/2}\nabla j(c) = 0$. This change of variable may be viewed as a preconditioning method, see e.g. [41,42] for detailed analysis in a different context. Then the optimization problem to be solved re-reads as:

$$\min_k \text{J}(k) \tag{10}$$

with $\text{J}(k) = j(c)$, j defined by (7) and the control vector k defined by (9). The unknown parameter k contains the three variables $Q_{in}(t)$, $b(x)$, $K(h)$ in their discrete form. These three variables are assumed to be uncorrelated: B is thus defined as a block diagonal matrix. We set:

$$B = \begin{pmatrix} B_Q & 0 & 0 \\ 0 & B_b & 0 \\ 0 & 0 & B_K \end{pmatrix} \tag{11}$$

Each block $B_\square$ is defined as a covariance matrix (positive definite matrix). The matrices B_Q and B_b are set as classical second order auto-regressive correlation matrices, see e.g. [40,42,43]:

$$(B_Q)_{i,j} = (\sigma_Q)^2 \exp\left(-\frac{|t_j - t_i|}{\Delta t_Q}\right) \text{ and } (B_b)_{i,j} = (\sigma_b)^2 \exp\left(-\frac{|x_j - x_i|}{L_b}\right) \tag{12}$$

The parameters Δt_Q and L_b act as correlation wave lengths. Given the observation frequency (1 day minimum), given the measurements accuracy (200m long "observation pixels") and given the typical Froude number of the observed river flows, adequate values for these parameters are: $\Delta t_Q = 24$ h and $L_b = 1$ km. We refer to [28] for a thorough analysis of the discharge inference in terms of frequencies and wave lengths. The matrix B_K is diagonal; it may be (roughly) set as: $B_K = diag(\sigma_\alpha^2, \sigma_\beta^2)$. The scalar values $\sigma_\square$ may be viewed as variances.

This VDA formulation above takes into account prior hydraulic scales through the parameters Δt_Q and L_b; the balance between the different control variables are set through the parameters $\sigma_\square$. Their values are detailed in the numerical results sections (Sections 5 and 6).

On the non over-fitting of data. Let us denote by δ the noise level such that for all locations $\|Z_{obs} - Z_{true}\|^2 \leq \delta$ with Z_{obs} the observed and Z_{true} the true WS elevation. Following the Morozov discrepancy principle, see e.g. [49] and references therein, the regularization parameter γ in (7) is chosen *a-posteriori* such that j does not decrease below the noise level. In the present numerical experiments, the convergence is stopped if $j_{obs}(c) \leq \chi\delta$ with $\chi \sim 0.9$.

4. Test rivers and data description

In this section the three test rivers and the different scenarios (depending on the availability of data) are presented. The developed methods to compute the first guesses (estimations used as the first value in the iterative VDA process) are detailed; a summary of all numerical experiments is presented.

4.1. River cases and Water Surface (WS) measurements

The inversion method capabilities are assessed on three rivers with two densities of WS observations representing increasing difficulty. They consist in 75 km of the Garonne River (France), 98 km of the Po River (Italy) and 147 km of the Sacramento River (California, USA), see Tab. 1. These three rivers have contrasting morphological characteristics in terms of the average variability of the slope and the shape of the cross section, resulting in hydrodynamic nonlinearities observable in WS signatures. The impact of WS measurements sampling on hydraulic controls identifiability is assessed with a SWOT Cal/Val scenario (daily observations) on the Po and Garonne Rivers,

and with the nominal SWOT spatio-temporal sampling on the Sacramento River (i.e. relatively unfrequent observations).

The SWOT-like WS observations used in the experiments consist in sets of $(Z_{r,p}, W_{r,p})_{R,P+1}$ either computed using synthetic errors or using the SWOT Hydrology Simulator. SWOT Hydrology Simulator outputs were only available for the Sacramento river for this study. For the two other rivers, synthetic SWOT-like data have been simulated by adding Gaussian errors to the outputs or the reference hydraulic models: $Z_{obs}(x,t) = Z_{true}(x,t) + \mathcal{N}(0, \sigma_Z)$ with $\sigma_Z = 25$ cm corresponding to the expected magnitude of SWOT measurement errors at the present spatial scale - according to the SWOT scientific requirements [1]. The observation period $\Delta t_{obs} = 1$ day for these two rivers corresponds to the to the calibration/validation (Cal/Val) phase of the mission.

For the Sacramento case, the observability corresponding to a nominal 21 days SWOT cycle is simulated by the SWOT Hydrology Simulator. The whole river portion is observed by SWOT (tracks number #249 and #527 respectively) at the 9th and 19th days of each 21 days repeat cycle; the simulated measurement error on Z is characterized by an average variance $\sigma_Z = 34$ cm, see Tab. 1. It is interesting to use this high (simulated) measurement error, in case the real SWOT errors would be locally higher than the scientific requirements. For the three cases the spatial sampling is $\Delta x_{obs} = 200$ m; the SWOT Hydrology Simulator outputs are averaged in space on ~ 200 m nodes (corresponding to the so-called RiverObs nodes).

Table 1. Hydraulic characteristics for each case

Case Name	Reach Length (km)	Max. Width (m)	Avg. Slope (m/km)	Avg. Flow (m3/s)	Froude Range (-)
Garonne	75	49/1,383	0.861	156	0.03-0.67
Po	98	116/5,515	0.145	1499	0.04-0.47
Sacramento	147	59/678	0.558	251	0.02-0.64

Table 2. Observations sources and characteristics for each case

Case Name	Hydrodynamic Model	Temporal Window	Observations Frequency
Garonne 44	HEC-RAS	06 jan. 2010 - 06 avr. 2010	1 day
Po 45	HEC-RAS	01 jan. 2002 - 01 avr. 2002	1 day
Sacramento	HEC-RAS + SWOT HR	01 jan. 2009 - 27 jun. 2009	21 days

4.2. Classes of inverse problems for ungauged and poorly gauged rivers

The future SWOT mission will observe worldwide rivers wider than 100 m with on average of 1 to 4 temporal revisits every 21 days at mid-latitudes, see e.g. [3]. The difficulty of the inverse problem depends on the availability of ancillary data. The inversion capabilities of the present inverse method(s) are tested on three scenarios of increasing difficulty. All these scenarios are based on a set of SWOT observations consisting in measurements of river surface deformations $(Z_{r,p}, W_{r,p})_{R,P+1}$. These three scenarios are:

(1) Only SWOT observations are available; this corresponds to the most challenging inverse problems on ungauged rivers. This case is addressed in Section 5.
(2) SWOT observations and multi-temporal priors on discharge are available. These

prior rough values (potentially a unique one) may be provided by worldwide hydrological models or discharge databases. This case is addressed in Section 6.

(3) SWOT observations and one in-situ measurement of bathymetry (denoted by b_{ref}) is available. In this case an accurate effective bathymetry surrounding ($\pm$100 km) the reference value b_{ref} can be inferred. This case is addressed in Section 6.

For all scenarios and river cases a worldwide river database - the SWOT river database [46,47] in construction - contains at least one discharge value (inter-annual average value of discharge modeled with a large scale water balance model or estimated from ancillary hydrological databases) and one friction estimate. Given the $R\times(P+1)$ WS observations, the complete control vector c formed by $(P+4R+2)$ unknowns, see (6), is computed by the VDA process. It is formed by $(P+1)$ unknown discharges, $4R$ unknown river bed elevations and a friction parameterization depending on h (forward model state) and two constant parameters.

