



Extraction of roughness parameters from remotely-sensed products for hydrology applications

Charlotte Emery^{1,2}, Kevin Larnier^{1,3}, Maxime Liquet¹, João Hemptinne¹, Arthur Vincent¹, and Santiago Peña Luque⁴

¹CS GROUP - France, Toulouse, France

²CERFACS, Toulouse, France

³INSA Toulouse, Toulouse, France

⁴CNES, Toulouse, France

Correspondence: Charlotte Emery (charlotte.emery@csgroup.eu)

Abstract. Along rivers, where local insitu gauges are unavailable, estimation of river discharge are indirectly derived from the Manning formula that relate discharge to geomorphological characteristics of the rivers and flow conditions. Most components of the Manning formula can currently be derived from spaceborne products except for two features: the unobserved always-wet bathymetry and the roughness coefficient. Global-scale applications use simplified equivalent riverbed shapes and empirical parameters while local-scale applications rely on finer model dynamics, field survey and expert knowledge. Within the framework of the incoming Surface Water and Ocean Topography (SWOT) mission, scheduled for a launch in 2022, and more particularly, the development of the SWOT-based discharge product, fine-resolution but global discharge estimates should be produced. Currently implemented SWOT-based discharge algorithms require prior information on bathymetry and roughness and their performances highly depend on the quality of such priors. Here we introduce an automatic and spaceborne-data-based-only methodology to derive physically-based roughness coefficients to use in one-dimensional hydrological models. The evaluation of those friction coefficients showed that they allow model performances comparable to calibrated models. Finally, we illustrate two cases of application where our roughness coefficients are used as-is to initiate the experiment: a data assimilation experiment designed to correct the roughness parameters and an application around the HiVDI SWOT-based discharge algorithm. In both cases, the roughness coefficients showed promising perspectives by reproducing, for the data assimilation application, and sometimes improving, in the SWOT discharge algorithm case, the calibrated-parameter-based performances.

1 Introduction

At the interface with the atmosphere and the oceans, continental water is at the core of the water cycle. It is therefore crucial to study the dynamics of continental waters to understand the Earth climatic system and close the water budget. More particularly, continental waters are freshwaters, essential to life form and human activities. Freshwaters represent only 2.5% of the total amount of water and an even smaller part is directly accessible, e.g in rivers and lakes. Nevertheless, the continental part of the water cycle is particularly active as freshwaters are continuously moving and renewed. Therefore, Oki and Kanae (2006) stated that freshwater availability is better measured through the estimation of water fluxes, and notably river discharge.



Nevertheless, river discharge remains a quantity particularly complicated to assess. Insitu stream gauges still stand as the most reliable measuring technique but the number of available and maintained insitu stations has always been limited and is steadily decreasing worldwide (Fekete and Vörösmarty, 2007; Hannah et al., 2011). Thankfully, for the past three decades, hydrology-related remotely-sensed data has considerably increased (Lettenmaier et al., 2015) and numerous studies have focused on ways to estimate discharge from remotely-sensed products. Among them, Tarpanelli et al. (2013) used the different water and land surface response from MODIS images near river bank to derive variations of river flow velocity and hence river discharge. Also using MODIS images, Van Dijk et al. (2016) examined the feasibility of relating observed inundated area to insitu discharge records in order to expand such statistical laws to ungauged basin while Paris et al. (2016) replicated traditional insitu techniques to measure discharge by deriving rating curves parameters relating pointlike water surface elevations from nadir altimetry and simulated discharges from a large-scale hydrological model over the Amazon basin.

Most recently, the incoming Surface Water and Ocean Topography (SWOT) mission, scheduled for a launch in 2022, and its expected river discharge product has pushed forward researches allowing to derive river discharge from satellite products (Durand et al., 2014, 2016; Hagemann et al., 2017; Bjerklie et al., 2018; Brinkerhoff et al., 2019; Tuozzolo et al., 2019; Larnier et al., 2021; Andreadis et al., 2020). Although all current 5 algorithms being developed relies on their own physical laws and/or mathematical/statistical models, they all plan to use SWOT river observations of water surface elevation (Z), river surface width (W) and slope (S) to derive an estimation of the river discharge. Moreover, they all rely, more or less directly, on the Manning-Strickler equation (Manning, 1891). Hence, the unobserved bathymetry, denoted A_0 , and the friction coefficient, modeled either as the Manning coefficient n or the Strickler coefficient K_s , remain unknown parameters. Therefore, SWOT discharge algorithm working groups agreed on simultaneously infer A_0 and n with the discharge (Durand et al., 2016). Mathematically speaking, the discharge inversion problem turns to be ill-posed but previous studies showed that if, among A_0 and n , one of them can be accurately estimated, the remaining parameter can be inferred together with the discharge (Larnier et al., 2021).

Therefore, a set of a priori parameters is preliminarily required to initialize the Manning-Strickler inversion problem. These parameters comprise direct priors values of the missing parameters or at least a prior distribution, e.g a mean value and a standard deviation. Henceforth, there is an emerging literature focusing on developing reliable priors (Peirong et al., 2017; Andreadis et al., 2020; Brinkerhoff et al., 2020; Larnier and Monnier, 2020). At first, it seems easier to get an accurate knowledge of the unobserved bathymetry A_0 as it a physically measurable variable. However, this measure appears complicated to apply at large scale. Indeed, current instruments on board of satellite can not "see" through water and measurements have to be obtained locally.

Even more, friction coefficients remain conceptual parameters that are not directly measurable on the field. Roughness coefficients are indispensable parameters or river-based modeling applications for both large-scale hydrology models such as ISBA-CTRIP (Decharme et al., 2019), MGB-IPH (Pontes et al., 2017) of CaMa-Flood (Yamazaki et al., 2011); and finer hydraulic models such as DassFlow-1D (Brisset et al., 2018), Mascaret (Goutal et al., 2012) and Telemac-2D (Zaoui et al., 2018). Those models need friction parameters to run and such parameters are issued from empirical laws and/or reference tables. But, above all, such parameters require a calibration step, prior to any application, to fix their value such that the model is able to reproduce an observed event. Other recurring applications use Data Assimilation (DA) to constrain input parameters,



such as roughness parameters, alone or along with the model state variables, such as discharge and water surface elevations. The drawback of the above approaches is that the roughness coefficients become model-dependent and can not be used outside
60 of the scope of the model.

Nevertheless, it is still possible to infer physically-based friction properties by inversion of the river flow velocity profiles. Number of studies based on this concept can be found in the literature. Working with a very controlled experiment, under completely known vegetation and flow velocities and/or discharges conditions, they assert the value of friction parameters for different types of vegetation, i.e. submerged (Jarvela, 2005; Ebrahimi et al., 2008; Pu et al., 2019) or not (Nepf, 1999;
65 Fathi-Moghadam et al., 2011), flexible (Jarvela, 2005; Pu et al., 2019) or rigid (Tanino and Nepf, 2008; Hamidifar et al., 2020) and even wood (Hygelund and Manga, 2003), in either the river channel (Kean and Smith, 2006; Ebrahimi et al., 2008; Fathi-Moghadam et al., 2011) or floodplains (Galema, 2009; Jung et al., 2011; Whittaker, 2014; Walczak et al., 2015). Such environments are difficult to reproduce outside of the experiment scope while also relying on fine-scale parameters that are not obtainable with remote-sensing techniques. Even Chaulagain (2018) tried to derive friction laws using remote-sensing airborne
70 LiDAR data but such products are unavailable at global-scale. In a more general way, these methods are hardly applicable to remote-sensing since there is no acceptable methodology to estimate flow velocities via remote-sensing.

From another perspective, founding works in hydrology established reference tables for the value of the friction coefficients given the soil composition, the surrounding vegetation, the shape of the river bed, the nature of the flow (Chow, 1964; Arcement and Schneider, 1989); and are still widely used. Most of this information can be found in global-scale products issued from
75 automatic chains using remotely-sensed data. Among them, the processing chain IOTA² (Inglada et al., 2017) allows the fully automatic generation of land cover maps at country scale. Using any existing soil-related databases in a supervised classification context, high resolution satellite optical images can be treated to produce corresponding land cover maps. Such global-scale information could be derived into global-scale friction information.

Therefore, the objective of the present study is to introduce an automatic methodology to derive physically-based, model-
80 free prior values of friction coefficients, using, among others, land cover maps produced with the IOTA² chain. Such products would therefore be usable globally. First, we describe the whole processing chain and emphasize on how it can be exported on any study domain given some reference data to train the IOTA² chain. Then, the quality of such priors is first compared to friction coefficients obtained from model calibration, via hydraulic simulations using the 1D-hydraulic Mascaret model (Goutal et al., 2012) applied over a portion of the upstream Garonne river in France and a portion of the Po river in Italy. Subsequently, we
85 validate our products through two application cases: a real-data DA experiment around the Mascaret model applied to the Po river domain and a discharge inference experiment with the HiVDI SWOT-discharge algorithm (Larnier et al., 2021) over both the Garonne and the Po, along with an additional application domain – namely a portion of the Brahmaputra river – to test our approach over another, less-gauged domain.



2 Methods and Materials

90 The present study aims at automatically derive friction parameters such as the Manning coefficient n or Strickler coefficient $K_s = \frac{1}{n}$ used for empirical estimation of free-surface flow discharge based on the Manning formula (Manning, 1891):

$$Q = \frac{1}{n} AR_h^{\frac{2}{3}} I^{\frac{1}{2}} = K_s AR_h^{\frac{2}{3}} I^{\frac{1}{2}}, \quad (1)$$

where Q is the free-surface flow discharge, A is the wetted area, R_h is the hydraulic radius and I the linear hydraulic head loss.

2.1 Method

95 2.1.1 Principle

Manning's n roughness coefficients depend on various factors, including among others, the shape of the channel or floodplain, the type of materials that compose the ground, the level of obstruction, etc. (Cowan, 1956) suggested a formula to determine the value of n in both the river channel and the surrounding floodplain. This formula assumes a base value for n (n_b) which is adjusted by several coefficients to account for different channel's properties:

$$100 \quad n = (n_b + n_1 + n_2 + n_3 + n_4) \times m, \quad (2)$$

where n_b [$\text{s.m}^{-\frac{1}{3}}$] is a base value of n for a straight, uniform and smooth channel in natural materials. Then, n_1 [$\text{s.m}^{-\frac{1}{3}}$] is a correction coefficient that accounts for surface irregularities, n_2 [$\text{s.m}^{-\frac{1}{3}}$] is a value for variations in shape and size of the channel cross-section, n_3 [$\text{s.m}^{-\frac{1}{3}}$] represents the effects of obstructions, n_4 [$\text{s.m}^{-\frac{1}{3}}$] is a value for vegetation and flow conditions, and m [-] is the meandering ratio of the channel.

105 Arcement and Schneider (1989) developed a practical guide for selecting Manning's n roughness coefficients in both the river channel and the adjacent floodplain following the Cowan's formula. For both the channel and the floodplain, the Guide presents a series of tables that map characteristics of the rivers with values of each parameter for the Cowan's formula. Then, it suggests a range of adjustment values given some qualitative conditions (e.g. "severe irregularity" or "minor obstruction"). While, in some cases, numerical values are provided to distinguish between two different classes (e.g. "minor obstruction in
110 a floodplain can be assumed if the obstructions occupy less than 15% of the total cross-sectional area"), the classification is mostly based on observational data (what an operator would actually observe on field).

The method developed in this paper adapts this procedure. Its novelty lies in relying on information from remote global datasets rather than on local observational data to compute each indicator. The approach allows to compute each element of the Cowan's formula and therefore estimate the value of n . While land cover maps are extensively used to derive most of those
115 terms, other sources of data are also used when more appropriate.

The general output format for our roughness coefficients would be 2-dimensional maps for roughness coefficients in the floodplain and 1-dimensionnal along-centerline data point for roughness coefficients in the main channel. Indeed, while only raster-based products are used to derive roughness coefficients in the floodplain, vector-based products are also employed to



Dataset	Availability		Reference
	Product available	Processing chain	
Land cover maps from IOTA ² chain	Yes, France only	Yes	Inglada et al. (2017)
SoilGrids	Yes	–	Poggio et al. (2019)
GRWL	Yes	–	Allen and Pavelsky (2018)
WorldDatabaseRivers	Yes	–	Frasson et al. (2019)

Table 1. Detail of global-scale datasets used to derive roughness.

120 derive roughness coefficients in the main channel. The next section details the global-scale remotely-sensed dataset used in our methodology.

2.1.2 Global remotely-sensed products used to derive friction coefficients

Table 1 sums up all global-scale products used for our approach and their availability.

125 **Land cover maps** The processing chain IOTA² (Inglada et al., 2017) allows the fully automatic generation of land cover maps at country scale and high resolution. The chain relies on satellite optical imagery (e.g. Landsat-8, Sentinel-2) and on existing land cover databases (e.g. Corine Land Cover, Randolph Glacier Inventory). Satellite imagery is collected over a complete year and pre-processed with temporal gapfilling and resampling techniques. Three spectral indices (NDVI, NDWI, brightness) are computed at every date to build the feature vector. The land cover databases are cleansed, filtered and merged to generate a reference dataset, that is defined with the targeted number of land cover classes. A first part of the reference dataset is used to train a supervised classifier, namely a Random Forest, while the second part is used to validate the land cover maps obtained with the trained classifier. The chain has been successfully applied over France to generate land cover maps for multiple years already.

Given the reference database used to train the IOTA² chain, the land cover map nomenclature will be different. Our methodology relies on 8 broader classes detailed in Table 2. In future applications, land cover maps should be aggregated to this reduced nomenclature to be later used in our chain.

135 **SoilGrids** (Poggio et al., 2019) is a system for global soil mapping. SoilGrids250m, its main product, provides, for 6 different depth-levels, global maps of soil properties at a resolution of 250 m. It allows to access at the pixel-level an estimation of the fraction of silt, sand and clay that constitutes the soil.

The Global River Widths from Landsat or GRWL (Allen and Pavelsky, 2018) was built to characterize the global coverage of rivers and streams wider than 30 m. One of the products offers a vector-based dataset of river centerlines, each segment being characterized by several properties including an estimation of the river width at the mean annual discharge.



