Validation of an Uncertainty Propagation Method for Moving-Boat Acoustic Doppler Current Profiler Discharge Measurements

Aurélien Despax, Jérôme Le Coz, David S Mueller, Alexandre Hauet, Blaise Calmel, Gilles Pierrefeu, Grégoire Naudet, Bertrand Blanquart, Karine Pobanz

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR031878

Key Points:

- A new uncertainty propagation method (OURSIN) for moving-boat Acoustic Doppler Current Profiler discharge measurements is introduced and implemented in QRevInt software
- The method proposes an uncertainty budget accounting for the main error sources, including discharge extrapolations
- The uncertainty difference from the outputs of two large-scale repeated measures experiments is −0.4% on average

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Le Coz,
jerome.lecoz@inrae.fr

Citation:

Despax, A., Le Coz, J., Mueller, D. S., Hauet, A., Calmel, B., Pierrefeu, G., et al. (2023). Validation of an uncertainty propagation method for moving-boat acoustic Doppler current profiler discharge measurements. *Water Resources Research*, *59*, e2021WR031878. https://doi.org/10.1029/2021WR031878

Received 22 DEC 2021
Accepted 4 JAN 2023

Validation of an Uncertainty Propagation Method for Moving-Boat Acoustic Doppler Current Profiler Discharge Measurements

Aurélien Despax[1,2], Jérôme Le Coz[1], David S. Mueller[3], Alexandre Hauet[4], Blaise Calmel[1], Gilles Pierrefeu[2], Grégoire Naudet[2,5], Bertrand Blanquart[6], and Karine Pobanz[2]

[1]INRAE, UR RiverLy, Lyon, France, [2]Compagnie Nationale du Rhône, Lyon, France, [3]Genesis HydroTech LLC, Louisville, KY, USA, [4]EDF-DTG, Grenoble, France, [5]SMAGGA, Brignais, France, [6]Independent Expert, Nancy, France

Abstract The moving-boat Acoustic Doppler Current Profiler (ADCP) gauging method is extensively used to measure the discharge of rivers and canals. Informed decisions related to water management require reliable estimates of the uncertainty of such measurements. The evaluation of the uncertainty is however a difficult task because of the complexity of the ADCP data workflow and the lack of reference discharges in rivers and canals. This study presents an evolution of the OURSIN method, which follows the framework of the Guide to the expression of Uncertainty in Measurement. The method has been implemented in the QRevInt software which provides an ADCP data quality review prior to the uncertainty analysis. As a computationally efficient alternative to the Monte Carlo approach, the uncertainty propagation combines elemental measurement uncertainty sources, transect-to-transect discharge variability, and estimation of the uncertainty of discharges in unmeasured areas by sensitivity analysis. The OURSIN uncertainty results are validated using empirical uncertainty estimates from two large-scale repeated measures experiments with variable site and flow conditions. The mean and the 95% quantiles of the uncertainty differences are −0.4% [−4.4%; +2.5%]. The OURSIN method provides a comprehensive uncertainty budget that is helpful for understanding the relative impacts of error sources and defining strategies to minimize them. The method should be further evaluated with experiments in different conditions, and additional uncertainty components and options such as GPS reference for boat velocity could be implemented.

1. Introduction

1.1. Moving-Boat ADCP Uncertainty

Acoustic Doppler Current Profilers (ADCPs) are now among the most-used instruments for measuring discharge in rivers throughout the world (Boldt & Oberg, 2015; Le Coz, 2017). The technology and general guidance for making ADCP discharge measurements are presented in various manuals and guides, for example, those established by the U.S. Geological Survey (USGS) (Mueller et al., 2013) or the WMO (2010). The ADCP is mounted on a boat or on a small float that transects a river cross-section. The field procedure can be a stationary deployment (section by section) or a moving-boat deployment, wherein each river crossing provides an individual discharge measurement called a transect. This study deals with moving-boat ADCP discharge measurements. The ADCP uses the transit-time of sound waves and the Doppler shift to measure water velocity and depth. Due to physical limitations of the instrument, velocities are measured throughout a limited portion of the cross-section (Mueller et al., 2013). The measured area is decomposed vertically into cells (or bins) that are distributed horizontally into n ensembles (see Figure 1). The discharge $Q_{i,j}$ through a cell is computed from the vector product of the boat velocity $\overrightarrow{v_i}$ and the water velocity $\overrightarrow{w_{i,j}}$ relative to the ADCP system as:

$$Q_{i,j} = \left(\overrightarrow{w_{i,j}} \times \overrightarrow{v_i}\right) \cdot \vec{k}\, \mathrm{d}z_{i,j}\ \ \mathrm{d}t_i \tag{1}$$

where $\vec{k}$ is the unit vertical vector, $\mathrm{d}z_{i,j}$ is the cell height, and $\mathrm{d}t_i$ is the time interval between ensembles.