4.3. First guess computation

In order to be as realistic as possible, first guesses of the VDA process are infered from available worldwide databases and/or hydrological models. However for many rivers worldwide in-situ measurements will not be accessible or even do not exist. First guess values $c^{(0)} = \left(Q_{in,0}^{(0)}, ..., Q_{in,P}^{(0)}; b_1^{(0)}, ..., b_R^{(0)}; \alpha^{(0)}, \beta^{(0)}\right)^T$ of the parameter vector, see (6) and (9), are obtained as follows:

- At inflow, we set: $\forall p \in [0...P]$, $Q_p = Q_{MAF} \equiv Q_{in}^{(0)}$.
  Q_{MAF} may be obtained by retrieving Mean Annual Flow (MAF) either from the SWOT a-priori river database under construction see e.g. [46] or from the global Water Balance Model (WBM) [48].
- A constant prior on Manning-Strickler friction coefficient $K^{(0)}$ is estimated from the SWOT a-priori river database. The $K^{(0)}$ values in the present case are 25, 33 and 25 respectively for the Garonne, Po, and Sacramento rivers, corresponding the $\alpha^{(0)}$ values with $\beta^{(0)} = 0$, see (3).
- A prior on the unobserved bathymetry $b^{(0)}$ (or equivalently the unobserved cross sectional areas $(A_{r,0})_r$ assuming a shape and knowing the low flow WS widths $(W_{r,0})_r$, see Section 2.3) can be obtained by the following three methods, see Tab. 3.
  - *"Manning"*- the unobserved low flow bathymetry $b^{(0)}$ is obtained by inverting the Manning equation (4) applied to a single flow line with $K^{(0)}$ and $Q_{in}^{(0)}$ constant in space and time,
  - *"Manning-multi"*- low flow bathymetry $b^{(0)}$ is obtained by inverting Eqn (13) with numerous flow lines that represent a large range of flow regimes and variabilities on the studied river. In that case, flow lines can be grouped by deciles and put in correspondence with deciles of discharge from the worldwide available WBM model or a discharge database.
  - *"Low Froude"*- In the case where one bathymetry measurement is available, an effective bathymetry is derived from WS observables and the low Froude equation (5).

For the three scenarios, *"Manning"*, *"Manning-multi"* and *"Low Froude"*, the first-guesses $Q^{(0)}$, $\alpha^{(0)}$ and $\beta^{(0)}$ are determined using the same methods described above.

Table 3. Summary of experiments (*: two overpasses at days 9 and 19 every 21 days repeat period)

Experiment	River	$b^{(0)}$	Δx_{obs}	Δt_{obs}	$\sigma_Z^{perturb}$	σ_Z	σ_Q	σ_{α_K}	σ_{β_K}	σ_b
A.1	Garonne	Manning	200 m	1 day	0	25 cm	15 m3/s	0.5	0.01	25 cm
A.2	Garonne	Manning-multi	200 m	1 day	0	25 cm	15 m3/s	0.5	0.01	25 cm
A.3	Garonne	Low Froude	200 m	1 day	0	25 cm	15 m3/s	0.5	0.01	25 cm
A.4	Garonne	Manning	200 m	1 day	25 cm	25 cm	15 m3/s	0.5	0.01	25 cm
A.5	Garonne	Manning-multi	200 m	1 day	25 cm	25 cm	15 m3/s	0.5	0.01	25 cm
A.6	Garonne	Low Froude	200 m	1 day	25 cm	25 cm	15 m3/s	0.5	0.01	25 cm
B.1	Po	Manning	200 m	1 day	0	25 cm	71 m3/s	0.5	0.01	25 cm
B.2	Po	Manning-multi	200 m	1 day	0	25 cm	71 m3/s	0.5	0.01	25 cm
B.3	Po	Low Froude	200 m	1 day	0	25 cm	71 m3/s	0.5	0.01	25 cm
B.4	Po	Manning	200 m	1 day	25 cm	25 cm	71 m3/s	0.5	0.01	25 cm
B.5	Po	Manning-multi	200 m	1 day	25 cm	25 cm	71 m3/s	0.5	0.01	25 cm
B.6	Po	Low Froude	200 m	1 day	25 cm	25 cm	71 m3/s	0.5	0.01	25 cm
C.1	Sacramento	Manning	200m	21 days*	0	25 cm	20 m3/s	0.5	0.01	25 cm
C.2	Sacramento	Low Froude	200m	21 days*	0	25 cm	20 m3/s	0.5	0.01	25 cm
C.3	Sacramento	Manning	200m	21 days*	34 cm	25 cm	20 m3/s	0.5	0.01	25 cm
C.4	Sacramento	Low Froude	200m	21 days*	34 cm	25 cm	20 m3/s	0.5	0.01	25 cm

Only the method used to compute $b^{(0)}$ differs for each scenario.

It is important to point out that if $b^{(0)}$ is estimated from the simple one-value Manning-Strickler relation, this first crucial estimation of the bathymetry is highly sensitive to any error made on the a-priori estimation of $(Q^{(0)}, K^{(0)})$.

These three scenarios are tested on all river cases, either with perfect or noisy observations, except for the Sacramento River were the second scenario cannot be computed. Indeed the small number of overpasses (17) is not sufficient to determine flow regimes an thus statistically relate observations with deciles of discharge in databases.

The criteria used to evaluate the performance of the estimations at assimilation times are the classical $RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Q_t^{est}-Q_t^{true})^2}{n}}$ and relative RMSE $rRMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{Q_t^{est}-Q_t^{true}}{Q_t^{true}}\right)^2}$ where Q_{in}^{est} (resp. Q_{in}^{true}) is the estimated/infered (resp. observed) inflow discharge of size n in time.

5. Hydraulic inferences from WS long time-series on ungauged rivers

This section presents the numerical inference of the complete control vector (6) in the 1D Saint-Venant model (1) in the case of ungauged rivers: WS observations distributed in space and time (SWOT like data) only are available. The inferences are performed in the SWOT Cal/Val scenario for the Po and Garonne rivers and for a real like SWOT scenario for the Sacramento River. In this last case, a preliminary analysis of the inverse problem based on the identifiability map (see [28]) helps to define the control frequency. Finally, a formal re-scaling of the Saint-Venant equations demonstrates the intrinsic ill-posedness feature of the present inverse problem; this (simple) calculation explains the bias observed in the numerical results.

5.1. SWOT Cal/Val scenario: Po and Garonne rivers

The infered inflow discharge for the Po and Garonne Rivers and this scenario are shown in purple on Fig. 2 (Scenarios A.4 and B.4 in Tab. 3). For each of the two rivers, the true daily inflow hydrograph over 90 days (green curve) is satisfactorily retrieved with $rRMSE_Q = 24.3\%$ on the Po River the and $rRMSE_Q = 9\%$ on the Garonne River at assimilation points (corresponding to observation points) - 8.5% and 4.8% respectively with perfect observations. Concerning the bathymetry, the results are listed in Tab. 5 and plotted on Fig. 5 using B^* (deviation from the trend) for better reading. The estimated bathymetry $b_{1..R}$ is improved by the VDA process on the Garonne River (prior $RMSE_{b^{(0)}} = 0.39$ decreases to $RMSE_b = 0.31$) ; it remains close to the first guess in the Po River case. Let us recall that the friction parameter K is model-equations-geometry dependent; its calibrated value compensates various modeling errors.

For both rivers considered as ungauged i.e. with prior $c^{(0)}$ defined from the a-piori river database only (scenarios A4,5 ; B4,5 and A1,2 ; B1,2 with perfect observations, Tab. 5), the inference of discharge remains robust and accurate. Most of the identification errors are absorbed by the friction coefficient. It can be noticed that an equifinality issue may remain between the bathymetry and the friction. Indeed, on the Po River while discharge is very well retrieved, neither the bathymetry nor the friction coefficient is significantly improved by the VDA process. Different (friction, bathymetry) pair values can produce at least similar (correct) discharge given a dataset. However this "optimal value" of $(K, b(x))$ solution of (10) can be refined by solving afterwards the algebraic model presented in Section 7, see Tab. 5.

The "*Manning*" method has been performed using both the WBM discharge [48] and the discharge from the SWOT a-priori river database, see e.g. [46] (respectively cases A.1.a, A.4.a, B.1.a and B.4.a; or A.1.b, A.4.b, B.1.b and B.4.b in Tab. 4). Recall that the "*Manning*" method used to determine the prior value of $b^{(0)}$ is based on a single pair $(Q^{(0)}, K^{(0)})$, therefore it is highly sensitive to any error in these first-guess values.

If measuring the inference accuracy in terms of rRMSE on Q, the results obtained for these 12 experiments show that rough priors $Q^{(0)}$, see paragraph 4.3, may lead to poor estimations of the discharge, see Tab. 5. For the A.1.a case $rRMSE_Q = 4.8\,\%$ which is really better than for the A.1.b case where $rRMSE_Q = 30.5\,\%$**.** For the A.4.a case $rRMSE_Q = 9.0\,\%$ and $rRMSE_Q = 33.0\,\%$ for the A.4.b case. The same observation can be made for the Po River.