Class key	Broad land cover classes
1	Artificial surfaces
2	High-height agricultural areas
3	Low-height agricultural areas
4	Forest
5	Mixed natural areas
6	Natural bare ground
7	Glaciers and perpetual snow
8	Water bodies

Table 2. Aggregated land cover classes considered in our methodology.

Indicator: SoilGrids soil type	Bed material (range/value [$\text{s.m}^{-\frac{1}{3}}$])	Adopted value [$\text{s.m}^{-\frac{1}{3}}$]
Sand	Coarse sand (0.026 - 0.035)	0.0310
Clay	Firm soil (0.020 - 0.025)	0.0225
Silt	Silty soil (0.020)	0.0200

Table 3. Indicators and values of n_b (bed material) for both the main channel and the floodplain.

GlobalDatasetForRivers (Frasson et al., 2019) took advantage of GRWL and the Shuttle Radar Topography Mission (SRTM) digital elevation model to create a global database of river width, slope, catchment area, meander wavelength, sinuosity, and discharge and is also distributed as a vector-based dataset of river centerlines.

The next sections detail how the above products are used to derive each term of the Cowan formula.

145 2.1.3 Computation of n_b

n_b represents the friction only due to the surface onto which water flows, for both the channel and the floodplain. Arcement and Schneider (1989) directly relates the value of n_b to the materials that constitute the soil surface or soil type, e.g. sand, concrete, firm soil, gravel, cobble, etc. Since *SoilGrids250m* provides the fraction of sand, silt and clay that composes the ground at the surface level, n_b is derived from this database. The value of n_b for each type of soil are then given in Table 3, taken from

150 Arcement and Schneider (1989).

The value of n_b at any given location is computed as the weighted average of the adopted values of n_b given the fraction of corresponding soil type. There are a few locations where SoilGrids provide *no data*. In this case n_b is computed as the average of the three adopted values of n_b values, equal to $0.0245 \text{ s.m}^{-\frac{1}{3}}$.

2.1.4 Computation of n_1

155 n_1 represents the effect of surface irregularities.



Indicator: Land cover map from IOTA ²	Degree of irregularity (range/value [s.m ^{-1/3}])	Adopted value [s.m ^{-1/3}]
Artificial surfaces	Smooth (0.000)	0.000
High height agricultural areas	Moderate (0.006-0.010)	0.008
Low height agricultural areas	Minor (0.001-0.005)	0.003
Forests	Severe (0.011-0.020)	0.016
Mixed natural areas	Moderate (0.006-0.010)	0.008
Natural bare ground	Smooth (0.000)	0.000
Glaciers and perpetual snow	Smooth (0.000)	0.000

Table 4. Indicators and values of n_1 (degree of irregularity) for the floodplain.

Indicator: L_{smooth}/L_{raw}	Degree of irregularity (range/value [s.m ^{-1/3}])	Adopted value [s.m ^{-1/3}]
$ 1 - \text{Indicator} < 0.005$	Smooth (0.000)	0.000
$0.005 < 1 - \text{Indicator} < 0.01$	Minor (0.001-0.005)	0.003
$0.01 < 1 - \text{Indicator} < 0.015$	Moderate (0.006-0.010)	0.008
$(\text{Indicator} - 1) > 0.015$	Severe (0.011-0.020)	0.016

Table 5. Indicators and values of n_1 (degree of irregularity) for the main channel

For the floodplain it is computed based on the land cover type. Arcement and Schneider (1989) relates the value of n_1 to the degree of irregularity of the floodplain surface specified as smooth, minor, moderate or severe. Table 4 shows how these four categories are mapped to the aggregated land cover classes derived from the IOTA² chain. Note that the mapping between Arcement and Schneider (1989) categories and the land cover classes suggested here (between column 1 and 2 of Table 4) was set according to our hydrology expertise.

Note that the "water bodies" class from Table 2 is ignored here as we are interested in the surface below the water to derive roughness. Therefore, the nomenclature from Table 4 can not be used for the main channel.

For the main channel, the method relies on cross-sectionnal profiles information. To evaluate the irregularity of the cross-sectionnal shape, each cross-sectionnal profile is first smoothed using an univariate cubic spline between the edges of the profile. Illustrations of the smoothing procedure are given in Appendix B. Then, a ratio between the original cross-section length and the smoothed profile length is calculated. Finally, Table 5 shows how this ratio is mapped against the categories suggested by Arcement and Schneider (1989) to estimate the value of n_1 in the main channel. Note that, currently, the threshold values were picked to match our knowledge of the study domains.

2.1.5 Computation of n_2

n_2 represents the effect of longitudinal variations in cross-section shape and size in the main channel. Note that n_2 is assumed to be equal to $0.0 \text{ s.m}^{-1/3}$ in the floodplain as suggested in Arcement and Schneider (1989).



Indicator: $\sigma \left(\frac{\partial W}{\partial x_s} \right)$	Variations in channel cross-section (range/value [$\text{s.m}^{-\frac{1}{3}}$])	Adopted value [$\text{s.m}^{-\frac{1}{3}}$]
Indicator < 0.025	Gradual (0.000)	0.000
0.025 < Indicator < 0.05	Alternating occasionally (0.001-0.005)	0.003
Indicator > 0.05	Alternating frequently (0.010-0.015)	0.013

Table 6. Indicators and values of n_2 (variations in cross-section shape and size) for the main channel.

Indicator: Land cover map from IOTA ²	Effect of obstruction (range/value [$\text{s.m}^{-\frac{1}{3}}$])	Adopted value [$\text{s.m}^{-\frac{1}{3}}$]
Artificial surfaces	Appreciable (0.020-0.030)	0.025
High height agricultural areas	Minor (0.005-0.019)	0.012
Low height agricultural areas	Negligible (0.000-0.004)	0.002
Forests	Minor (0.005-0.019)	0.012
Mixed natural areas	Minor (0.005-0.019)	0.012
Natural bare ground	-	0.000
Glaciers and perpetual snow	-	0.000

Table 7. Indicators and values of n_3 (degree of irregularity) for the floodplain.

The value of n_2 in the channel is determined based on how frequently the channel shape and size alternates between large and small cross sections. The channel widths sequence, extracted from GRWL along the study domain, is used as a proxy to evaluate this feature. More specifically, the first derivative of the width signal is derived to measure the range of width variations. In the end, the value of n_2 is determined according to the standard deviation of the width first derivative value. Table 6 shows the categories created with this indicator that map with the categories suggested by Arcement and Schneider (1989) to estimate the value of n_2 . Additional illustrations of GRWL width variations and its first derivative along the study domains are given in Appendix B. Similarly to the thresholds used for n_1 , note that the threshold values are currently also picked according to match our knowledge of the study domains.

2.1.6 Computation of n_3

n_3 represents the effect of obstructions.

Since the scope of the proposed method is framed by the SWOT mission, its goal is to estimate the value of n for large rivers (width>30m) for which obstructions within the channel are rare and can be neglected. The value of n_3 is therefore assumed to be equal to $0.0 \text{ s.m}^{-\frac{1}{3}}$ in the channel.

The value of n_3 in the floodplain is derived from land use maps. Table 7 shows how the land cover nomenclature is associated with the categories suggested by Arcement and Schneider (1989) to estimate the value of n_3 .



Indicator: Land cover map from IOTA ²	Amount of vegetation (range/value [s.m ^{-1/3}])	Adopted value [s.m ^{-1/3}]
Artificial surfaces	Small (0.001-0.010)	0.006
High height agricultural areas	Large (0.025-0.050)	0.038
Low height agricultural areas	Small (0.001-0.010)	0.006
Forests	Very large (0.050-0.100)	0.075
Mixed natural areas	Medium (0.011-0.025)	0.018
Natural bare ground	-	0.000
Glaciers and perpetual snow	-	0.000

Table 8. Indicators and values of n_4 (amount of vegetation) for the floodplain.

Indicator: Sinuosity	Degree of meandering (value [-])	Adopted value [-]
Indicator < 1.2	Minor (1.00)	1.00
1.2 < Indicator < 1.5	Appreciable (1.15)	1.15
Indicator > 1.5	Severe (1.30)	1.30

Table 9. Indicators and values of m (degree of meandering) for the main channel.

2.1.7 Computation of n_4

n_4 represents the effect of vegetation.

As for the estimation of n_3 , the effect of vegetation in the main channel is assumed to be negligible and set to $0.0 \text{ s.m}^{-1/3}$.

190 The value of n_4 in the floodplain is derived from land use maps. Table 8 shows how the IOTA² classification can be mapped with the categories suggested by Arcement and Schneider (1989) to estimate the value of n_4 .

2.1.8 Computation of m

m is a coefficient that accounts for the effect of meandering. According to the Guide the value of m can be assumed to be equal to 1.0 in the floodplain.

195 The sinuosity parameter calculated globally by Frasson et al. (2019) is used as a proxy to infer the value of m in the main channel, see Table 9.

2.1.9 Method applicability at global scale

This method aims at being applicable globally. This may be possible as most of the required data are publicly available and have global coverage, see Table 1.

200 However we should first point out that the IOTA² products are currently only available over France. Thus generating land cover maps from the IOTA² chain on its own is mandatory for study areas outside France. Running the IOTA² chain is manageable but requires both training and availability of the required inputs. From our experiments (e.g. Po river, see Section 2.2.2



and Brahmaputra river, see Section 3.2), it may be difficult to find cloud-free Sentinel 2 images which are required to compute Land-cover maps using the IOTA² chain. Still, cloud-free Sentinel 2 images can be produced using, for example, the MAJA software (Baetens et al., 2019).
205

The methodology also requires cross-sectionnal profiles to compute n_1 in the main channel. Such profiles are generally obtained locally from field surveys. Still, one could reconstruct profiles down to the lowest observed water level from remotely-sensed products; by combining, for example, measurements of elevations – from nadir altimetry – simultaneously with watermasks – from optical imagery or even watermask processing chains such as SURFWater (Peña Luque et al., 2021) – from which river widths can be extracted, similarly to approaches used in Hostache et al. (2010); Desrochers et al. (2020). Others approaches are possible, see the application over the Brahmaputra river in Section 3.2
210

2.2 Study Domain and Data

Our method was first tested over two European study domains: a portion of the upstream Garonne river in France and a portion of the Po river in Italy.

2.2.1 Garonne domain 215

The study area consists in a 75-km-long reach of the Garonne river, from Toulouse to Castelsarrasin, see Figure 1. The Garonne river is the largest river in South-West of France, it drains an area of about 10,000 km² in Toulouse. The studied segment has an average slope of 0.9 m/km, the main channel width ranges from 100 to 300 m and the floodplain width can reach 4 km. The topography of the reach is described by a set of 145 cross-sections. Each cross-section is the compound product of a local topographic survey achieved during the early 2000s (main channel) and an extraction from the French national Digital Elevation Model (DEM) RGE ALTI (floodplain).
220

Discharge data has been collected over the year 2019 at multiple gauges, located either directly on the Garonne or on its main tributaries. Water level data has also been collected over the same time period at a single gauge located in the middle of the reach (Verdun-sur-Garonne).

Land cover maps of France obtained from the IOTA² chain are already freely available. The map latest version classifies land use in 23 different categories. This native nomenclature was deemed too high for developing the quasi-automatic procedure to estimate roughness coefficients. A new land cover map was therefore created by aggregating the original 23 classes down to the 8 macro-classes introduced in Tab 2. Table 10 shows how the original classes were mapped to the aggregated classes while Figure 2 displays the resulting land cover map.
225

The resulting roughness coefficients - defined as Strickler coefficients (the inverse of the Manning coefficient) - are illustrated in Figure 3 for the floodplain and in Figure 4 for the main channel. Moreover, general statistics on the Strickler coefficient values are displayed in Table C2 of Appendix C.
230



Land use (CES OSO version 2019)	Aggregated classes
Continuous urban fabric Discontinuous urban fabric Industrial and commercial units Road surfaces	Artificial surfaces
Rapeseed fields Small-grain fields Sunflower fields Soybean fields Corn fields Rice fields	High-height agricultural areas
Protein crops Tubers/roots Intensive grasslands Natural grasslands	Low-height agricultural areas
Orchards Vineyards Broad-leaved forests Coniferous forests	Forests
Woody moorlands	Mixed natural areas
Beaches and dunes Bare rock	Natural bare ground
Glaciers and perpetual snow	Glaciers and perpetual snow
Water bodies	Water bodies

Table 10. Mapping between the 23 original land use classes obtained from the IOTA² chain (CES OSO version 2019) and the 8 newly aggregated classes

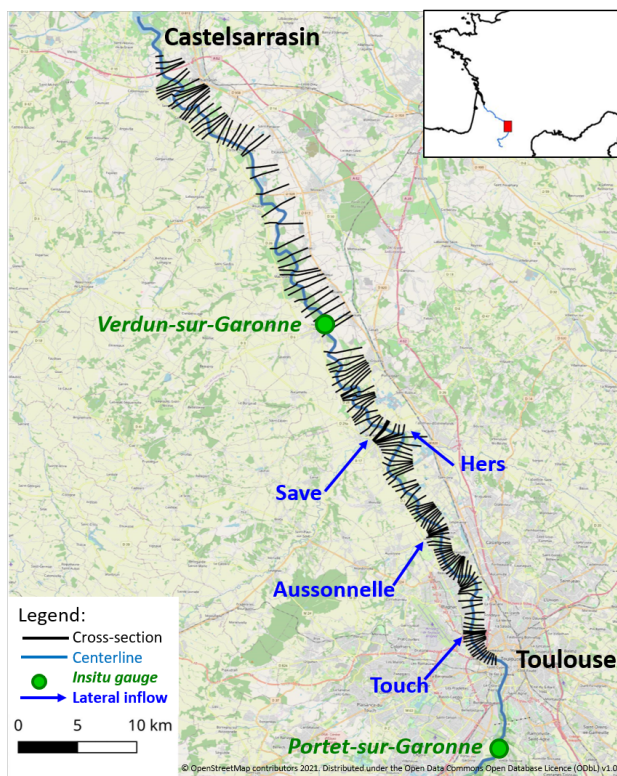


Figure 1. Map of the Garonne study domain.