In the unmeasured areas (see Figure 1), discharge has to be estimated (Mueller, 2016). The missing or invalid cells and ensembles are interpolated based on contiguous valid data. The discharge is extrapolated near the riverbed (bottom discharge), near the water surface (top discharge), and near the banks (right and left discharges). Thus, the total discharge Q_k in each transect k is a sum of partial discharges: the measured discharge (the sum of individual cell discharges $Q_{i,j}$), the interpolated discharges used to fill missing cells and ensembles, respectively

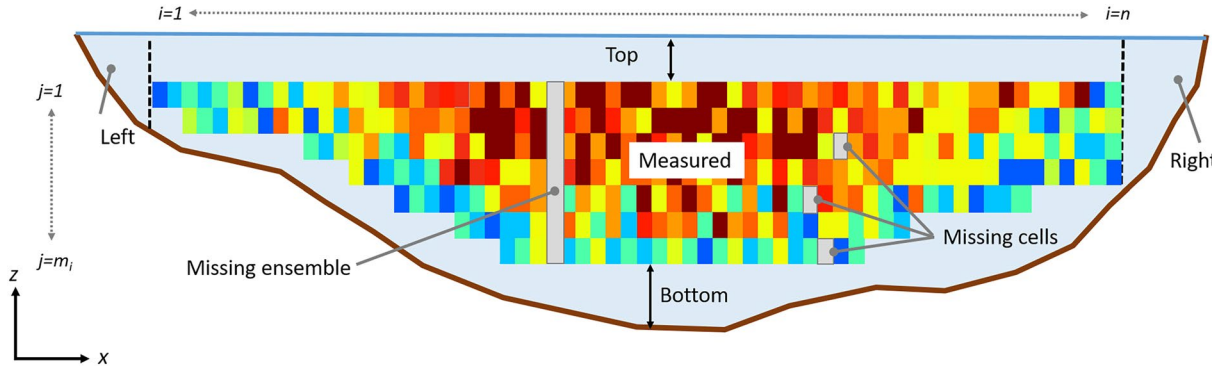

Figure 1. Decomposition of an Acoustic Doppler Current Profiler (ADCP) cross-sectional transect showing measured areas (colored cells), missing samples (dark gray), and unmeasured areas (in blue). The x-axis is the transverse direction and the z-axis is the vertical direction.

(Q_{invcell} and Q_{invens}), and the extrapolated discharges in the top, bottom, left, and right edge unmeasured areas, respectively (Q_{top}, Q_{bot}, Q_{left}, and Q_{right}):

$$Q_k = \sum_{i=1}^{n} \sum_{j=1}^{m_i} Q_{i,j} + Q_{\text{invcell}} + Q_{\text{invens}} + Q_{\text{top}} + Q_{\text{bot}} + Q_{\text{left}} + Q_{\text{right}} \tag{2}$$

where n is the total number of recorded ensembles and m_i is the number of recorded cells in ensemble i.

We define the measured discharge as the first term of the sum in Equation 2: $Q_{\text{meas}} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} Q_{i,j}$. The measured discharge divided by the total discharge is the measured discharge ratio $\eta = Q_{\text{meas}}/Q_k$. In our definition (as opposed to the definition of Teledyne RDI, for instance), discharges from the missing cells and ensembles are not included in the measured discharge: if cell (i, j) (or ensemble i) is invalid, then $Q_{i,j} = 0$ and the interpolated discharges contribute to the Q_{invcell} and Q_{invens} terms. An ADCP discharge measurement $\overline{Q}$ is the average of a number P of single-transect discharges Q_k from successive crossings of the stream or river under approximately steady flow conditions:

$$\overline{Q} = \frac{1}{P} \sum_{k=1}^{P} Q_k \tag{3}$$

The discharge average should include pairs of reciprocal transects to minimize any potential directional biases in measured discharges (Huang, 2019; Le Coz et al., 2008; Mueller et al., 2013). Best practices vary across agencies. The USGS recommends performing at least one pair of reciprocal transects acquired during at least 720 s (Mueller et al., 2013; Oberg & Mueller, 2007), while in France, a minimum of three pairs of reciprocal transects acquired during at least 900 s is specified (Le Coz et al., 2008).