However, as shown in Figure 2, the infered temporal variations of Q remain excellent in all cases. Indeed, if we compute the same rRMSE but using $Q^*(x,t) = \frac{\bar{Q}_{true}}{\bar{Q}^{(0)}} Q(x,t)$ (i.e. the shift of the prior is taken into account) the obtained values are: $rRMSE_{Q*} = 7.6\,\%$ for the A.1.b case, $rRMSE_{Q*} = 13.3\,\%$ for the A.4.b case, $rRMSE_{Q*} = 11.0\,\%$ for the B.1.b case and $rRMSE_{Q*} = 24.9\,\%$ for the B.4.b case. Thus the rRMSE obtained are comparable with the rRMSE obtained using the good prior. This high accuracy of the discharge temporal variation independently of the priori quality is mathematically explained in the next paragraph 5.2.

These results show that the estimations are highly sensitive to the accuracy of the inflow prior if using the simple *"Manning"* method. On the contrary, if using multiple accurate discharges values (here from the GRDC database) that is using the *"Manning-Multi"* method, these estimations are much more robust.

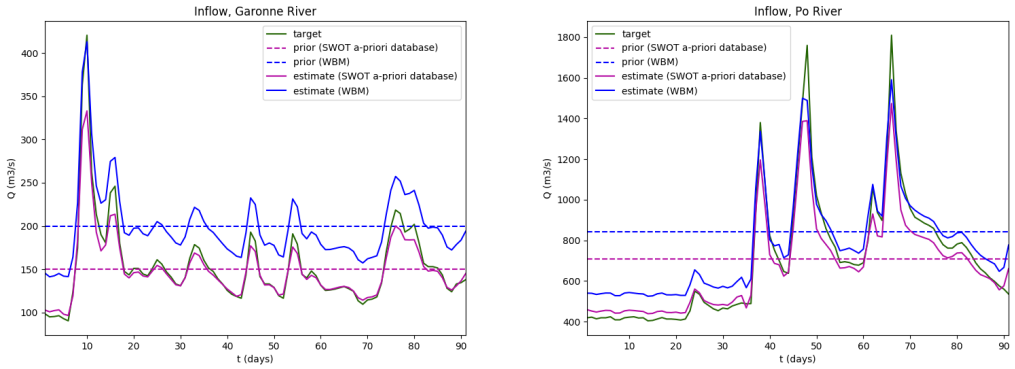

Figure 2. Inflow discharge (true=target, prior, infered=estimate) with daily SWOT-like observations. (Left) Garonne River Case A.3. (Right) Case B.3

Table 4. Scores of the inversions performed from different priors ("a" for SWOT a priori database, "b" for WBM) defining the hydrograph first guess $Q^{(0)}$.

Case	River	σ_Z^{obs}	Q_{prior} (m^3/s)	$RMSE_{b^{(0)}}$ (m)	$RMSE_Q$ (m^3/s)	$rRMSE_Q$ (%)	$rRMSE_{Q*}$ (%)	$RMSE_b$ (m)
A.1.a	Garonne	0	150	0.18	13.4	5.1	5.0	0.43
A.1.b	Garonne	0	270	0.51	40.6	30.5	7.5	0.73
A.4.a	Garonne	25 cm	150	0.39	24.0	10.3	9.8	0.31
A.4.b	Garonne	25 cm	270	0.54	41.8	33.0	13.3	0.45
B.1.a	Po	0	710	0.61	83.9	8.5	8.5	0.75
B.1.b	Po	0	841	0.91	88.4	17.1	11.0	0.92
B.4.a	Po	25 cm	710	0.92	185.9	24.8	24.8	1.04
B.4.b	Po	25 cm	841	1.12	235.3	31.5	24.9	1.24

5.2. On the equifinality issue & the importance of one ancillary data

The numerical results above show that the space time variations of the infered discharge are very accurate but may present a bias, see Fig. 2. However the bias vanishes if the (scalar) prior value $Q^{(0)}$ is sufficiently accurate; $Q^{(0)}$ may be the mean value of Q. As a consequence if the a-priori value(s), see paragraph 4.3, are far from reality, the rRMSE criteria on the estimation may be large, see e.g. case A.4.b in Tab. 4; while the rRMSE computed from $Q^*(x,t)$ remains excellent, Fig. 2. This bias (in other words this equifinality issue) can be explained by the simple re-scaling calculation of the flow equations below.

Let $\bar{Q}$ be a scalar value e.g. a mean value of Q or even K. The variables (A, Q, h) are re-scaled as follows: $(A_*, Q_*, h_*) = (A, Q, h)/\bar{Q}$. We set: $A_* = h_* W$; W the measured surface width. The mass equation divided by $\bar{Q}$ reads: $\partial_t(A_*) + \partial_x(Q_*) = 0$. Therefore as it is well known, in terms of mass, rescaling Q implies to rescale A by the same factor (or equivalently rescaling the water depth h). The momentum equation divided by $\bar{Q}$ reads:

$$\partial_t(Q_*) + \partial_x\left(\frac{Q_*^2}{A_*}\right) + gA_*\,\partial_x Z = -gA_* S_f$$

with $S_f \equiv S_f(A, Q; h; K) = \frac{1}{K^2}\frac{|Q|Q}{A^2 h^{4/3}}$.
A short calculation shows that: $S_f(A, Q, h; K) = S_f(A_*, Q_*, h_*; \bar{K}^{-2/3}K)$. Therefore, given the WS measurements (W, Z), the 1D Saint-Venant equations (1) with K as Manning-Strickler's coefficient is equivalent to the same equations but in the re-scaled variables (A_*, Q_*, h_*) with the Manning-Strickler coefficient equal to $(\bar{K}^{-2/3}K)$.
As a consequence the infered solution $(A, Q; K)$ obtained by inverting the Saint-Venant model is those corresponding to the prior value(s) introduced into the vector c_{prior}, see (9) and Section 4.3. (Recall that the first guess value $Q^{(0)}$ and values considered at downstream for the rating curve conditions are consistent).

This formal re-scaling of the equation and the numerical results, Fig. 2, show that the inversions from WS measurements only are accurate to retrieve the time variations of the inflow discharge $Q_{in}(t)$ (therefore the space-time variation throughout the domain) but it may be up to a multiplicative factor (the bias). This equifinality issue is solved (the bias vanishes) as soon as a reference value of one of the three fields A_0 (or equivalently b) Q or K is provided. One good global mean value may be enough. The low-complexity algebraic system (13) is solved to obtain the first guess values (see Section 4.3). Such a context is considered in next section.

It is worth noticing that the same inversions based on the (scalar) Manning-Strickler equation or on the algebraic system (4) (and not on the dynamic Saint-venant equations) present equifinality issues even stiffer. Indeed, given a set of WS measurements $(Z_{r,p}, W_{r,p})_{R,P+1}$ and an effective low flow bathymetry $A_{0,r}\ \forall r$ (therefore the effective cross-sectional area $A_{r,p} = A_{r,0} + \delta A_{r,p}\ \forall r \in [1..R]$ is given), the inference of the *ratio* $Q_{r,p}/K_{r,p}$ is possible from (4); but not the pair $(Q_{r,p}, K_{r,p})$! Infering the discharge value from the WS measurements and local Manning-Strickler's laws i.e. the low Froude system (4), requires additional information on the friction coefficient; otherwise the resulting uncertainty on Q is proportional to the uncertainty on K^{-1}, and reciprocally. Of course, given a strong prior information on K or Q (e.g. an a-priori pdf), the corresponding uncertainty on the infered value may be computed.

5.3. Identifiability map & preliminary analysis

The discharge inference capabilities depend on the space-time sampling of the observations and on the flow dynamics. An instructive reading of the inverse problem can be obtained by plotting the "identifiability map" introduced in [28]. The identifiability map represents in the (x, t) plane the complete information: the "space-time windows" observed by SWOT, the hydrodynamic wave propagation (1D Saint-Venant model in fluvial regime) and the misfit to the "local equilibrium" (local misfit between the steady state uniform flow and the dynamic flow), see [28] for details. This preliminary analysis enables to roughly estimate the inflow time intervals that can be identified by the VDA process; indeed the inflow discharge arise from these observed "space-time windows". This qualitative reading of the inverse problem is instructive since it enables to roughly estimate if the sought information has been observed or not. In the present case (Sacramento River case), the reach of 147 km long is completely observed by two SWOT tracks respectively at the 9th and 19th days for each 21 days repeat period. As the Sacramento River is considered ungauged in this experiment, the real identifiability map cannot be computed. To circumvent this point, we compute an *a-priori* identifiability map as follows:

- A VDA inference of the control vector c is performed from prior information given by the "*Manning*" method (Section 4.1) and SWOT observables.
- Using this control vector c, a forward model run is performed (the Saint-Venant model (1)).
- Finally, the "equilibrium misfit" $(S_f - S_0)$ (see [28] for detailed explanations) and the wave propagation times (T_{wave}) are calculated using the output of the forward model.