2.2.2 Po domain

The study area consists of a 96-km-long reach, from Cremona to Borgoforte, see Figure 5. The Po river is the largest river in Italy, it drains an area of about 50,000 km² in Cremona. The studied segment has an average slope of 0.15 m/km and the main channel width ranges from 150 to 500 m. The floodplain is marked by the presence of two continuous dike systems, whose width ranges from 400 m to 4 km. The topography of the reach is described by a set of 91 cross-sections whose data points have been extracted from a 2005 2m-resolution DEM that was made available by the Po River Basin Authority (AdBPo). This DEM was obtained by a fusion of Lidar data for the floodplain and underwater sonar and ground survey data for the main channel.

Discharge and water level data has been collected over the period 2016-2019 at multiple in situ gauges from <https://simc.arpae.it/dext3r/>. Water level data has also been collected over the same time period at three virtual stations located on or very close to the Po river. This data has been downloaded from the online platform Hydroweb that makes available water level time series of continental waters (large lakes, reservoirs, major rivers) obtained after processing nadir altimetry measurements.

Land cover maps generated by the IOTA² chain were not available for this study domain. The chain was therefore directly run to create a land cover map, using Corine Land Cover (CLC) 2018 as reference data for training. CLC was however not used

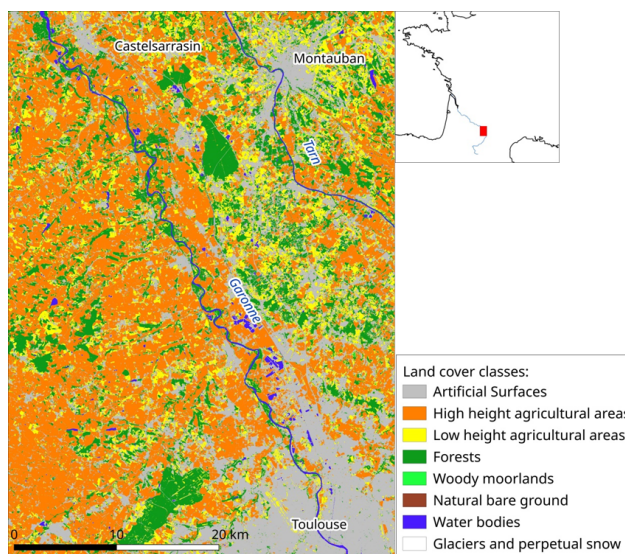


Figure 2. Land cover map from the IOTA² chain with 8 classes over the Garonne domain.

as is, its classes were aggregated into the 8 meta-classes described in Table 2 (see the details of the aggregation in Appendix A). This resulted in the creation of a land cover map consistent with the one available for the Garonne study domain, see Figure 6.

The resulting roughness coefficients - defined as Strickler coefficients (the inverse of the Manning coefficient) - are illustrated in Figure 7 for the floodplain and in Figure 8 for the main channel. Moreover, general statistics on the Strickler coefficient values are displayed in Table C3 of Appendix C.

2.3 Validation

2.3.1 Approach

The roughness coefficient of a river portion is not a quantity that can be measured on the field. Because of that, the results generated by the method developed in this paper cannot be compared to any existing dataset. The method, however, generates roughness coefficients that can be used as input to design a hydraulic model, whose river-related outputs can be compared to observed data.

The approach adopted to validate the method consists of two steps. First, a hydraulic model is required over a study domain whose characteristics are within the scope of SWOT. This model, either constructed for the sake of validation or obtained from another source, serves as a reference model. Its roughness coefficients must be set after a calibration procedure, whose complexity should be limited to reflect the global scope of the method. The results obtained from this model are compared to observed data, which allows to compute a reference level of performance achievable over the study domain, using metrics such as the Nash-Sutcliffe Efficiency (NSE). Second, a hydraulic model is constructed based on the roughness coefficients estimated

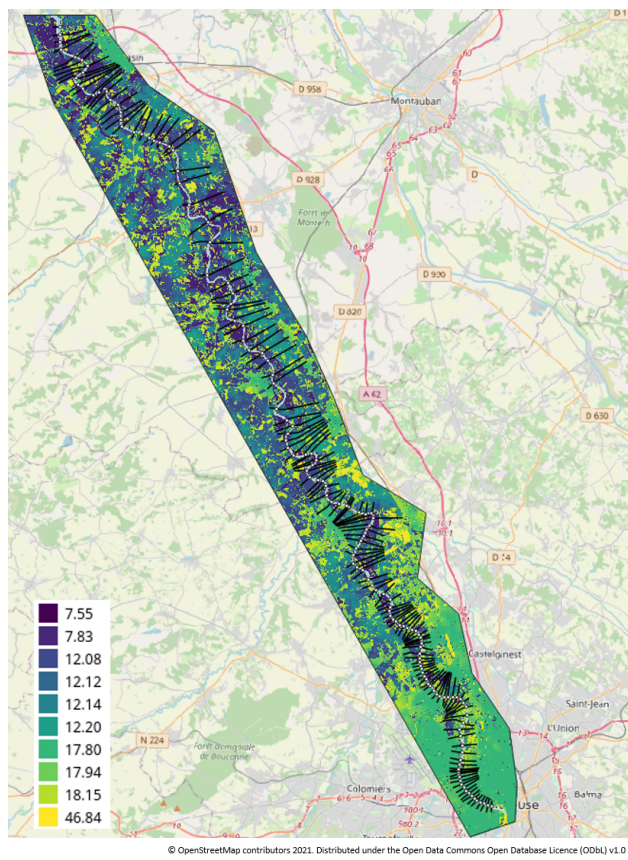


Figure 3. Map of inferred Strickler coefficients over the Garonne domain floodplain. The legend is set to mark down roughness decile values.

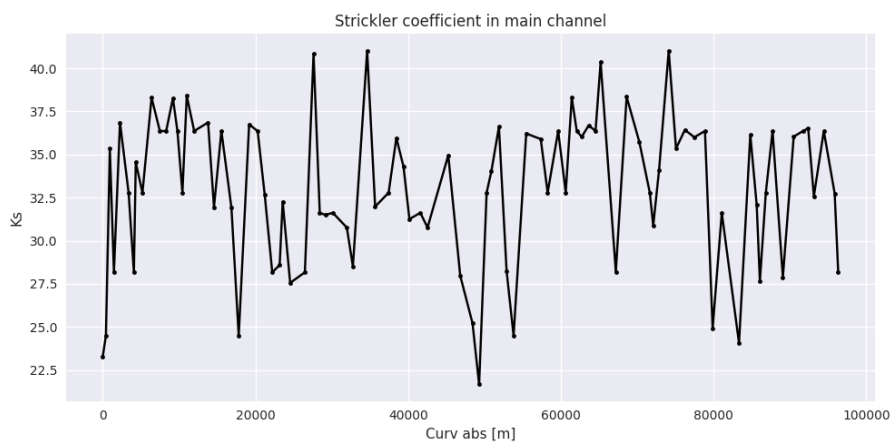


Figure 4. Curve of inferred Strickler coefficients over the Garonne domain along the main channel.

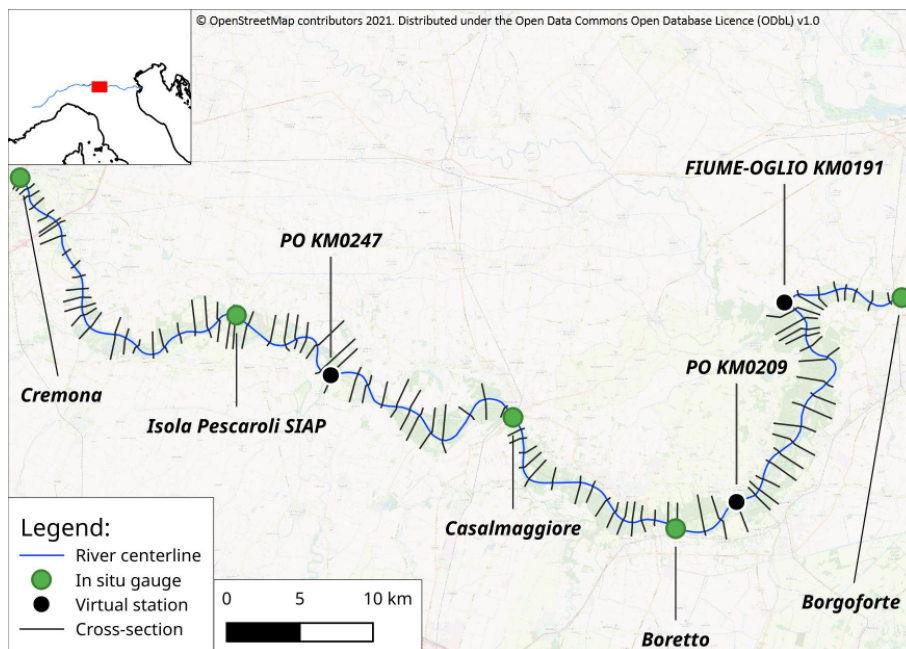


Figure 5. Map of the Po study domain.

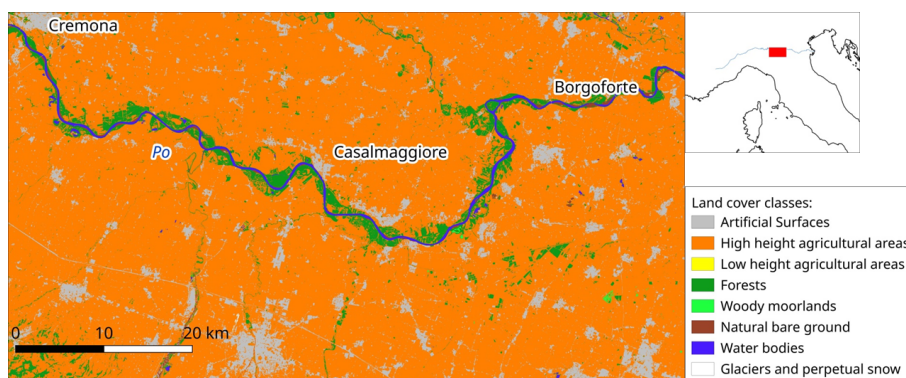


Figure 6. Land cover map from the IOTA² chain with 8 classes over the Po domain.

from the method developed in this paper, this model is referred to as the estimation model. Its results are compared to the same
 265 observed data and the method is deemed validated if the estimation performance is close to the reference performance.

The method described previously is first applied to the two study segments of the Garonne and Po rivers. For the channel n_1 ,
 n_2 and m are directly determined for each cross-section. The values of n_b , n_1 , n_3 and n_4 for the floodplain, and n_b , n_3 and n_4
 for the channel, are determined at each cross-section vertex, and then aggregated over their length with a weighted average to
 obtain a pair of values per cross-section. The Cowan's formula (2) is finally used to calculate the value of n per cross-section
 270 for both the channel and the floodplain.

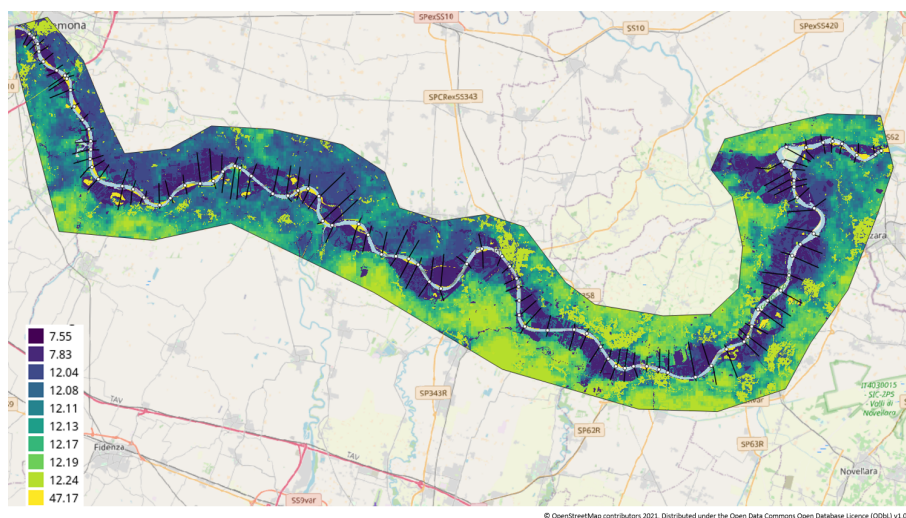


Figure 7. Map of inferred Strickler coefficients over the Po domain floodplain. The legend is set to mark down roughness decile values.



Figure 8. Curve of inferred Strickler coefficients over the Po domain along the main channel.

2.3.2 Models implementation over the study domains

1D hydraulic models for both the Garonne and Po study domains are constructed with Mascaret, set up to solve the shallow water equations with the finite difference Preissman scheme.

For the Garonne river, the upstream boundary condition is the discharge time series collected from the gauge Portet-sur-
275 Garonne (only 10 km upstream of the actual model boundary). The downstream boundary is a rating curve built on the assumption of a normal flow. Input discharge from the 4 largest tributaries of the reach are integrated as point source discharge. The collected cross-sections are used to represent the river topography. For the reference model, the roughness coefficient was

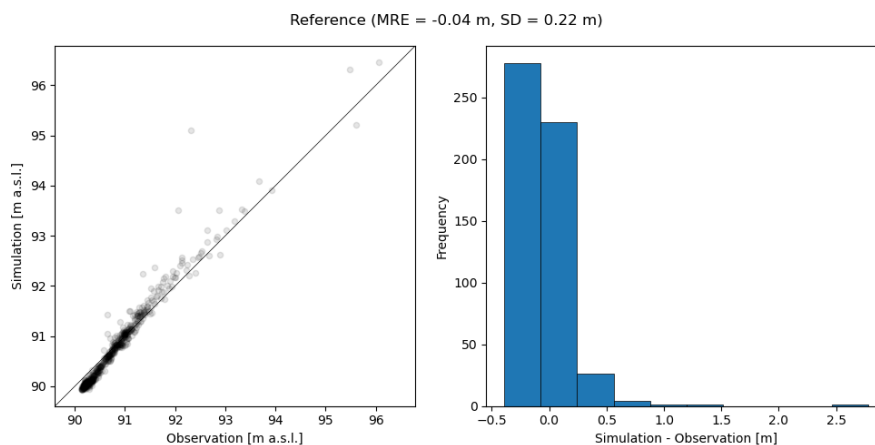


Figure 9. Comparison of the simulated and observed water levels for the reference model of the Garonne at Verdun-sur-Garonne.

assumed to be spatially homogeneous within the reach while differentiated between the main channel and the floodplain. The model was calibrated with data collected in 2019. The calibration procedure consisted in finding the pair of main channel and floodplain roughness coefficients that optimize the NSE value computed at the in situ gauge Verdun-sur-Garonne. The calibration led to set the main channel Strickler coefficient to $24 \text{ m}^{1/3} \cdot \text{s}^{-1}$ and the floodplain one to $16 \text{ m}^{1/3} \cdot \text{s}^{-1}$, corresponding to a NSE of 0.90.