The obtained identifiability map is plotted on Fig. 3. Due to the observation layover, there are few unobserved zones for the second SWOT pass #527, for example at $x = 80$ km and $t = 19, 40, 61$ days. Given the WS observation of each reach at a given time, on Fig. 3 purple dots at the upstream BC represent the foot of each hydraulic characteristics - the upward hydraulic information propagation. Finally the identifiability of $Q_{in}(t)$ seams to be possible on time windows of $\sim$3 days ($\sim$2 days for the flood peak recession at day 51) before observation times.

The "equilibrium misfit"$(S_f - S_0)$ is represented on the map; it highlights where and when the flow is not locally steady and uniform. The magnitude of this equilibrium misfit tends to increase when a flood wave is traveling through the domain as shown by the observation at day 51, and day 61 corresponding to another peak entering the studied domain, Fig. 3 and 4. The map indicates that the WS deformations due to the flood peak between day 51 and 61 have not been observed. Moreover the mean estimated upward information propagation time is of 72h.

Consequently, from this analysis we define 3 assimilation points every 12h, hence corresponding to half the mean propagation time, before observations at days 51 and 61. Those observations may contain information respectively on the first flood peak recession and the third flood peak rise. The mean estimated upward propagation time is 36h (estimation coming from "observable times", which are represented with vertical purple dots on the right side of the figure). Since these "observable times" are computed a-posteriori and using an estimation of the wave speeds, we dismissed the extrema values (we retain 70% of the values).

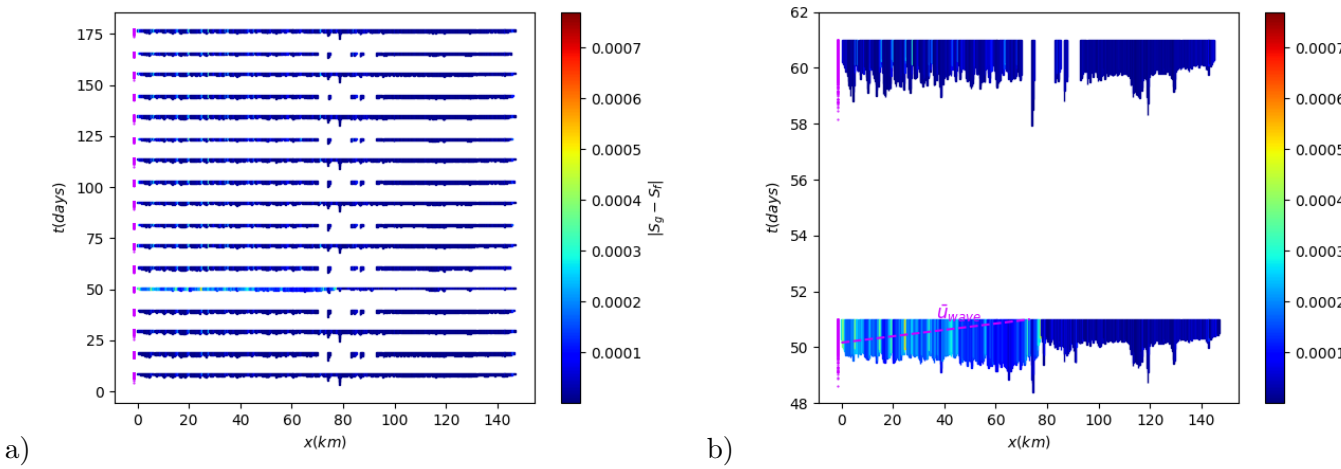

Figure 3. Identifiability map: overview in the (x,t)-plane of the inverse problem features: observables, hydraulic wave speed (and "equilibrium misfit"being absolute value of the source term in the Saint-Venant equations, color bar). For each observation of the domain in time, the vertical spreading corresponds to the time $\frac{\alpha}{u}\Delta x$ necessary for the upstream wave to cross an observed cell of size Δx; α is simply a dilatation factor for a sake of readability. . (a) The complete (x,t)-plane. (b) Zoom on the most varying time interval.

5.4. SWOT scenario: Sacramento River

The accuracy and the robustness of the algorithms proposed in this study are demonstrated on a real SWOT scenario, that is unfrequent observations compared to the hydrograph frequencies. The identifiability maps has provided some crucial information to set up the VDA process. The inflow discharge, bathymetry and friction infered by VDA on the Sacramento River are shown respectively on Fig. 4 (a) and Tab. 5. The discharge identification is accurate at each observation time and for the 3 points preceding observed flood peaks at $t = 51$ and 61 days. It is worth noticing that a basic approach consisting in infering discharge at observation times only would lead to a less accurate hydrograph inference as shown in Fig. 4 (b). The flood peak between days 51 and 61 is not retrieved since the preliminary analysis in terms of wave propagation above has shown (see Section 5.3) that the peak effects have not been observed.

Following [28], the identifiability index $I_{ident} = \frac{T_{wave}}{\Delta t_{obs}}$ of the river case is evaluated for a wave propagation time (T_{wave}) of about $72h$ (with $\Delta t_{obs} \sim 10$ days). This results to $I_{ident} \sim 0.3$, that is a relatively low identifiability index. Interestingly, even with this low identifiability value, the inference of discharge is accurate at observation points but also at the identification times defined before flood observations (following the analysis based on the identifiability map).

Finally we point out that the infered discharge is accurate while the bathymetry and friction need to be refined using other information or equations; this is done in next section using the low complexity systems presented in Section 2.3.

6. Inferences from WS observations and one in-situ measurement

Many worldwide rivers are not fully ungauged. In this section the benefit of using additional ancillary data is investigated. A single bathymetry measurement can be

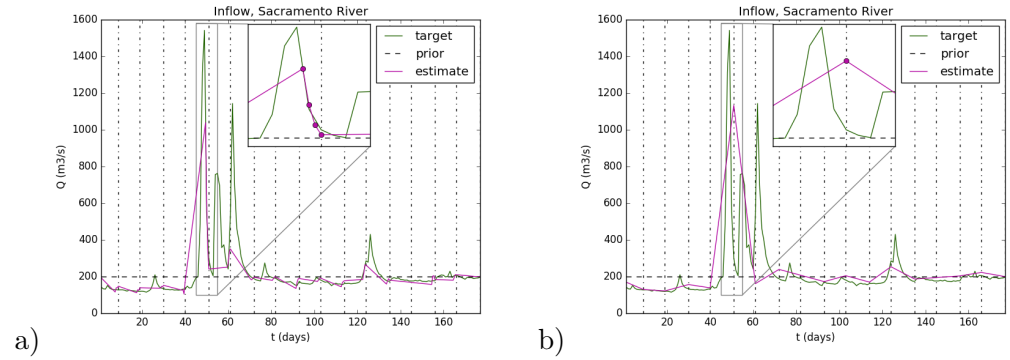

Figure 4. Inflow discharge (target=true, prior and infered=estimate with SWOT-HR observations), Sacramento River. Computations with $N = 4$ assimilation points every ($0h$,$12h$,$24h$,$36h$) before observation times (a) or at observation times only (b).

valued with the Low Froude model (5): an accurate first guess bathymetry can be infered and used in the control vector $c^{(0)}$ (case *"Low Froude"* Section 4.3).
A dedicated original bilinear system is derived to exploit discharge values potentially available in time and/or space. The latter being provided by a regional scale hydrological model or a gauging station. Next this first estimation of (A_0, K) is used as a first guess in the control vector $c^{(0)}$ (case *"Manning-multi"* Section 4.3).

On the accuracy of the low Froude - low complexity algebraic systems (4) and (5). As already mentioned, the systems (4) and (5) have been numerically assessed. For the Po and Garonne rivers, the HEC-RAS model (US Department of Defense, Army Corps of Engineers) has been used to generate synthetic observations; next Gaussian noises with realistic variances have been added, see Section 4.1. For the Sacramento river, SWOT Hydrology Simulator outputs have been considered. Next these direct models have provided the considered "true" flows measurements. Next given the effective true bathymetry b and the synthetic data $(Z, WS^{1/2})$, the system (4) has been solved and its solution has been compared to the "true" one. The obtained difference between the discharge values deduced from (4) and the "true" values equals approximatively $\pm 7\%$, see Fig. 7.

The obtained difference between the unobserved flow area A_0 infered from (5) and W_0, and the effective "true" value equals approximatively $\pm 10\%$ in $\approx$ 100 km long reaches, see Fig. 6.

It is worth noting that this bathymetry estimation presents an increasing "drift" with the distance to the reference measurement location x_{ref}. As discussed in [7], this drift is a-priori due to the non consistency of the steady state mass conservation $\partial_x Q = 0$. However this low Froude bathymetry is accurate enough if applied to a ~ 100 km portion only (for a single measurement h_{ref}). (Recall that neither tributary flows nor groundwater exchanges are taken into account in the present modeling).

6.1. Inference of A_0 from one in-situ bathymetry measurement

As already pointed out, infering the pair $(Q(t); K)$ given the bathymetry b is much less challenging; this is one of the context analyzed in detail in [28]. Therefore infering a reliable prior bathymetry before the iterative VDA process may be highly interesting. Here the control vector c, see (6), is composed by the upstream discharge, the

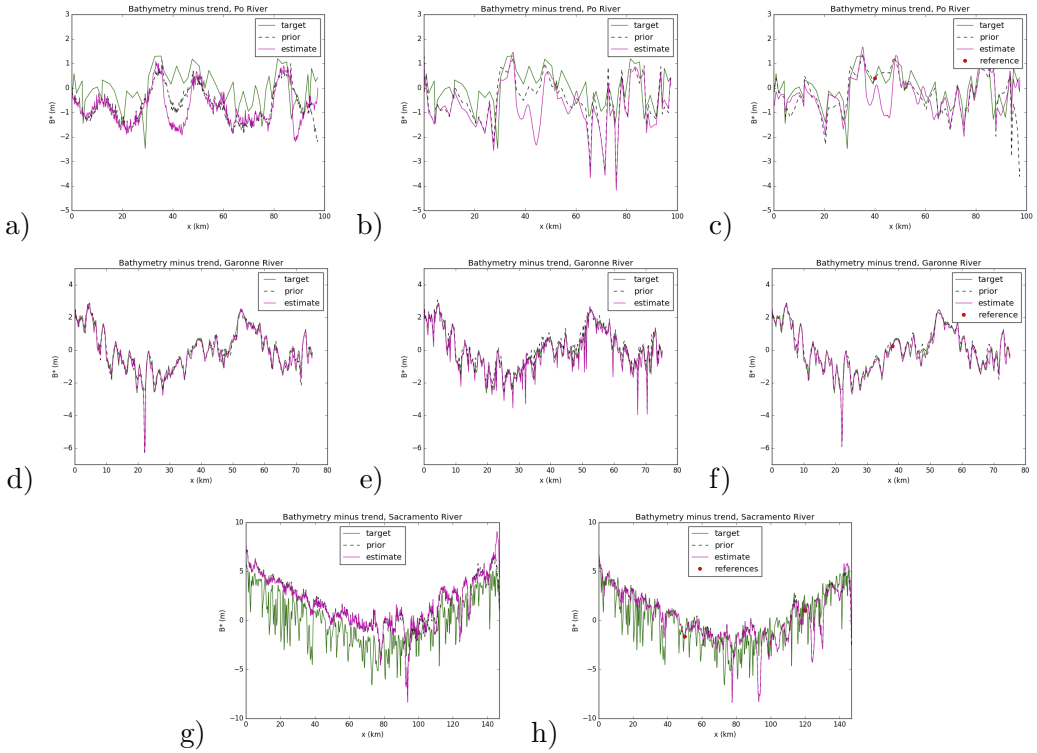

Figure 5. True and infered bathymetry deviation from trend $(B^*)_{1..R}$ compared to prior bathymetry estimations, noisy observations (Section 4.3): (a,d,g) "Manning". (b,e) "Manning-multi". (c,f,h) "Low Froude".

bathymetry and the friction. In this paragraph the different alternatives presented in Section 4.3 for estimating low flow bathymetry are tested on the three rivers, see Fig. 5. Each method gives fairly good estimations of $(b_r)_{1..R}$ (represented here as B^* the bed elevation b minus an average bathymetry trend). For each river, unsurprisingly, the best estimate is given by the low Froude approach, Eqn (5) requiring one in-situ point. Using the approaches based on the Manning equation and one flow line ($b^{(0)}$ *"Manning"*, Eqn (4)) leads to a shift of the infered prior bathymetry (Po and Sacramento rivers) whereas it is avoided if using the low complexity approach involving multiple flow lines better sampling flow regimes ($b^{(0)}$ *"Manning-multi"*, Eqn (13)).

In the Sacramento river case, a drift (increasing error in space) appears in the *"Low Froude"* bathymetry for nodes far from the reference point. Indeed if the basic hypothesis $\partial_x Q = 0$ of *"Low Froude"* model is not satisfied (that is an unsteady flow), such a drift appears due to the nature of the differential equation (a first order differential equation), see [7] for a detailed investigation. In order to avoid this drift, a segmentation of the river into two zones is performed and two reference bathymetry points are used for the *"Low Froude"* bathymetry prior estimation, Fig. 5 (Bottom Right).

6.2. Inference of (A_0, K) from one gauging station

This section aims at formulating an additional "low complexity model" to take advantage of discharge data available in space and/or time, for example provided by a large (regional) scale hydrological model or a gauging station. Again the derived equations

natively integrate the measurements in its coefficients.
It is assumed that spatially distributed values of Q are available in the domain for each overpass $p \in [0...P]$, that is $(Q_{r,p})_{r,p}$.
Let us assume that the friction varies both in space and time, that is $(K_{r,p})_{r,p}$. Then (4) reads as follows:

$$c_{r,p} \cdot K_{r,p}^{3/5} A_{r,0} + d_{r,p} \cdot K_{r,p}^{3/5} = Q_{r,p}^{3/5} \text{ for all } (r,p) \tag{13}$$

with $c_{r,p} = W_{r,p}^{-2/5} S_{r,p}^{3/10}$ and $d_{r,p} = c_{r,p}\, \delta A_{r,p}$. The coefficients $(c_{r,p}\ , d_{r,p})$ can be evaluated from the altimetry measurements. System (13) is multi-linear in $(K_{r,p}^{3/5} A_{r,0}, K_{r,p}^{3/5}, \bar{Q}_p)$. It contains $R(P+1)$ equations. It can be employed differently depending on the available information and the unknowns. If considering the full set of unknowns $(K_{r,p}^{3/5}, A_{r,0}, Q_{r,p})$, it is an underdetermined system since it has $R(2(P+1)+1)$ unknowns; therefore in this case it cannot be solved since it has an infinity of solutions.

Let us define the diagonal matrices D_c and D_d of dimensions $[R(P+1)]^2$ by: $M_c = diag\,(c_{r,p})$, $M_d = diag\,(d_{r,p})$. Let us define the vectors : $\tilde{K} = (K_{r,p}^{3/5})_{r,p} \in R^{R(P+1)}$, $A = (A_{r,0})_{r,} \in R^R$ and $\tilde{Q} = (Q_{r,p}^{3/5})_{r,p} \in R^{R(P+1)}$. Then (13) reads:

$$D_c\, Bil(\tilde{K}, A) + D_d \tilde{K} = \tilde{Q} \text{ in } R^{R(P+1)} \tag{14}$$

with the bilinear operator $Bil(\tilde{K}, A)_{r,p} = K_{r,p}^{3/5} A_{r,0}\ \forall r \forall p$. If $\tilde{Q}$ is given then (14) has $(P+2)R$ unknowns (the two vectors $\tilde{K}$ and A).