For the Po river, the upstream boundary condition is set to the discharge time series observed at Cremona and the downstream condition to the water level time series observed at Borgoforte. Input discharge from the tributaries are considered as negligible. The cross-sections that were manually delineated are used to represent the river topography, their width has been limited to the first system of dikes to prevent the model from inundating areas normally protected behind this system. Again, for the reference model, the roughness coefficient was assumed to be spatially homogeneous within the reach and still differentiated between the main channel and the floodplain. The model was calibrated over the period 2016-2018, by finding the pair of main channel and floodplain roughness coefficients that optimizes the value of a mean NSE computed over water level time series at 4 in situ gauges (Cremona, Isola Pescaroli SIAP, Casalmaggiore, Boretto). The calibration procedure led to set the main channel Strickler coefficient to $34 \text{ m}^{1/3} \cdot \text{s}^{-1}$ and the floodplain one to $8 \text{ m}^{1/3} \cdot \text{s}^{-1}$, corresponding to a mean NSE of 0.94. It also showed that the model was far less sensitive to the floodplain roughness coefficient value than the main channel one.

2.3.3 Results

Simulations are conducted with the constructed models based on data from 2019. Their results are compared to water level time series obtained at in situ gauges. The following metrics are used for comparison: NSE, Mean Relative Error (MRE), Standard Deviation (SD), Root Mean Squared Error (RMSE) and Percentage Bias (PBIAS).

Figure 9 and Figure 10 show a comparison of the simulated and observed water levels for the reference and estimation models of the Garonne at Verdun-sur-Garonne. The MRE is close to 0 (-0.04) for the reference model, which is expected since

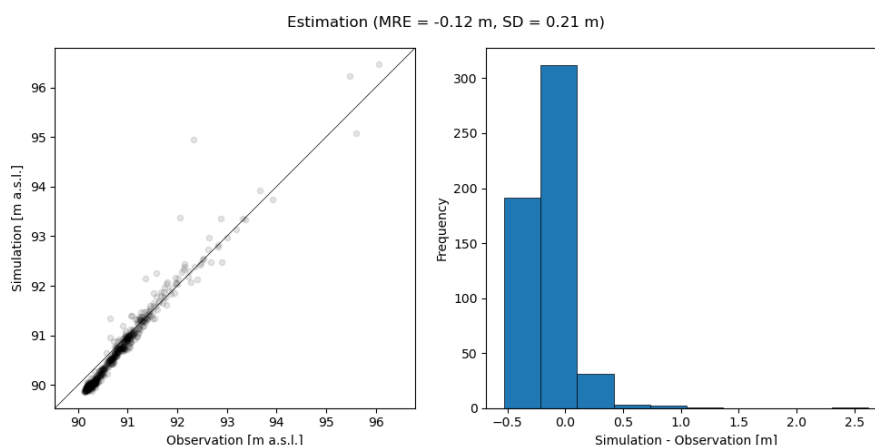


Figure 10. Comparison of the simulated and observed water levels for the estimation model of the Garonne at Verdun-sur-Garonne.

	RMSE		PBIAS		NSE	
	Reference	Estimation	Reference	Estimation	Reference	Estimation
Verdun-sur-Garonne	0.22	0.24	0.04	0.14	0.90	0.89

Table 11. Metrics of the simulated and observed water levels for the reference and the estimation model of the Garonne at the Verdun-sur-Garonne gauge.

it was calibrated to better fit the water level time series collected at Verdun-sur-Garonne. A bias of -0.12 m is obtained with the
 300 estimation model and the SD has slightly decreased, going from 0.22 m to 0.21 m.

Table 11 shows statistics obtained at Verdun-sur-Garonne. The bias previously noted is found again in the PBIAS value of 0.14% for the estimation model compared to a value of 0.04% for the reference model (a positive PBIAS indicates a model underestimation). Both the RMSE and the NSE are deteriorated, although just slightly. These statistics demonstrate that the results obtained with the estimation model are very close to those obtained with the reference model.

305 Similarly, Figures 11 and 12 show a comparison of the simulated and observed water levels for the reference model of the Po, aggregated among 4 gauges (Cremona, Isola Pescaroli SIAP, Casalmaggiore, Borretto). The MRE is close to 0 for the reference model, which is expected since it was calibrated to better fit these water level time series. A global bias of -0.33 m is obtained with the estimation model and the SD is slightly increased from 0.40 m to 0.48 m.

Table 12 shows the results obtained at each gauge location. The global bias already noted is found in each PBIAS value
 310 (a positive PBIAS indicates a model underestimation). There is however no clear pattern that emerges from the results, while the RMSE and NSE are deteriorated at the gauges Cremona, Casalmaggiore and Boretto, they are improved at the gauge Isola Pescaroli SIAP. The NSE values are all above 0.94 for the reference model, and stay above 0.84 for the estimation model. This shows that despite the deterioration observed between the reference and the estimation models, the results obtained with the estimation model still have a good performance level.

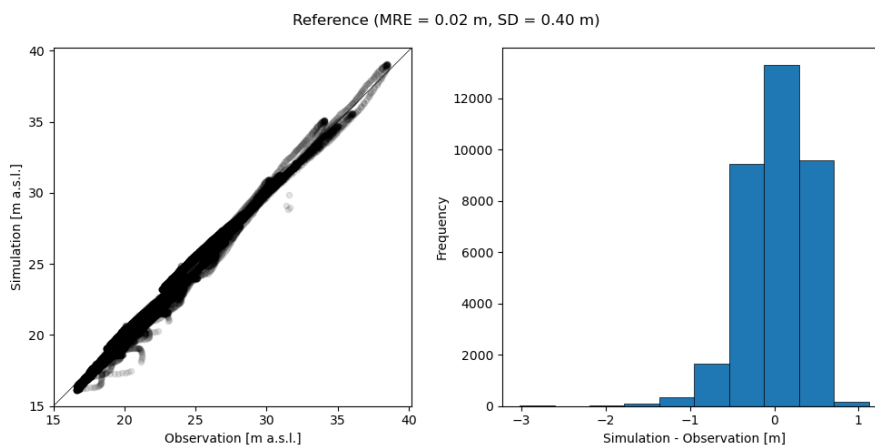


Figure 11. Comparison of the simulated and observed water levels for the reference model of the Po aggregated over 4 in situ gauges.

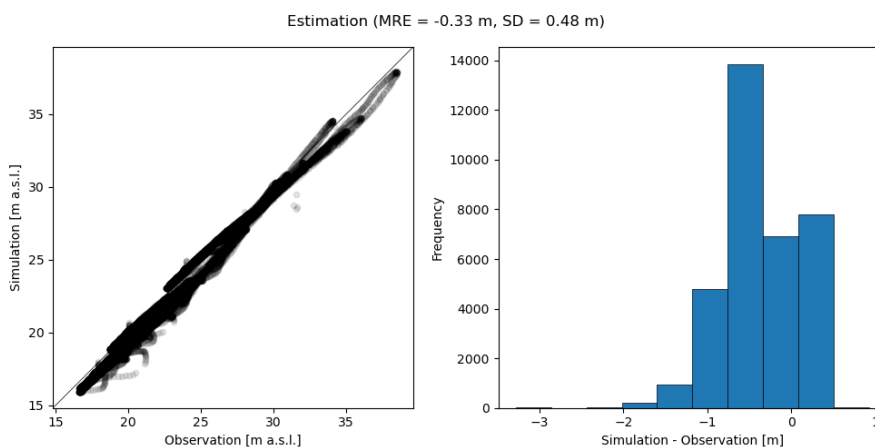


Figure 12. Comparison of the simulated and observed water levels for the estimation model of the Po aggregated over 4 in situ gauges.

	RMSE		PBIAS		NSE	
	Reference	Estimation	Reference	Estimation	Reference	Estimation
Cremona	0.13	0.59	0.26	1.97	1.00	0.91
Isla Pescaroli	0.52	0.32	-1.96	-0.97	0.94	0.98
Casalmaggiore	0.34	0.42	-0.53	0.88	0.97	0.96
Boretto	0.50	0.85	2.22	4.20	0.94	0.84

Table 12. Metrics of the simulated and observed water levels for the reference and the estimation model of the Po at 4 in situ gauges.



315 It is not surprising that our estimated model performs slightly worse than the calibrated model, as our approach to define roughness coefficient is purely physically-based and the calibrated model encompasses model structural errors to match the observations. Still, our estimated model performs fairly well and allows us to serenely use those roughness coefficients in our applications.

3 Applications

320 These new remote-data-fusion-based friction can serve various applications in hydrology. A first and obvious field of application would be to use these friction coefficients in one-dimensional shallow water models (such as HEC-RAS, Mascaret or DassFlow-1D among others). A second field of application could be fine-scale 2D shallow water models for surveys or research on flood prediction. Indeed our chain does not only provide friction coefficients in the main channel and in the near floodplain, but could also be extended to the entire floodplain using the 2D maps of IOTA² and ancillary data (SoilGrids, etc.). In this
325 case the values computed by our chain could be used as-is or would require a small tuning. One must recall that the friction coefficients in shallow water models encompass the modeling errors and so may vary from one model to another and also to take into account modeling differences between one-dimensional and two-dimensional models. But having a good starting point for a calibration reduces the computational cost.

Finally these coefficients could also be used as-is or as good first guess for DA in other models that are based on shallow
330 water equations or their simplification (Manning equation, Kinematic wave, etc.). For instance they could be used in the large scale model MGB-IPH (Diffusive wave equations) or in some of the SWOT discharge algorithms, such as MetroMan (Durand et al., 2014), geoBAM (Brinkerhoff et al., 2019) or HiVDI (Larnier et al., 2021). Here we present two applications. The first application is a DA experiment in the one-dimensional shallow water model Mascaret while the second application shows the benefits of using the computed friction coefficients as a first guess for the HiVDI SWOT discharge algorithm.

335 3.1 Data Assimilation experiment

A first case of application for the remote-data-fusion-based friction coefficients will consist in a Data Assimilation (DA) experiment around the Mascaret 1D-hydraulic model (Goutal et al., 2012).

DA consists in correcting parameters and/or state variables of a numerical physical model from external observations to improve its forecasting capabilities. DA can be tackled with two main approaches: "state" estimation or "parameter" estimation.
340 For the first configuration, model state variables are directly constrained and updated. For the second configuration, model inputs are corrected, i.e. parameters defining model's equations, e.g. Strickler coefficients, or even boundary conditions, e.g. upstream discharge forcing 1D-hydraulic model. When applied as parameter estimation, DA approaches can be related to a calibration procedure. In hydraulics, such applications are mostly focused on friction coefficients (Habert et al., 2016; Goery et al., 2017; Zaoui et al., 2018; Mirouze et al., 2020) as little is known on the value of these parameters.

345 The SMURF package – i.e. System for Modeling with Uncertainty Reduction, and Forecasting – is an open source Python modular system developed at CERFACS set for running and cycling DA systems. It is currently mainly used for hydraulic



applications, more precisely within the preparation to the SWOT mission. For more details, the reader should read Mirouze and Ricci (2021). Note that SMURF can orchestrate both kind of assimilation introduced above. Here, SMURF is used to orchestrate the application of the Ensemble Kalman filter (EnKF) DA algorithm around the Mascaret model.

350 The EnKF (Evensen, 2003) is a widely-spread extension of the classical Kalman filter (Kalman, 1960) developed for numerical models with nonlinear dynamics. Error covariance matrices are stochastically estimated directly from an ensemble of model runs so they do not rely on a linearity assumption. For more details on the EnKF implementation, please see Appendix D.

Here, we realized two DA experiments around the Po model over the period from Oct, 01 2019 to Dec, 31 2019, as this period covers three flooding events. Both experiment assimilates real-data from insitu and altimetry gauges. The difference
355 between the two relies on the friction coefficients used as first guess to initiate the experiment: first, the uniform calibrated values and second, the roughness from our methodology.

Water levels measurements from 4 insitu stations and 3 virtual stations from Sentinel-3A are assimilated. All observations were converted into equivalent water surface elevations referenced to the same geoid as the Mascaret model. The observation error for the EnKF is modeled as a white noise with a given standard deviation depending on the gauge type. A dispersion of 10
360 cm was taken for the insitu gauges and the averaged error provided by Hydroweb (<http://hydroweb.theia-land.fr/>) was used for the virtual gauges, see Table 13 for the detailed observation error value. Insitu measurements are delivered between every 15 or 30 minutes with possible missing data within the entire timeseries while virtual gauges have a revisit period of 27 days, with possibly missing data. We set an assimilation window of 1 day for the EnKF. During a given assimilation cycle, all available measurements for the current day are assimilated and compared to the closest simulated timestep.

365 For both experiments, the spatial distribution of Strickler coefficients in both the main channel – K_{smin} variable – and the floodplain – K_{smaj} variable – are simultaneously updated. Note that Mascaret does not differentiate a right and left floodplain so only one overall value of Strickler coefficient is required for the floodplain. Given the spatial coverage of the observations to be assimilated, the Strickler coefficients are set uniform over 7 zones delimited by the gauges position on the 1D domain. Therefore, the control vector to update through data assimilation contains 14 elements: 7 Strickler coefficients for the main
370 channel and 7 Strickler coefficients for the floodplain. The control error is set as a uniform law which range is initialized as $\pm 30\%$ of the prior value in the river channel and as $\pm 50\%$ of the prior value for the floodplain. Table 13 sums up all control and observation variables configuration for both experiments.