For K constant in space but varying in time (e.g. K is defined as a power law in h) then $\tilde{K} = (K_p^{3/5})_p \in R^{(P+1)}$. In this case, the $(P+1+R)$ unknowns $(K_p^{3/5}, A_{r,0})$ can be computed by solving the overdetermined bi-linear system (14) e.g. by employing a trusted region reflective algorithm.

For K constant in time but varying in space, $\tilde{K} = (K_r^{3/5})_r \in R^R$, the system (14) has $2R$ unknowns. Therefore if $P \geq 1$ (more than 2 overpasses), the solution $(K_r^{3/5}, A_{r,0})$ can be computed e.g. by a trusted region reflective algorithm.

Then the RHS of (14) reads: $\bar{Q} = \left((\bar{Q}_0^{3/5}, ..., \bar{Q}_0^{3/5}), ..., (\bar{Q}_p^{3/5} ... \bar{Q}_p^{3/5}), ..., (\bar{Q}_P^{3/5} ... \bar{Q}_P^{3/5})\right) \in R^{R(P+1)}$. Given $\bar{Q}$, estimations of $(A_{r,0})_r$ are obtained by computing the pairs $(K_{r,p}, A_{r,0})_{r,p}$ solution of (14) (with the values of $(K_{r,p})$ not necessarily used).

Recall the estimations of $Q = (Q_{r,p}) \in R^{R(P+1)}$ may be provided at large scale by hydrological models, then (14) may provide effective unobserved wetted areas $(A_{r,0})_r$. Also $Q = (Q_{r,p}) \in R^{R(P+1)}$ may be provided by the VDA process containing a bias (Section 3) too, then (14) may provide a the effective unobserved wetted areas $(A_{r,0})_r$ corresponding to Low-Froude flows at equilibrium.

6.3. Inferences of $(Q(t), b(x), K)$ by VDA using one in-situ measurement

The impact of a prior bathymetry value on the inversions accuracy is presented in this section. For each river, the infered inflow hydrograph starting from a "Low Froude bathymetry" prior, see Section 6.1 and Fig. 5, is shown on Fig. 6. For the Sacramento

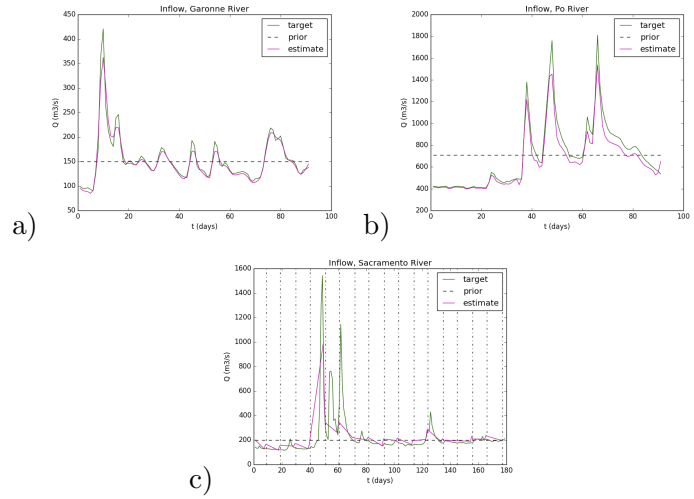

Figure 6. infered inflow discharge with the bathymetry priors infered using the Low-Froude model and one (1) in-situ point ($b^{(0)}$ "Low Froude"). Daily SWOT-like observations (noise $\sigma_Z = 0.25$) (a, b) and SWOT-HR observations (c).

river, both VDA runs C.3 and C.4, performed either with *"Manning"* or *"Low Froude"* bathymetry priors, result in comparable errors ($rRMSE_Q \sim 20\%$). The bathymetry is slightly degraded but the inversion remains robust. For the Garonne River the inversions performed with the 3 different bathymetry first guesses result in comparable errors ($rRMSE_Q = 8.5\%$ and 9.1%, runs C.3 and C.4). The finest bathymetry infered by VDA results from the use of *"Manning"* or *"Low Froude"* bathymetry first guesses - that provide fairly accurate priors in that case. For the Po River, using a *"Low Froude"* bathymetry results in the best error ($rRMSE_Q = 18.3\%$, run B.6) compared to other bathymetry first guesses (runs B.4 and B.5). This bathymetry is also slightly improved by the VDA process.

7. Real-time estimations of Q given (A_0, K)

The inversion method based on VDA is based on SWOT dataset measurements. Such a VDA approach may be viewed as an optimal estimation in a least-square sense with respect to the observations datasets; therefore it is highly advisable to perform it on a complete hydrological cycle (a year). The method provides an estimation of the rivers bathymetry therefore an intrinsic feature of the observed rivers (intrinsic generally at years time scale at least). The method setup is not obvious; moreover it is a bit CPU-time consuming. For all these reasons, this inversion method cannot be performed in real-time. Then a good strategy in the context of SWOT mission is the following.
The inversion method based on VDA is performed on a complete year SWOT observation set; it is the calibration (or "learning") period and relatively accurate values of $(A_{r,0}) \in R^R$ and $(Q_{r,p}) \in R^{R(P+1)}$ are estimated. Given these $(Q_{r,p}, A_{r,0})_{r,p}$ values, an effective space-time dependent friction coefficient $(K_{r,p})_{r,p}$ corresponding to the Low-Froude algebraic flow model (14) is straightforwardly obtained. Recall that the friction parameter K is fully model dependent; it is not an intrinsic feature of the flows we may keep. This low Froude - low complexity system (4) is quite accurate, see e.g. Fig. 7. Moreover it can be performed in real-time. Therefore given new observations

Table 5. Scores of the inversions performed from different priors defining the bathymetry first guess $b^{(0)}$, see Section 4.3.

Case	River	σ_Z^{obs}	b_{prior}	$RMSE_{b^{(0)}}$ (m)	$RMSE_Q$ (m^3/s)	$rRMSE_Q$ (%)	$RMSE_b$ (m)
A.1	Garonne	0	Manning	0.18	13.4	5.1	0.27
A.2	Garonne	0	Manning-multi	0.17	10.0	4.2	0.16
A.3	Garonne	0	Low Froude	0.16	11.9	5.2	0.22
A.4	Garonne	25 cm	Manning	0.39	24.0	10.3	0.31
A.5	Garonne	25 cm	Manning-multi	0.42	17.9	8.5	0.38
A.6	Garonne	25 cm	Low Froude	0.35	19.4	9.1	0.27
B.1	Po	0	Manning	0.61	83.9	8.5	0.75
B.2	Po	0	Manning-multi	0.73	91.2	10.2	0.88
B.3	Po	0	Low Froude	0.48	76.4	7.8	0.45
B.4	Po	25 cm	Manning	0.92	185.9	24.8	1.04
B.5	Po	25 cm	Manning-multi	0.83	183.7	24.3	1.07
B.6	Po	25 cm	Low Froude	0.64	144.8	18.3	0.60
C.1	Sacramento	0	Manning	2.29	30.1	11.5	2.33
C.2	Sacramento	0	Low Froude	1.95	54.5	15.9	2.01
C.3	Sacramento	34 cm	Manning	2.45	124.7	19.3	2.49
C.4	Sacramento	34 cm	Low Froude	1.84	141.2	20.2	1.82

$(Z_r, W_r)_R$, typically satellite observations acquired after the "learning period", the corresponding discharge values Q_r can be computed in real-time simply by evaluating the Manning-Strickler equation (4). Such real-time estimations are presented in Section 7.

7.1. Friction coefficient re-calibration

Given $(Q_{r,p}, A_{r,0})_{r,p}$ by the VDA process, a new (effective) friction coefficient corresponding to the low Froude - low complexity model is computed ("re-calibration" of K). It is done following one of the following two methods.

1) Space-time dependent values $K_{r,p}$ are obtained by solving (4). Next we set: $K_r = mean_p(K_{r,p})$. This mean value is computed on a subset of the WS observations e.g. by considering the 2^{nd} - 8^{th} deciles of the flow profiles (20 overpasses were selected in this study). Indeed this corresponds to a relatively wide and representative range of flow regimes without the extremes.
2) Space-time dependant values $K_{r,p}$ are obtained by solving (4) as in Method 1) above. Water depth $h_{r,p}$ are straightforwardly obtained as $(Z_{r,p} - b_r)$, with the bathymetry values infered by VDA. Finally for each reach r, we make fit the power-law $\hat{K}_r(h) = \alpha_r h^{\beta_r}$ with the values previously obtained.