DA performances for the real-data experiment can only be evaluated at the observation gauges, see Table 14 and Figures 13-14. Overall, the second experiment performs slightly better than the first one, which validates more our inferred roughness
375 coefficient. After the assimilation in the first experiment, water level RMSE are little or not reduced compared to the free run while the contribution of the DA is clearer with the second experiment. Moreover, in the second experiment, DA displays interesting results with mostly improved RMSE but decreasing enhancement the more downstream the observation gauge. This could be easily explained by the downstream control imposed by the model where a water level timeseries is used as downstream boundary condition and constrains simulated water surface elevations including the two most downstream gauges.

380 The physically-based roughness coefficients are here used as initial prior values for DA experiments. It is true that part of the fine-resolution information was lost by aggregating over 7 larger areas but this decision was based on the availability of



Zone	Obs error [m]	Ksmin			Ksmaj		
		bounds	Exp #1	Exp #2	bounds	Exp #1	Exp #2
1	0.1	25.0-60.0	34.0	40.546	2.0-25.0	8.0	9.833
2	0.1	25.0-60.0	34.0	36.478	2.0-25.0	8.0	10.172
3	0.18	25.0-60.0	34.0	36.085	2.0-25.0	8.0	9.598
4	0.1	25.0-60.0	34.0	39.04	2.0-25.0	8.0	9.693
5	0.1	25.0-60.0	34.0	35.136	2.0-25.0	8.0	10.088
6	0.13	25.0-60.0	34.0	41.702	2.0-25.0	8.0	10.067
7	0.36	25.0-60.0	34.0	39.616	2.0-25.0	8.0	9.408

Table 13. For each zone ordered from upstream to downstream along the Po model (rows), observation error standard deviation for the associated gauge located at the upstream on the zone (column 2); allowed values range for Ksmin (column 3); first experiment initial mean prior value for Ksmin (column 4); second experiment initial mean prior value for Ksmin (column 5); allowed values range for Ksmaj (column 6); first experiment initial mean prior value for Ksmaj (column 7); second experiment initial mean prior value for Ksmaj (column 8).

Zone	RMSE Z [m]	
	Exp #1	Exp #2
	calibrated roughness	estimated roughness
1	0.199 / 0.152	0.687 / 0.150
2	0.442 / 0.230	0.303 / 0.213
3	0.819 / 0.518	0.692 / 0.256
4	0.451 / 0.243	0.836 / 0.237
5	0.616 / 0.262	1.069 / 0.259
6	0.225 / 0.521	0.239 / 0.442
7	0.008 / 0.088	0.223 / 0.174

Table 14. Real-data experiment results before (left-hand side of the slash character) and after (right-hand side of the slash character) assimilation per zone. Performances are evaluated by RMSE for Z at observation gauges: free run RMSE (left-hand side of the slash character) are compared to analysis run RMSE (right-hand side of the slash character).

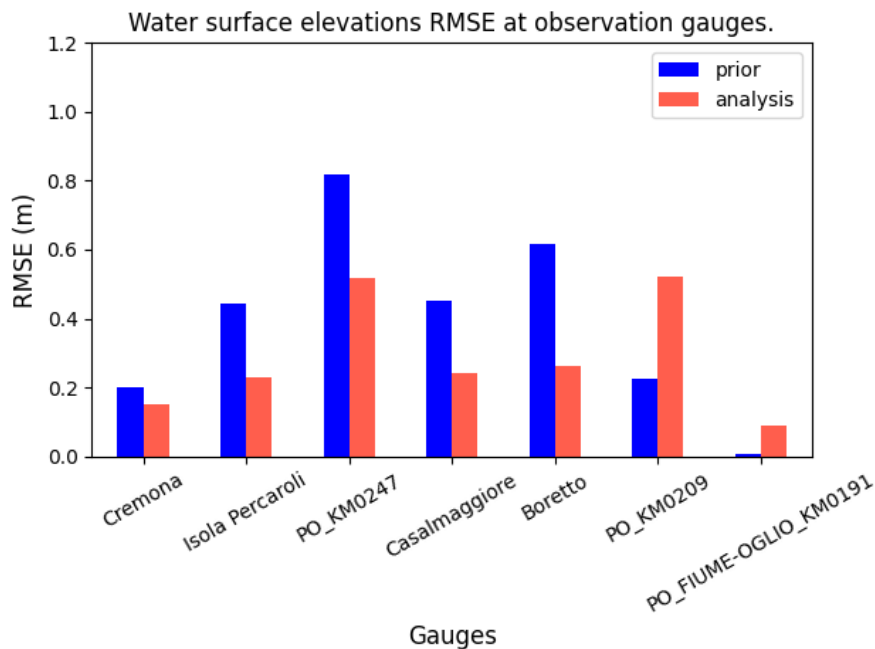


Figure 13. Model performances over simulated water surface elevation Z at observation gauges before (blue "prior" bars) and after (red "analysis" bars) assimilation for the Experiment #1 that uses uniform calibrated roughness coefficients.

the observations, as it would be complicated to assess the quality of updated roughness coefficient over unobserved part of the study domain. Better performances could be surely obtained by better tuning and setting the DA framework but such setup is out of the scope of the present study. Despite these conditions, the physically-based roughness coefficients present as
385 interesting priors parameters to initialize such DA experiments.

3.2 SWOT Discharge Algorithm

A second case of application will consist in assessing the benefits of using the remote-data-fusion-based friction coefficients as a first guess for the assimilation of synthetic SWOT observations in the HiVDI discharge algorithm (Larnier et al., 2021; Larnier and Monnier, 2020). The performance of such applications will be compared to the current implementation that uses a rough
390 estimation of friction coefficients from the river width. As SWOT has not been launched yet, the discharge algorithms have been benchmarked so far on synthetic SWOT observations, namely the PEPSI challenges' datasets (Durand et al., 2016; Frasson et al., 2021). These datasets consist in synthetic SWOT observations of river portions with various flow regimes, computed using calibrated shallow water models. Some of the algorithms were also tested on observations from the SWOT Science simulator that simulates the full error budget of the Karin instrument or from AirSWOT data (airborne similar instrument).

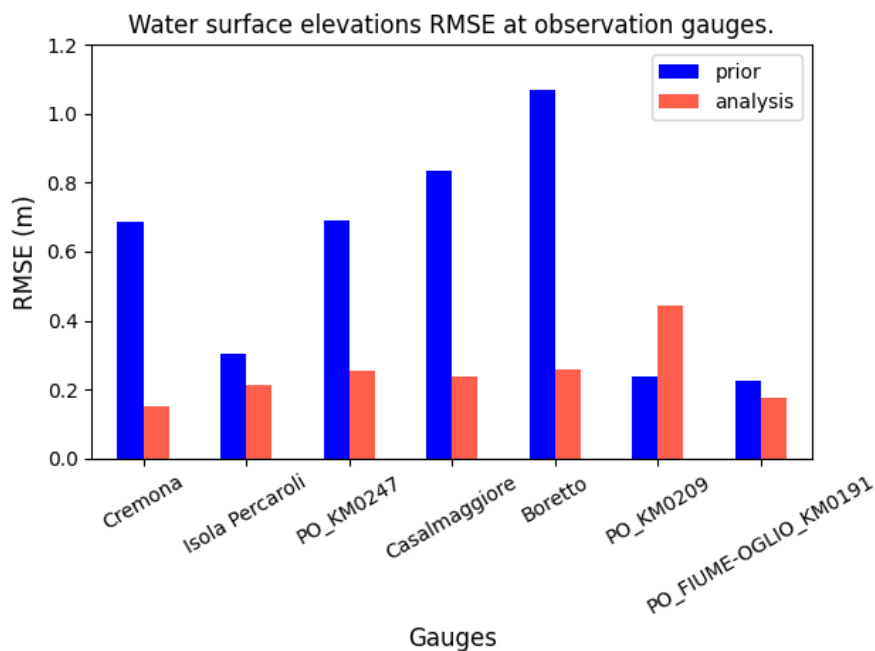


Figure 14. Model performances over simulated water surface elevation Z at observation gauges before (blue "prior" bars) and after (red "analysis" bars) assimilation for the Experiment #2 that uses roughness coefficients from our method.

395 The analyses of these benchmarks have highlighted the strong dependency on the accuracy of the estimated discharges to the priors, namely bathymetry, friction coefficients and mean discharge value. So it is mandatory to find new solutions to estimate good priors. The application of our chain to estimate the priors of friction coefficients is in line with this objective.

Usually the first guess of friction coefficients are computed from the mean width of the considered river portion. It may vary from 20 (small rivers) to 40 (very large rivers). This rough estimation plus the fact that it provides a uniform value for the first
400 guess can lead to wrong estimations of discharge after the assimilation of the synthetic SWOT observations.

Here we realized the comparison on three cases of the PEPSI challenges. The first two cases are the Garonne river and the Po river cases. These cases were selected as the friction coefficients for these rivers have already been estimated in this study. The third case is the Brahmaputra case (India). This last case has been selected based on the availability of input data for running the IOTA² chain. For this case, the Copernicus land-cover at 100m resolution (Buchhorn et al., 2020) was used for the
405 calibration of the IOTA² classifier.

The PEPSI dataset does not contain any information about the separation between main channel and the floodplain. But as the objective of these datasets was to benchmark algorithms on "nominal conditions", namely natural rivers with no or very little flooding, we conducted our experiments under the assumption that the flow is restricted to the main channel. Thus we only used friction coefficients estimated for the main channel. Moreover, no cross-sectional profiles were previously available. As
410 introduced in Section 2.1.9, we combined simultaneous observations of width and water surface elevations, provided by the



Table 15. Statistics on the estimated values of friction coefficients for the three PEPSI cases.

Case	This method ($m^{1/3} \cdot s^{-1}$)			Former Value ($m^{1/3} \cdot s^{-1}$)
	Min	Mean	Max	
Garonne river	22.24	25.53	27.05	25.0
Po river	34.28	38.96	42.37	30.0
Brahmaputra river	40.25	40.90	41.85	50.0

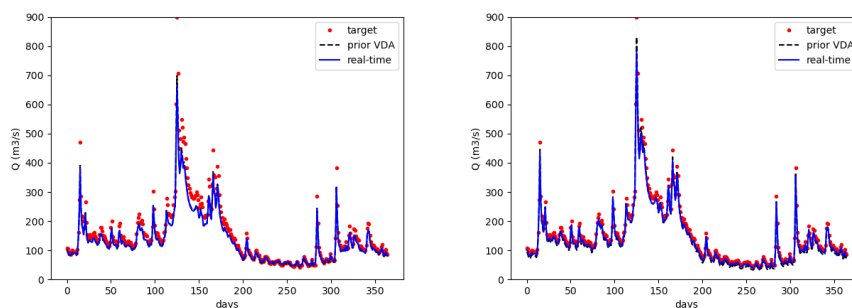


Figure 15. Estimated discharge for the Garonne river, using the former constant friction coefficient (left) and the new method (right)

PEPSI dataset, to reconstruct each cross-sectionnal profile individual shape. Using a handcrafted routine, we then draw the profiles along the riverline such that they remain quasi-orthogonal to the riverline while never crossing each other.

For the Garonne River and Po river cases, friction coefficients already computed in our study were spatially interpolated using curvilinear abscissae in the PEPSI 1 datasets. For the Brahmaputra river, the friction coefficients were directly computed at the location of the cross-sections in the PEPSI 2 dataset using our new method. The statistics on the values estimated using our method are listed in Table 15 with comparison to the constant values formerly estimated.

The results obtained for the three cases with comparison with previous results are displayed in Figure 15 for the Garonne river, in Figure 16 for the Po river and in Figure 17 for the Brahmaputra river. Two usual metrics for the benchmarks of SWOT discharge algorithms were computed on these results, namely the Nash-Sutcliffe Efficiency (NSE) and the Normalized Root Mean Squared Error (NRMSE). The values obtained are listed in Table 16.

These results show that using the friction coefficients estimated by our method as a prior looks very promising. Indeed for the Garonne and Po rivers cases the results outperform those obtained with the former method with a 47 % reduction of errors (in terms of NRMSE) for the Garonne river and 13 % for the Po River. For the Brahmaputra river, the results obtained with the new method slightly underperform those obtained with the former method with an 3 % increase of errors. However the results are comparable. These results show the potential benefits of this new method to better constrain the SWOT discharge algorithms.

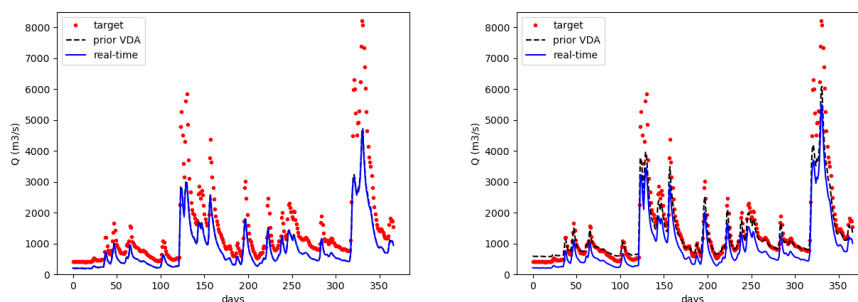


Figure 16. Estimated discharge for the Po river, using the former constant friction coefficient (left) and the new method (right)

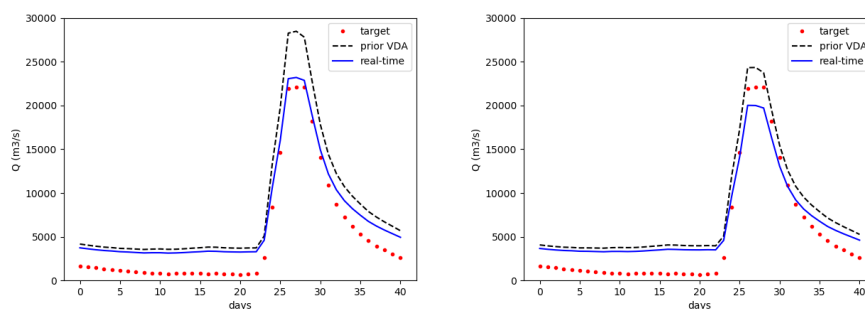


Figure 17. Estimated discharge for the Brahmaputra river, using the former constant friction coefficient (left) and the new method (right)

4 Conclusions

The present study introduces an automatic methodology to derive physically-based roughness coefficients for hydrology applications. The methodology is applied via a processing chain based on the Cowan's formula and can infer roughness coefficients in both the floodplain and the riverbed at different format. A two-dimensional raster image gathers the roughness coefficient in the floodplain while a node-based shapefile contains the roughness coefficients along the main channel centerline. The interesting aspect of such a chain is that it only uses remotely-sensed, publically available datasets and/or existing pre-processing chains. It can therefore be applied at global-scale for a wide range of applications.

Among the input data, land cover maps generated by the IOTA² chain are at the core of the methodology to derive roughness coefficients in the floodplain. Such maps are currently only available yearly over France but the IOTA² chain itself can be run over different region given any adapted training data. During the present project, we used the already-available French land cover maps over a 75-km-long reach of the upstream Garonne river near Toulouse in France and we run by ourselves the IOTA² chain over two additional domain: a 96-km-long reach along the middle Po river in Italy and a 250-km-long portion of the Brahmaputra river. The method was first tested over the Garonne and Po domains and the third one was only used in applications. The interesting aspect with using successively these three domains is the diminution of available local – and field



Table 16. Efficiency metrics computed on the HiVDI results using former and new method.

Case	Former method		New method	
	NRMSE (%)	NSE (-)	NRMSE (%)	NSE (-)
Garonne river	20.59	0.91	10.83	0.98
Po river	60.64	0.57	52.57	0.68
Brahmaputra river	45.57	0.89	47.07	0.88

– data to be used. Namely, land cover maps from the IOTA² chain are only available over France; local DEM and insitu data were acquired only for the Garonne and Po domain while profiles had to be reconstructed and simulated data had to be used for the Brahmaputra river. Therefore, we were able to apply our methodology over both well-gauged and ungauged domains.

Our chain aims at estimating physically-based roughness coefficients, therefore, for specific modeling applications, our coefficients still require some further tuning and calibration to match the model physics. Still, we evaluated the quality of such parameters against calibrated ones via hydraulic simulations using the Mascaret 1D-hydraulic model and by comparing the simulated water levels to observed water levels obtained with insitu data. For both testing domains, namely the Garonne and Po domain, neither the simulations using the calibrated coefficients nor the simulations using the estimated coefficients performed clearly better than the other. These encouraging results allow us to serenely use those roughness coefficients as-is as first guess to initialize different kinds of applications for hydrology. One should keep in mind though that the produced roughness coefficients are better suited for one-dimensional applications.

Finally, the roughness coefficients generated with our method were used as first guess in two applications: first, a real-data DA experiment and second, an execution of one of the SWOT discharge algorithm, here HiVDI. The DA experiment purpose was to correct simultaneously the model's roughness coefficients in both the main channel and the floodplain, for the Mascaret model applied over the Po domain by assimilating insitu and altimetry observations of water levels (for a total of 7 observing gauges). The experiment showed the capacity of neatly reducing the water level RMSE. Performances were less obvious downstream but could be explained by the model implementation forcing the water level as downstream boundary condition. Thus, the roughness coefficients generated by our method can be easily used in DA applications.

The second application is more related to SWOT-based applications. Indeed, one of the purposes of generating global scale "good" priors of roughness coefficients is to later on use them within SWOT discharge algorithms. The HiVDI algorithm, one of the official in-development SWOT discharge algorithm, is applied over the Brahmaputra domain, one of the PEPSI dataset domains, using our coefficients as first guess for the friction parameters. To evaluate the benefits of our roughness coefficients, the performances of the experiment were compared to the usual implementation using the mean width to estimate the roughness coefficients. Results were very promising as experiments using our roughness coefficients performed better or slightly worse (for the Brahmaputra only) than the classical implementation. Given that it has been shown that the overall performances of SWOT discharge algorithms highly depend on the quality of the priors, such results show the potential benefits of this new methodology to infer roughness coefficients and better constrain SWOT discharge algorithm.



The study offer several perspectives. First, given training data and cloud-free Sentinel 2 images are gathered at global scale, one could aim at extend the methodology to build global scale maps of roughness coefficients. Riverlines from Frasson et al. (2019) or the SWOT a priori river database could be used to set the spatial coverage and resolution of the resulting product. Roughness would be easily derived in the floodplain given that only global-scale raster products are used. The main challenge would lie in deriving cross-sectional profiles at global scale. Meanwhile, we are aware that several features of our methodology were parameterized according to our own knowledge in general hydrology and on the study domains we chose. To adjust our choices and/or assert them more, it would interesting to lead a full sensitivity analysis of the roughness coefficients estimated by our methodology to the different terms of the Cowan formula and to the inputs.

Code and data availability. The IOTA² chain is available through the framagit platform at <https://framagit.org/iota2-project/iota2>. The toolbox used to build cross-sectionnal profiles over the Brahmaputra river domain are hosted on Kevin Larnier's private github account, please directly contact Kevin Larnier's (kevin.larnier@csgroup.eu). To access the roughness coefficient extraction chain, please directly contact co-authors Santiago Pena Luque (santiago.penaluque@cnes.fr) or Charlotte Emery. All ancillary databases used for this study are freely available online. The GRWL vector product is available at <https://zenodo.org/record/1297434#.YJOrbSY68UE>. The World-DatabaseRivers database is available at <https://zenodo.org/record/2582500#.YNismTo69hF>. The RGE ALTI French DEM is available at <https://geoservices.ign.fr/documentation/diffusion/telechargement-donnees-libres.html>. French land cover maps are available yearly at <https://www.theia-land.fr/ceslist/ces-occupation-des-sols/>. Insitu river data were downloaded from the French Banque Hydro website (<http://www.hydro.eaufrance.fr/>). The Po basin DEM has been downloaded from <https://adbpo.gov.it/carta-del-po-e-dtm/>. Finally, when available, altimetry water surface elevations from virtual stations were downloaded from the Theia-Hydroweb website at <http://hydroweb.theia-land.fr/>.

Appendix A: Corine Land Cover classes aggregation

The following tables detail how Corine Land Cover classes were aggregated to match our 8 meta-classes.

Appendix B: Illustration of n_1 and n_2 Cowan's formula terms over the study domains

B1 Garonne domain

B1.1 Term n_1

Figures B1-B2 illustrate some raw profiles and their smoothed version that give different range of the n_1 term.

B1.2 Term n_2

Figure B3 displays the widths value from GRWL along the domain and the associated first derivative used to estimate n_2 .



CLC classes	code	Aggregated class
Continuous urban fabric	111	Artificial surfaces
Discontinuous urban fabric	112	
Industrial and commercial areas	121	
Road and rail network with associated areas	122	
Port areas	123	
Airports	124	
Dump sites	132	
Construction sites	133	
Green urban areas	141	
Sport and leisure facilities	142	

Table A1. Aggregation of CLC land cover classes related to artificial surfaces

CLC classes	code	Aggregated class
Non-irrigated arable land	211	High-height agricultural areas
Permanently irrigated land	212	
Rice fields	213	
Vineyards	221	
Fruit trees and berry plantations	222	
Olive groves	223	
Annual crops associated with permanent crops	241	
Complex cultivation patterns	242	
Land principally occupied by agriculture , with significant areas of natural vegetation	243	
Sparsely vegetated areas	333	
Natural grasslands	321	Low-height agricultural areas
Moors and heathland	322	
Sclerophyllous vegetation	323	
Burnt areas	334	

Table A2. Aggregation of CLC land cover classes related to agricultural areas



CLC classes	code	Aggregated class
Agro-forestry areas	244	Forests
Broad-leaved forest	311	
Coniferous forest	312	
Mixed forest	313	
Transitional woodland-shrub	324	
Inland marshes	411	Mixed natural areas
Peatbogs	412	
Mineral extraction sites	131	Natural bare grounds
Beaches, dunes, sands	331	
Bare rocks	332	

Table A3. Aggregation of CLC land cover classes related to natural surfaces.

CLC classes	code	Aggregated class
Glaciers and perpetual snow	335	Glaciers and perpetual snow
Salt marshes	421	-
Salines	422	
Intertidal flats	423	Water bodies
Water courses	511	
Water bodies	512	
Coastal lagoons	521	
Estuaries	522	
Sea and ocean	523	

Table A4. Aggregation of CLC remaining land cover classes.

B2 Po domain

495 B2.1 Term n_1

Figures B4-B5 illustrate some raw profiles and their smoothed version that give different range of the n_1 term.

B2.2 Term n_2

Figure B6 displays the widths value from GRWL along the domain and the associated first derivative used to estimate n_2 .

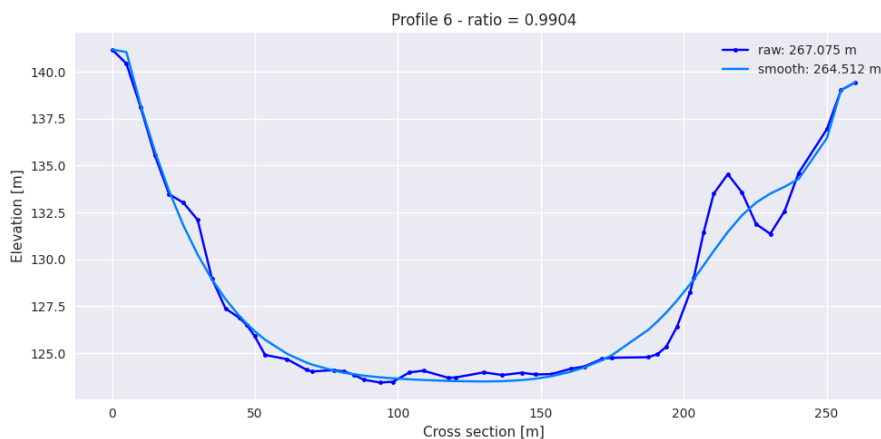


Figure B1. Comparison of the 6th raw and smoothed cross section profiles over the Garonne domain. The length ratio is of 0.9904 giving $n_1 = 0.003$.

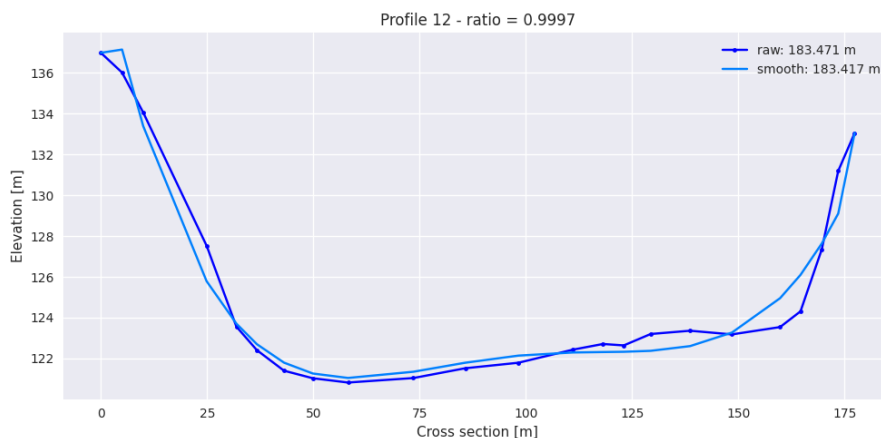


Figure B2. Comparison of the 12th raw and smoothed cross section profiles over the Garonne domain. The length ratio is of 0.9997 giving $n_1 = 0.0$.

Appendix C: Descriptive statistics on estimated roughness coefficients

500 C1 Garonne domain

C2 Po domain

Appendix D: Ensemble Kalman filter implementation with SMURF

This appendix details EnKF implementation choices specific to SMURF and the DA experiments presented in Section 3.1.



Land type	mean	std	min	25%	50%	75%	max
Artificial surfaces	17.89	0.143	16.58	17.79	17.88	17.99	19.14
High-height agricultural areas	12.13	0.052	11.47	12.10	12.13	12.16	12.62
Low-height agricultural areas	28.05	0.329	24.83	27.88	28.07	28.24	31.01
Forest	7.82	0.034	7.56	7.81	7.82	7.84	8.05
Mixed natural areas	15.95	0.192	14.87	15.85	15.91	16.01	16.85
Natural bare soils	39.97	1.40	34.11	39.63	39.90	40.44	46.19
Ice and snow	-	-	-	-	-	-	-
Water - Main channel	24.89	1.91	18.90	23.40	25.98	26.30	27.59

Table C1. Descriptive statistics of estimated Strickler coefficients over the Garonne domain over the floodplain per land cover land type (rows 1-7) and over the main channel (row 8): mean (column 1), standard deviation (column 2), minimum (column 3), first quartile (column 4), median (column 5), third quartile (column 6) and maximum (column 7). Note that the "Water" land type refers to the main channel.

Table C2. Descriptive statistics of estimated Strickler coefficients over the Garonne domain over the floodplain per land cover land type (rows 1-7) and over the main channel (row 8): mean (column 1), standard deviation (column 2), minimum (column 3), first quartile (column 4), median (column 5), third quartile (column 6) and maximum (column 7). Note that the "Water" land type refers to the main channel. Note also that there is no occurrence of the "low-height agricultural areas" over this domain.

Land type	mean	std	min	25%	50%	75%	max
Artificial surfaces	18.02	0.139	16.61	17.95	18.02	18.10	19.17
High-height agricultural areas	12.14	0.046	11.46	12.09	12.13	12.19	12.64
Low-height agricultural areas	-	-	-	-	-	-	-
Forest	7.82	0.049	7.56	7.81	7.82	7.84	8.06
Mixed natural areas	15.92	0.129	14.90	15.87	15.90	15.95	16.84
Natural bare soils	40.86	1.908	34.21	40.44	40.81	41.04	47.18
Ice and snow	-	-	-	-	-	-	-
Water - Main channel	33.05	4.338	21.67	30.82	32.79	36.36	41.02

Table C3. Descriptive statistics of estimated Strickler coefficients over the Po domain over the floodplain per land cover land type (rows 1-7) and over the main channel (row 8): mean (column 1), standard deviation (column 2), minimum (column 3), first quartile (column 4), median (column 5), third quartile (column 6) and maximum (column 7). Note that the "Water" land type refers to the main channel. Note also that there is no occurrence of the "low-height agricultural areas" over this domain.