These two methods are adequate in the present river cases since no over-bank flow occurs. However they are applied to the Garonne and Po rivers only since for Sacramento River the number of observations is not sufficient to select reliable deciles of the flow profiles, see Fig. 7. To mimic a real calibration/validation setup, data of the Po and Garonne rivers were divided in two sets. The first set corresponds to the Cal/Val period (90 days) considered for the VDA inversion method. The second set (9 months) were then used for the validation.

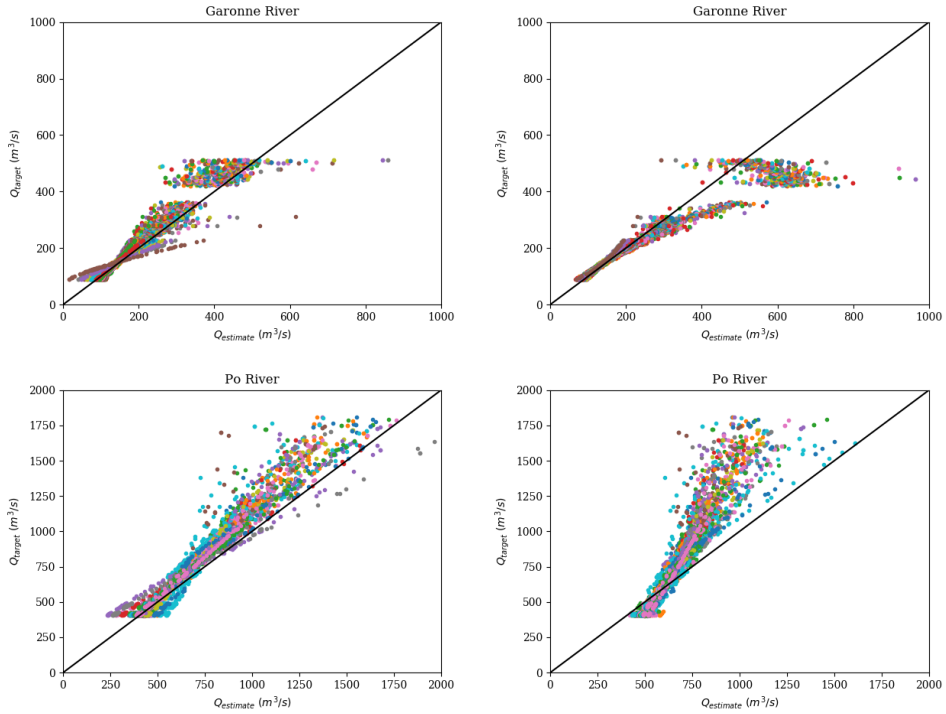

Figure 7. Discharge computations from the low complexity relations obtained after the re-calibration of the friction K obtained by the VDA process; Plot Q_{true} vs $Q_{infered}$ using (4) on the complete observations set. (Top). Discharge Garonne River; (Bottom) Po River. (Left) With $K(h)$ hence K_p; (Right) With K_r.

7.2. Real-time estimations of Q

Given the re-calibrated friction coefficients $(K_r)_{1..R}$, the discharge is computed by solving the Manning-Strickler equations system (4) with its coefficients provided by the newly acquired SWOT observations. The obtained values in the case of the Po and Garonne Rivers are plotted on Fig. 7 (for all flow lines). For both rivers, using a friction coefficient K_r constant in time provides the best results. The resulting errors are: $rRMSE_Q = 6.3\%$ for the Garonne River and $rRMSE_Q = 10.6\%$ for the Po River. Note that these performances are evaluated at all observation times and all nodes whereas for the VDA inference (Section 5) the performances were evaluated only for the upstream node of the river domain i.e. only for $r = 0$.

The second method for re-calibrating the friction $K_{r,p}$ provides a lower accuracy but still acceptable; the errors are: $rRMSE_Q = 9.1\%$ for the Garonne River and $rRMSE_Q = 25.3\%$ for the Po River.

Recall that the flow model is calibrated from the observed flow lines hence presenting a minimal and a maximal WS elevation. The re-calibrated friction coefficient K is related to this flow lines range (hence regime range). If K is defined as a power-law $K_{r,p}$ and if the newly acquired WS elevation is greater than the previously observed WS than the friction power-law is prone to an over-estimation. That is the reason why we encourage to prefer the first method (considering K_r) to the second one (power-law).

Finally the proposed approach based on the low complexity systems employed after a re-calibration of the friction coefficient turns out to be very promising to perform the discharge in real-time, that is in an operational way.

8. Conclusion

This study proposes a new hierarchical computational inversion method to infer the discharge $Q(t)$, an effective bathymetry $b(x)$ with a corresponding friction coefficient K from altimetry Water Surface (WS) measurements, more specifically data from the forthcoming SWOT satellite [1]. The inversion method is based on a combination of an advanced Variational Data Assimilation (VDA) formulation applied to the classical Saint-Venant equations (1D shallow-water) and original algebraic systems (low-Froude, locally permanent flows). The VDA formulation takes into account adequate scale dependency and a-priori error measurement amplitudes. The algebraic systems natively integrate the measured quantities; they may be differently employed depending on the fields that are not known. They enable to exploit in a consistent manner databases to define the first guesses (prior information) of the iterative VDA process. Moreover they enable to compute in real-time the discharge past the "calibration-learning period" (i.e. the assimilation of a complete hydrological year dataset).

Three rivers, $\sim$ 100 km long each, have been considered with two scenarios of observation: the SWOT Cal-Val orbit with $\sim$ 1 day period (or any equivalent multiple-sensor measurements) and SWOT like data with $\sim$ 21 days period (with 1 to 4 passes at mid-latitudes). The corresponding inversions are highly challenging since relying on relatively sparse observations (both in space and time) compared to the potential flow dynamics. Indeed the flows present low identifiability indexes as defined in [28]. Preliminary analyses based on the identifiability maps introduced in [28] enable to define adequate time grids for the identification process.

For ungauged rivers and/or in total absence of good prior information on the flow, the inversion algorithm provides accurate space-time variations of Q with an effective bathymetry $b(x)$ and a corresponding friction coefficient $K(h)$ (K function of the water depth h). However if the prior mean value of one of the three infered fields (typically those of Q) is far from reality, a bias on the infered hydrograph $Q(t)$ may remain. But as soon as a good mean value is provided (e.g. a mean discharge value from a database or from a large scale hydrological model), or a single reference value of bathymetry, the bias vanishes: the discharge $Q(t)$ is perfectly recovered even in terms of amplitudes (RMSE of a few percent at observation times are obtained).

Past the calibration period by VDA, the estimated values of $Q(t)$ and $b(x)$ obtained by the VDA process are kept and a new (effective) friction coefficient $K(x)$ corresponding to the low complexity flow model is computed. Next, this low complexity (algebraic) model can provide estimations of the discharge Q *in real-time from newly acquired satellite data.*

This new and complete inverse method fulfil the conditions of an operational solution to the estimations of rivers discharge at global scale from the forthcoming SWOT satellite mission (launch planned in 2021).

All the present equations and algorithms are implemented into the open-source computational software DassFlow [32]. On-going investigations focus on inversions applied to larger rivers portions presenting lateral fluxes and complete river networks.

Authors' contributions & acknowledgments

The corresponding author has elaborated the hierarchical flow model equations and the VDA formulation. The first author has performed all the numerical results, after having greatly enriched the computational software DassFlow1D from the previous version developed by P. Brisset [28]. The third author has greatly contributed to the

real data model setups and their analysis (and suggested the sensitivity to first guesses and densified temporal identifiability). The last author has nicely contributed to the VDA formulation during his beginning of PhD.

The authors K. Larnier (software engineer at CS corp.) and J. Verley have been funded by CNES.

The authors acknowledge M. Durand from Ohio State University for providing the Sacramento SWOT-HR dataset; also H. Roux from IMFT and INPT-Toulouse University has provided a fine expertise on the Garonne River dataset.

This article is dedicated to our bright colleague Jonas Verley, cool friend, who has disappeared so fast.

References

[1] Rodríguez E, et al. Swot science requirements document. JPL document, JPL. 2012;.