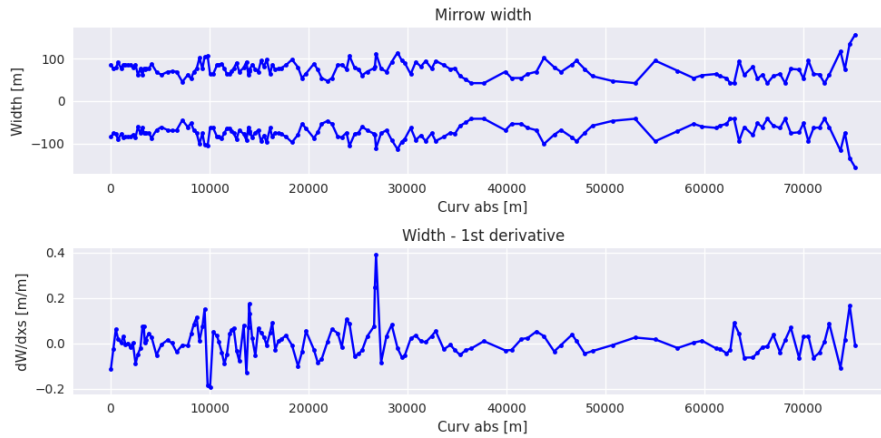


Figure B3. (Top) Mirror river width (centered around 0) with respect to the curvilinear abscissa over the upstream Garonne domain. (Bottom) First derivative with respect to the curvilinear abscissa of the above river width.

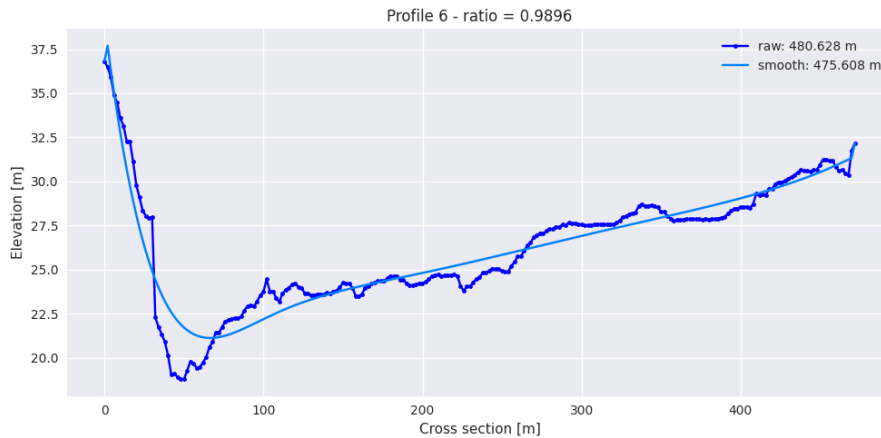


Figure B4. Comparison of the 6th raw and smoothed cross section profiles over the Po domain. The length ratio is of 0.9896 giving $n_1 = 0.008$.

The control vector \mathbf{x} contains the 7 Strickler coefficients for the main channel and the 7 Strickler coefficients for the flood-
 505 plain used in Mascaret for a hydraulic simulation. It is initialized with the friction coefficients issued from the methodology
 presented in Section 2.1 averaged over the 7 zones between observation gauges. At a given assimilation cycle k , the prior
 version of the control vector, before assimilation, is denoted \mathbf{x}_k^b , with b standing for "background".

As for any Kalman-type DA algorithm, the EnKF comprises two steps per assimilation cycle k :

1. A prediction step, during which the control vector is propagated from time t_k to t_{k+1} using the numerical model $\mathcal{M}_{k,k+1}$
 510 – here Mascaret.

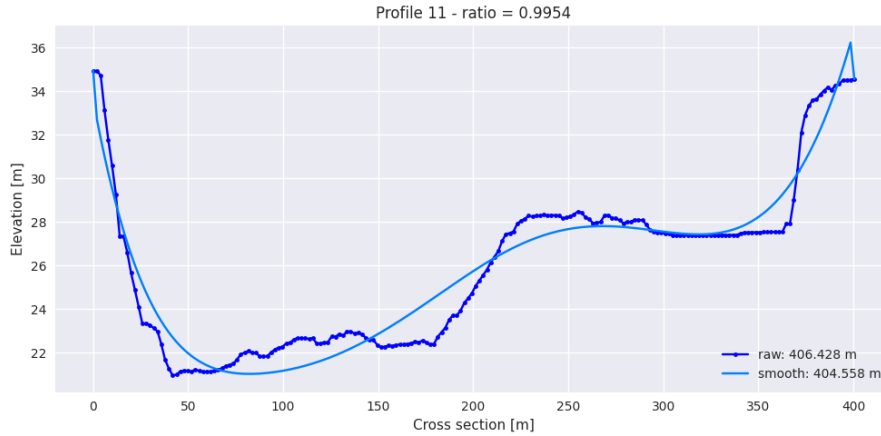


Figure B5. Comparison of the 11th raw and smoothed cross section profiles over the Po domain. The length ratio is of 0.9954 giving $n_1 = 0.003$.

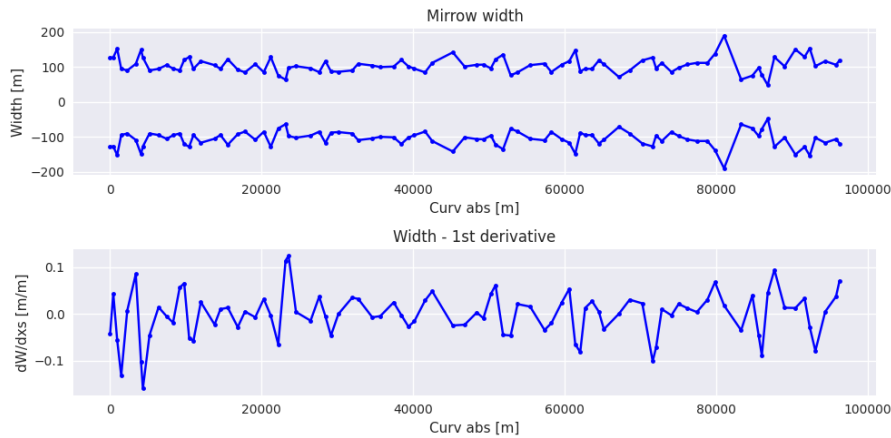


Figure B6. (Top) Mirror river width (centered around 0) with respect to the curvilinear abscissa over the Po domain. (Bottom) First derivative with respect to the curvilinear abscissa of the above river width.

515

2. An analysis step, generating the analysis control vector denoted \mathbf{x}_k^a , with a standing for "analysis", updated given the prior control vector, the available observations \mathbf{y}_k^o and their associated uncertainties, gathered in matrices \mathbf{B}_k for the control vector and \mathbf{R}_k for the observations. Moreover, when the control variables are of a different nature of the observations, as in our current experiment where control variables are Strickler coefficients and the observations are water surface elevations, an observation operator denoted \mathcal{H}_k (and \mathbf{H}_k for its matrix form) is used to map the control variables in the observation space, giving, the Kalman-type analysis equation:

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{B}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b)), \quad (\text{D1})$$



Now, the EnKF specificities state that the control-variable-related error covariance matrices can be stochastically estimated from an ensemble of simulations. Instead of working with a single instance of the control vector, the EnKF works with a control ensemble $\mathbf{X}_{e,k}^b$ containing n_e instances of the control vector, each drawn from a chosen random distribution characterizing the uncertainty model:

$$\mathbf{X}_{e,k}^b = \begin{bmatrix} \mathbf{x}_k^{[1],b} & \mathbf{x}_k^{[2],b} & \dots & \mathbf{x}_k^{[n_e],b} \end{bmatrix}. \quad (\text{D2})$$

Note that in our application, we chose $n_e = 100$ to ensure an large-enough ensemble regarding sample errors but still reasonable to avoid long execution time. In SMURF, a uniform law is used to perturb the control vector such as:

$$\forall l = 1 \dots n_e : \mathbf{x}_k^{[l],b} \sim \mathcal{U} \left(\overline{\mathbf{x}_k^b} \pm p\% \right), \quad (\text{D3})$$

where $\overline{\mathbf{x}_k^b}$ is the mean prior control vector from the previous-cycle control ensemble and p is the range around the mean values. Recall that p was set to 50 for Ksmaj and 30 for Ksmin. Note that this implementation implies that each control variable in the control vector is natively independent from all other variables.

Along with the control ensemble $\mathbf{X}_{e,k}^b$, its counterpart in the observation space is then generated giving:

$$\mathcal{H}(\mathbf{X}_{e,k}^b) = \begin{bmatrix} \mathcal{H}(\mathbf{x}_k^{[1],b}) & \mathcal{H}(\mathbf{x}_k^{[2],b}) & \dots & \mathcal{H}(\mathbf{x}_k^{[n_e],b}) \end{bmatrix}. \quad (\text{D4})$$

Using these two ensembles, control-variable-related error covariance matrices are directly estimated as:

$$[\mathbf{BH}^T]_{e,k} = \frac{1}{n_e - 1} \sum_{l=1}^{n_e} \left(\mathbf{x}_k^{[l],b} - \overline{\mathbf{X}_{e,k}^b} \right) \left(\mathcal{H}(\mathbf{x}_k^{[l],b}) - \overline{\mathcal{H}(\mathbf{X}_{e,k}^b)} \right)^T \quad (\text{D5})$$

and

$$[\mathbf{HBH}^T]_{e,k} = \frac{1}{n_e - 1} \sum_{l=1}^{n_e} \left(\mathcal{H}(\mathbf{x}_k^{[l],b}) - \overline{\mathcal{H}(\mathbf{X}_{e,k}^b)} \right) \left(\mathcal{H}(\mathbf{x}_k^{[l],b}) - \overline{\mathcal{H}(\mathbf{X}_{e,k}^b)} \right)^T \quad (\text{D6})$$

SMURF then apply the principle of observation randomization (Burgers et al., 1998) where the observation vector is also randomized such that:

$$\mathbf{Y}_{e,k}^o = \begin{bmatrix} \mathbf{y}_k^{[1],o} & \mathbf{y}_k^{[2],o} & \dots & \mathbf{y}_k^{[n_e],o} \end{bmatrix}, \text{ with } \forall l = 1 \dots n_e : \mathbf{y}_k^{[l],o} = \mathbf{y}_k^o + \epsilon_k^{[l],o}, \epsilon_k^{[l],o} \sim \mathcal{N}(0, \mathbf{R}_k). \quad (\text{D7})$$

In the litterature, there are two ways to apply Equation D1. One can apply it individually to each member to generate an analysis control ensemble or one can apply it to the mean of the background control ensemble. In SMURF the second option is implemented.

Finally, note that, in the classical definition of the Kalman filter, there exists an analysis equation to also update the background error covariance matrix \mathbf{B}_k from one assimilation cycle to another. In SMURF original implementation, decision was made not to use this equation and keep the initial definition of the control variable errors. This implementation allows to maintain enough dispersion in the control ensemble and ensure EnKF efficiency. However, for our application, we decided to propagate the analysis control error matrix from one assimilation cycle to another.



Author contributions. All first four authors participated in the definition of the method. Charlotte Emery conducted all the DA experiments using the Mascaret model and the SMURF framework. Maxime Liquet prepared all the data for the Po river and conducted the simulation and validation experiments. Kevin Larnier performed all the experiments for the application for the SWOT discharge algorithms and provided related datasets. Joao Hemptinne implemented and tested the method. Arthur Vincent provided support to use and apply the IOTA² chain over the Po and Brahmaputra domains. Finally, Santiago Pena Luque acted as the project coordinator and also provided input data to run the IOTA² chain

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This paper relays the work led by CS GROUP-France team for the CNES R&T project entitled "Parameters extraction to estimate friction coefficients from land cover maps generated by the IOTA² chain". It was supported by CNES within the framework of the SWOT-AVAL project. The authors would like to thank Mannaig L'Haridon and Catherine Bouzinac for their work and support at the project initiation. The authors also would like to thank Angelica Tarpanelli from the Italian Research Council for her considerable help in getting insitu data over the Po study domain. Finally, the authors would like to express their gratitude to Elena Poggi from Autorità di bacino distrettuale del fiume Po and Sandro Nanni from Agenzia regionale per la protezione ambientale Emilia-Romagna for their support in getting and understanding insitu data over the Po river.