[2] Calmant S, Cretaux JF, Remy F. Principles of radar satellite altimetry for application on inland waters. In: Baghdadi N, Zribi M, editors. Microwave remote sensing of land surface. Elsevier; 2016. p. 175 – 218.

[3] Biancamaria S, Lettenmaier DP, Pavelsky TM. The swot mission and its capabilities for land hydrology. Surveys in Geophysics. 2016;37(2):307–337.

[4] Chow V. Handbook of applied hydrology. McGraw-Hill Book Co, New-York, 1467 pages. 1964;.

[5] Bjerklie DM. Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information. Journal of hydrology. 2007;341(3):144–155.

[6] Durand M, Neal J, Rodríguez E, et al. Estimating reach-averaged discharge for the river severn from measurements of river water surface elevation and slope. Journal of Hydrology. 2014;511:92–104.

[7] Garambois PA, Monnier J. Inference of effective river properties from remotely sensed observations of water surface. Advances in Water Resources. 2015;79:103–120.

[8] Yoon Y, Garambois PA, Paiva R, et al. Improved error estimates of a discharge algorithm for remotely sensed river measurements: Test cases on Sacramento and Garonne Rivers. Water Resources Research. 2016;52(1):278–294.

[9] Durand M, Gleason C, Garambois PA, et al. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. Water Resources Research. 2016;.

[10] Roux H, Dartus D. Use of parameter optimization to estimate a flood wave: Potential applications to remote sensing of rivers. J of Hydrology. 2006;328:258–266.

[11] Ricci S, Piacentini A, Thual O, et al. Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. Hydrol Earth Syst Sci. 2011; 15:3555–3575. 11.

[12] Munier S, Polebistki A, Brown C, et al. Swot data assimilation for operational reservoir management on the upper niger river basin. Water Resources Research. 2015;51(1):554–575.

[13] Sasaki Y. Some basic formalisms in numerical variational analysis. Citeseer; 1970.

[14] Le Dimet FX, Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus A: Dynamic Meteorology and Oceanography. 1986;38(2):97–110.

[15] Cacuci D, Navon I, Ionescu-Bugor M. Computational methods for data evaluation and assimilation. Taylor and Francis CRC Press: Boca Raton; 2013.

[16] Panchang V, O'Brien J. On the determination of hydraulic model parameters using the adjoint state formulation. Modeling marine system. 1989;1:5–18.

[17] Chertok D, Lardner R. Variational data assimilation for a nonlinear hydraulic model. Applied mathematical modelling. 1996;20(9):675–682.

[18] Sanders B, Katopodes N. Control of canal flow by adjoint sensitivity method. Journal of irrigation and drainage engineering. 1999;125(5):287–297.
[19] Bélanger E, Vincent A. Data assimilation (4d-var) to forecast flood in shallow-waters with sediment erosion. Journal of Hydrology. 2005;300(14):114 – 125. Available from: http://www.sciencedirect.com/science/article/pii/S0022169404002914.
[20] Honnorat M, Monnier J, Le Dimet FX. Lagrangian data assimilation for river hydraulics simulations. Computing and Visualization in Science. 2009;12(5):235–246.
[21] Honnorat M, Monnier J, Rivière N, et al. Identification of equivalent topography in an open channel flow using lagrangian data assimilation. Computing and visualization in science. 2010;13(3):111–119.
[22] Honnorat M, Lai X, Monnier J, et al. Variational data assimilation for 2d fluvial hydraulics simulations. In: CMWR XVI-Computational Methods for Water Ressources. Copenhagen, june 2006.; 2006.
[23] Lai X, Monnier J. Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: mathematical method and test case. Journal of Hydrology. 2009; 377:1–11. 1-2.
[24] Hostache R, Lai X, Monnier J, et al. 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. 2010;390:257–268. 3-4; Available from: http://www.sciencedirect.com/science/article/pii/S0022169410004166.
[25] Monnier J, Couderc F, Dartus D, et al. Inverse algorithms for 2D shallow water equations in presence of wet dry fronts. application to flood plain dynamics. Advances in Water Resources. 2016;97:11–24.
[26] Gejadze I, Monnier J. On a 2d zoom for the 1d shallow water model: Coupling and data assimilation. Computer methods in applied mechanics and engineering. 2007;196(45):4628–4643.
[27] Marin J, Monnier J. Superposition of local zoom models and simultaneous calibration for 1d–2d shallow water flows. Mathematics and Computers in Simulation. 2009;80(3):547–560.
[28] Brisset P, Monnier J, Garambois PA, et al. On the assimilation of altimetric data in 1d saint-venant river flow models. Adv Water Ress. 2018;119:41–59.
[29] Gessese AF, Sellier M, Van Houten E, et al. Reconstruction of river bed topography from free surface data using a direct numerical approach in one-dimensional shallow water flow. Inverse Problems. 2011;27(2):025001.
[30] Gejadze I, Malaterre PO. 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):405–430. Fld.4273; Available from: http://dx.doi.org/10.1002/fld.4273.
[31] Oubanas H, Gejadze I, Malaterre PO, et al. River discharge estimation from synthetic swot-type observations using variational data assimilation and the full saint-venant hydraulic model. Journal of Hydrology. 2018;559:638–647.
[32] DassFlow. Data assimilation for free surface flows. Mathematics Institute of Toulouse - INSA group - CNES - CNRS - ICUBE; 2018. Available from: http://www.math.univ-toulouse.fr/DassFlow.
[33] Carlier M. Hydraulique générale et appliquée. Paris, France: Eyrolles; 1982.
[34] Bjerklie DM, Dingman SL, Bolster CH. Comparison of constitutive flow resistance equations based on the manning and chezy equations applied to natural rivers. Water Resources Research. 2005;41(11).
[35] Frasson R, Wei R, Durand M, et al. Automated river reach definition strategies: Applications for the surface water and ocean topography mission. Water Resources Research. 2017;53(10):8164–8186. Available from: http://dx.doi.org/10.1002/2017WR020887.
[36] Gilbert JC, Lemaréchal C. Some numerical experiments with variable-storage quasi-newton algorithms. Mathematical programming. 1989;45(1-3):407–435.
[37] Hascoët L, Pascual V. The Tapenade Automatic Differentiation tool: Principles, Model,

and Specification. ACM Transactions On Mathematical Software. 2013;39(3).
[38] Bouttier F, Courtier P. Data assimilation concepts and methods march 1999. Meteorological training course lecture series ECMWF. 2002;:59.
[39] Monnier J. Variational data assimilation: from optimal control to large scale data assimilation. Open Online Course, INSA Toulouse; 2018.
[40] Lorenc A, Ballard S, Bell R, et al. The met. office global three-dimensional variational data assimilation scheme. Quarterly Journal of the Royal Meteorological Society. 2000; 126(570):2991–3012.
[41] Haben S, Lawless A, Nichols N. Conditioning and preconditioning of the variational data assimilation problem. Computers & Fluids. 2011;46(1):252–256.
[42] Haben S, Lawless A, Nichols N. Conditioning of incremental variational data assimilation, with application to the met office system. Tellus A. 2011;63(4):782–792.
[43] Tarantola A. Inverse problem theory and methods for model parameter estimation. Vol. 89. SIAM; 2005.
[44] Larnier K. Modélisation thermohydraulique d'un tronçon de garonne en lien avec l'habitat piscicole : Approches statistique et déterministe [dissertation]. Toulouse; 2010.
[45] Di Baldassarre G, Schumann G, Bates P. Near real time satellite imagery to support and verify timely flood modelling. Hydrological Processes. 2009;23(5):799–803.
[46] Andreadis K, Schumann G, Pavelsky T. A simple global river bankfull width and depth database. Water Resources Research. 2013;49(10):7164–7168.
[47] Allen GH, Pavelsky TM. Global extent of rivers and streams. Science. 2018;361(6402):585–588.
[48] Wisser D, Fekete B, Vörösmarty C, et al. Reconstructing 20th century global hydrography: a contribution to the global terrestrial network-hydrology (gtn-h). Hydrology and Earth System Sciences. 2010;14(1):1.
[49] Kaltenbacher B, Neubauer A, Scherzer O. Iterative regularization methods for nonlinear ill-posed problems. Vol. 6. Walter de Gruyter; 2008.