560 References

- Allen, G. H. and Pavelsky, T. M.: Global extent of rivers and streams, *Science*, 361, 585–588, <https://doi.org/10.1126/science.aat0636>, 2018.
- Andreadis, K. M., Brinkerhoff, C. B., and Gleason, C. J.: Constraining the Assimilation of SWOT Observations With Hydraulic Geometry Relations, *Water Resources Research*, 56, <https://doi.org/https://doi.org/10.1029/2019WR026611>, 2020.
- Arcement, G. J. and Schneider, V. R.: Guide for Selecting Manning’s Roughness Coefficients for Natural Channels and Flood Plains, U.S. Geological Survey Water Supply, 2339, <https://doi.org/10.3133/wsp2339>, 1989.
- 565 Baetens, L., Desjardins, C., and Hagolle, O.: Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure, *Remote Sensing*, 11, <https://doi.org/10.3390/rs11040433>, 2019.
- Bjerklie, D. M., Birkett, C. M., Jones, J. W., Carabjal, C., Rover, J. A., Fulton, J. W., and Garambois, P.-A.: Satellite remote sensing estimation of river discharge: Application to the Yukon river Alaska, *Journal of Hydrology*, 561, 1000–1018, <https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.04.005>, 2018.
- 570 Brinkerhoff, C., Gleason, C., and Ostendorf, D.: Reconciling At-a-Station and At-Many-Stations Hydraulic Geometry through River-Wide Geomorphology, *Geophysical Research Letters*, 46, <https://doi.org/10.1029/2019GL084529>, 2019.
- Brinkerhoff, C., Gleason, C., Feng, D., and Lin, P.: Constraining Remote River Discharge Estimation Using Reach-Scale Geomorphology, *Water Resources Research*, 56, <https://doi.org/10.1029/2020WR027949>, 2020.
- 575 Brisset, P., Monnier, J., Garambois, P.-A., and Roux, H.: On the assimilation of altimetric data in 1D Saint-Venant river flow models, *Advances in Water Resources*, 119, 41–59, <https://doi.org/10.1016/j.advwatres.2018.06.004>, 2018.
- Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N. E., Herold, M., and Fritz, S.: Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe 2020, <https://doi.org/https://doi.org/10.5281/zenodo.3939050>, 2020.
- 580 Burgers, G., Leeuwen, P. J. V., , and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, *Monthly Weather Review*, 126, 1719–1724, [https://doi.org/10.1175/1520-0493\(1998\)126<1719:ASITEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2), 1998.
- Chaulagain, S.: An investigation into remote sensing techniques for describing hydraulic roughness, Ph.D. thesis, University of New Mexico, 2018.
- Chow, V.: Handbook of applied hydrology, McGraw-Hill Book Co., New-York, 1467 pages, 1964.
- 585 Cowan, W. L.: Estimating hydraulic roughness coefficients, *Agricultural Engineering*, 37, 473–475, 1956.
- Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.: Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global offline hydrological applications, *Journal of Advances in Modeling Earth Systems*, 11, 1207–1252, <https://doi.org/https://doi.org/10.1029/2018MS001545>, 2019.
- Desrochers, N. M., Trudel, M., Peters, D. L., Siles, G., and R., L.: Hydraulic Model Calibration Using Water Levels Derived from Time Series High-resolution SAR Images, *Journal of Hydraulic Engineering*, 146, [https://doi.org/10.1061/\(asce\)hy.1943-7900.0001687](https://doi.org/10.1061/(asce)hy.1943-7900.0001687), 2020.
- 590 Durand, M., Neal, J., Rodriguez, E., Andreadis, K., Smith, L., and Yoon, Y.: Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope, *Journal of Hydrology*, 511C, 92–104, <https://doi.org/j.jhydrol.2013.12.050>, 2014.
- Durand, M., Gleason, C. J., Garambois, P.-A., Bjerklie, D., Smith, L. C., Roux, H., Rodriguez, E., Bates, P. D., Pavelsky, T. M., Monnier, J., Chen, X., Di Baldassare, G., Fiset, J.-M., Flipo, N., Frasson, R. P. d. M., Fulton, J., Goutal, N., Hossein, F., Humphries, E., Minear, J. T., Mukolwe, M. M., Neal, J. C., Ricci, S., Sanders, B. F., Schumann, G., Schubert, J. E., and Vilmin, L.: An intercomparison of
- 595



- remote sensing river discharge estimation algorithms from measurements of river height, *Water Resources Research*, 52, 4527–4549, <https://doi.org/10.1002/2015WR018434>, 2016.
- Ebrahimi, N. G., Fathi-Moghadaù, M., Kashefipour, S. M., Saneie, M., and Ebrahimi, K.: Effects of Flow and Vegetation States on River Roughness Coefficients, *Journal of Applied Sciences*, 8, 2118–2123, <https://doi.org/10.3923/jas.2008.2118.2123>, 2008.
- 600 Evensen, G.: The Ensemble Kalman filter: theoretical formulation and practical implementation, *Ocean Dynamics*, 53, 343–367, <https://doi.org/10.1007/s10236-003-0036-9>, 2003.
- Fathi-Moghadam, M., Drikvandi, K., Lashkarara, B., and Hammadi, K.: Determination of friction factor for rivers with non-submerged vegetation in banks and floodplains, *Scientific Research and Essays*, 6, 4714–4719, <https://doi.org/10.5897/SRE11.751>, 2011.
- Fekete, B. M. and Vörösmarty, J.: The current status of global river discharge monitoring and potential new technologies complementing
605 traditional discharge measurements, *Proceedings of the PUB Kick-off Meeting*, 309, 20–22, 2007.
- Frasson, R., Pavelsky, T., Fonstad, M., Durand, M., Allen, G., Schumann, G., Lion, C., Beighley, R. E., and Yang, X.: Global relationships between river width, slope, catchment area, meander wavelength, sinuosity, and discharge, *Geophysical Research Letters*, 46, 3252–3262, <https://doi.org/10.1029/2019GL082027>, 2019.
- Frasson, R., Durand, M., Larnier, K., Gleason, C., Andreadis, K., Hagemann, M., Dudley, R., Bjerklie, D., Oubanas, H., Garam-
610 bois, P.-A., Malaterre, P., Lin, P., Pavelsky, T., Monnier, J., Brinkerhoff, C., and David, C.: Exploring the factors controlling the error characteristics of the Surface Water and Ocean Topography mission discharge estimates, *Water Resources Research*, <https://doi.org/10.1029/2020WR028519>, 2021.
- Galema, A.: Evaluation of vegetation resistance descriptors for flood management, Ph.D. thesis, University of Twente, 2009.
- Goeury, C., Ponçot, A., Argaud, J.-P., Zaoui, F., Ata, R., and Audouin, Y.: Optimal calibration of Telemac-2D models based on a data
615 assimilation algorithm, in: *TELEMAC-MASCARET user conference*, Graz University of Technology, 2017.
- Goutal, N., Lacombe, J.-M., Zaoui, F., and el Kadi Abderrezzak, K.: MASCARET: A 1-D open-source software for flow hydrodynamic and water quality in open channel networks, *River Flow 2012 - Proceedings of the International Conference on Fluvial Hydraulics*, 2, 1169–1174, 2012.
- Habert, J., Ricci, S., Thual, O., Le Page, E., Piacentini, A., Goutal, N., Jonville, G., and Rochoux, M.: Reduction of the uncertainties in the
620 water level-discharge relation of a 1D hydraulic model in the context of operational flood forecasting, *Journal of Hydrology*, 532, 52–64, <https://doi.org/10.1016/j.jhydrol.2015.11.023>, 2016.
- Hagemann, M. W., Gleason, C. J., and Durand, M. T.: BAM: Bayesian AMHG-Manning inference of discharge using remotely sensed stream width, slope, and height, *Water Resources Research*, 53, 9692–9707, <https://doi.org/https://doi.org/10.1002/2017WR021626>, 2017.
- Hamidifar, H., Keshavarzi, A., and Rowiński, P. M.: Influence of Rigid Emerged Vegetation in a Channel Bend on Bed Topography and Flow
625 Velocity Field: Laboratory Experiments, *Water*, 12, <https://doi.org/10.3390/w12010118>, 2020.
- Hannah, D. M., Demuth, S., van Lanen, H. A. J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., and Tallaksen, L. N.: Large-scale river flow archives: importance, current status and future needs, *Hydrological Processes*, 25, 1191–1200, <https://doi.org/https://doi.org/10.1002/hyp.7794>, 2011.
- Hostache, R., Lai, X., Monnier, J., and Puech, C.: Assimilation of spatial distributed water levels into a shallow water flood model. Part II :
630 using a remote sensing image of Mosel river, *Journal of Hydrology*, 390, 257–268, <https://doi.org/10.1016/j.jhydrol.2010.07.003>, 2010.
- Hygelund, B. and Manga, M.: Field measurements of drag coefficients for model large woody debris, *Geomorphology*, 51, 175–185, [https://doi.org/10.1016/S0169-555X\(02\)00335-5](https://doi.org/10.1016/S0169-555X(02)00335-5), 2003.



- Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., and Rodes, I.: Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series, *Remote Sensing*, 9, <https://doi.org/10.3390/rs9010095>, 2017.
- 635 Jarvela, A.: Effect of submerged flexible vegetation on flow structure and resistance, *Journal of Hydrology*, 307, 233–241, <https://doi.org/10.1016/j.jhydrol.2004.10.013>, 2005.
- Jung, I., Park, J., Park, G., Lee, M., and Kim, S.: A grid-based rainfall-runoff model for flood simulation including paddy fields, *Paddy Water Environ*, 9, 275–290, <https://doi.org/10.1007/s10333-010-0232-4>, 2011.
- Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, *Transactions of the ASME - Journal of Basic Engineering*, 82, 640 35–45, <https://doi.org/10.1115/1.3662552>, 1960.
- Kean, J. and Smith, J.: Form drag in rivers due to small-scale natural topographic features, *Journal Of Geophysical Research*, 111, <https://doi.org/10.1029/2006JF000467>, 2006.
- Larnier, K. and Monnier, J.: Hybrid Neural Network – Variational Data Assimilation algorithm to infer river discharges from SWOT-like data, *Nonlinear Processes in Geophysics Discussions*, 2020, 1–30, <https://doi.org/10.5194/npg-2020-32>, 2020.
- 645 Larnier, K., Monnier, J., Garambois, P. A., and Verley, J.: River discharge and bathymetry estimation from SWOT altimetry measurements, *Inverse Problems in Sciences and Engineering*, 29, 759–789, <https://doi.org/10.1080/17415977.2020.1803858>, 2021.
- Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., and Wood, E. F.: Inroads of remote sensing into hydrologic science during the WRR era, *Water Resources Research*, 51, 7309–7342, <https://doi.org/10.1002/2015WR017616>, 2015.
- Manning, R.: On the flow of water in open channels and pipes, *Transactions of the Institution of Civil Engineers of Ireland*, 20, 161–207, 650 1891.
- Mirouze, I. and Ricci, S.: Smurf: System for Modelling with Uncertainty Reduction, and Forecasting, *Journal of Open Research Software*, 9, 2–10, <https://doi.org/https://doi.org/10.5334/jors.312>, 2021.
- Mirouze, I., Ricci, S., and Goutal, N.: The impact of observation spatial and temporal densification in an ensemble Kalman Filter, in: *TELEMAC-MASCARET User Conference*, Toulouse, France, <https://doi.org/10.5281/zenodo.3549572>, 2020.
- 655 Nepf, H. M.: Drag, turbulence, and diffusion in flow through emergent vegetation, *Water Resources Research*, 35, 479–489, 1999.
- Oki, T. and Kanae, S.: Global Hydrological Cycles and World Water Resources, *Science*, 313, 1068–1072, <https://doi.org/10.1126/science.1128845>, 2006.
- Paris, A., Dias de Paiva, R., Santos da Silva, J., Madeiros Moreira, D., Garambois, P.-A., Collischonn, W., Bonnet, M.-P., and Seyler, F.: Stage-discharge rating curves based on satellite altimetry and modeled discharge in the Amazon basin, *Water Resources Research*, 52, 660 3787–3814, <https://doi.org/10.1002/2014WR016618>, 2016.
- Peña Luque, S., Ferrant, S., Cordeiro, M., and Martinez, J.-M.: Sentinel-1&2 Multitemporal Water Surface Detection Accuracies, Evaluated at Regional and Reservoirs Level, *Remote Sensing*, Review in progress, 2021.
- Peirong, L., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global reconstruction of naturalized river flowlines at 2.94 million reaches, *Water Resources Research*, 55, 6499–6516, 665 <https://doi.org/10.1029/2019WR025287>, 2017.
- Poggio, L., Duque Moreira de Sousa, L., Heuvelink, G., Kempen, B., Batjes, N., Leenaars, J., Mantel, S., Bai, Z., Turdukulov, U., Ruiperez Gonzalez, M., Carvalho Ribeiro, E., Rossiter, D., and van den Bosch, H.: SoilGrids: consistent soil information to assess and map soil functions at global scale, pp. 50–51, ISRIC, 2019.



- Pontes, P. R. M., Fan, F. M., Fleischmann, A. S., de Paiva, R. C. D., Buarque, D. C., Siqueira, V. A., Jardim, P. F., Sorribas, M. V., and
670 Collischonn, W.: MGB-IPH model for hydrological and hydraulic simulation of large floodplain river systems coupled with open source
GIS, *Environmental Modelling and Software*, 94, 1–20, <https://doi.org/https://doi.org/10.1016/j.envsoft.2017.03.029>, 2017.
- Pu, J., Hussain, A., Guo, Y., Vardakastanis, N., Hanmaiahgari, P., and Lam, D.: Submerged flexible vegetation impact on open chan-
nel flow velocity distribution: An analytical modelling study on drag and friction, *Water Science and Engineering*, 12, 121–128,
<https://doi.org/10.1016/j.wse.2019.06.003>, 2019.
- 675 Tanino, Y. and Nepf, H.: Laboratory Investigation of Mean Drag in a Random Array of Rigid, Emergent Cylinders, *Journal of Hydraulic
Engineering-asce - J HYDRAUL ENG-ASCE*, 134, [https://doi.org/10.1061/\(ASCE\)0733-9429\(2008\)134:1\(34\)](https://doi.org/10.1061/(ASCE)0733-9429(2008)134:1(34)), 2008.
- Tarpanelli, A., Brocca, L., Lacava, T., Melone, F., Moramarco, T., Faruolo, M., Pergola, N., and Tramutoli, V.: Toward the es-
timation of river discharge variations using MODIS data in ungauged basins, *Remote Sensing of Environment*, 136, 47–55,
<https://doi.org/https://doi.org/10.1016/j.rse.2013.04.010>, 2013.
- 680 Tuozzolo, S., Lind, G., Overstreet, B., Mangano, J., Fonstad, M., Hagemann, M., Frasson, R. P. M., Larnier, K., Garambois, P.-A., Monnier,
J., and Durand, M.: Estimating river discharge with swath altimetry: A proof of concept using AirSWOT observations, *Geophysical
Research Letters*, 46, 1459–1466, <https://doi.org/https://doi.org/10.1029/2018GL080771>, 2019.
- Van Dijk, A. I. J. M., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groeve, T., and Schellekend, J.: River gauging at global scale using
optical and passive microwave remote sensing, *Water Resources Resources*, 52, 6404–6418, <https://doi.org/doi:10.1002/2015WR018545>,
685 2016.
- Walczak, N., Hammerling, M., Spychała, M., and Nieć, J.: The influence of vegetations in floodplains on flood risks, *Journal of Ecological
Engineering*, 16, 160–167, <https://doi.org/10.12911/22998993/60473>, 2015.
- Whittaker, P.: Modelling the Hydrodynamic Drag Force of Flexible Riparian Woodland, Ph.D. thesis, University of Cardiff, 2014.
- Yamazaki, D., Kanae, S., Kim, H., and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing
690 model, *Water Resources Research*, 47, <https://doi.org/doi:10.1029/2010WR009726>, 2011.
- Zaoui, F., Goeury, C., and Audouin, Y.: Ensemble integrations of Telemac Mascaret for the optimal model calibration, in: TELEMAC-
MASCARET user conference, Norwich, United Kingdom, 2018.