During an ADCP discharge measurement, several errors may occur due to the limited accuracy of the ADCP, the estimation of unmeasured discharges, and site- and operator-induced errors. Error sources in ADCP discharge measurements have been listed by Muste et al. (2004), González-Castro and Muste (2007), Kim and Yu (2010), or Despax et al. (2019). Actually, as soon as ADCPs were used to measure discharge in rivers, the measurement errors and the corresponding uncertainties were investigated by pioneers like Simpson and Oltmann (1991). Given the use of ADCP measurements as inputs to flow monitoring and decision-making (Hamilton & Moore, 2012; Pagano et al., 2014), the uncertainty has to be estimated carefully to promote robust decisions related to water resources and natural hazards (McMillan et al., 2017).

Prior to the uncertainty analysis (UA), a data quality review has to be performed using a quality assurance/quality control (QA/QC) process (Oberg et al., 2005). A powerful tool to conduct a QA/QC process is the QRev software developed by the USGS (Mueller, 2016, 2020), now available as QRevInt, a fork developed by Genesis HydroTech LLC (Mueller, 2021) with the guidance and contributions from an international board of hydrological agencies (https://www.genesishydrotech.com/qrevint). QRevInt helps to clean ADCP measurements from avoidable errors and to homogenize the discharge computations irrespective of the instrument manufacturer and model. This study assumes that the general rules and guidance for making ADCP discharge measurements (Le Coz et al., 2008; Mueller et al., 2013) are followed and a QA/QC process is conducted prior to the UA.

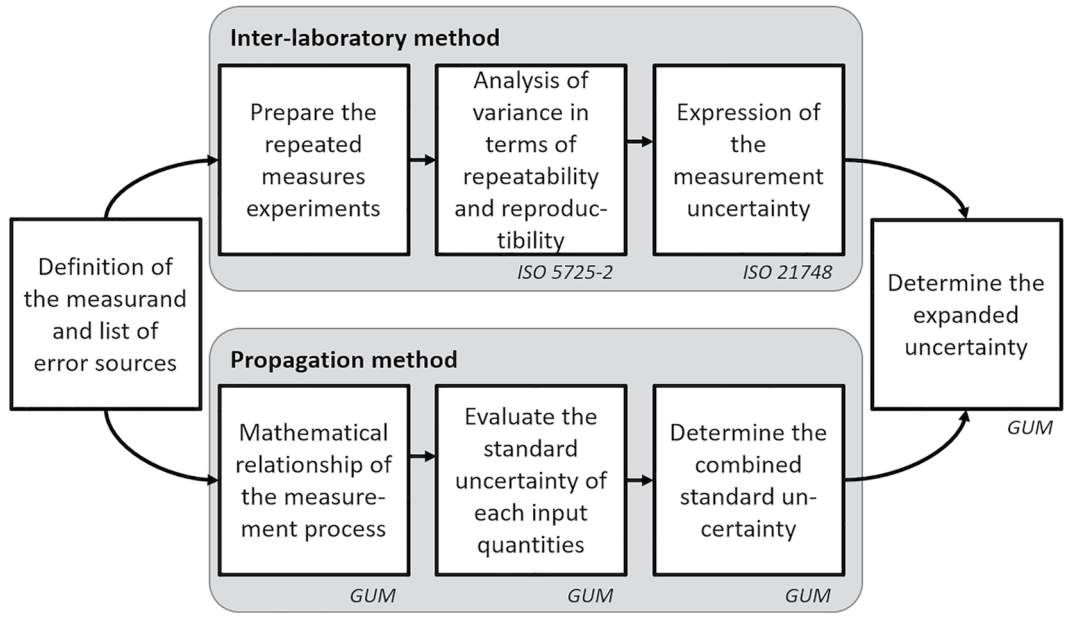

Figure 2. Synoptic chart of inter-laboratory comparisons and propagation methods as supplementary approaches for uncertainty analysis. Adapted from Blanquart (2013).

The uncertainty of a discharge measurement, like any other measurand, can be estimated using two approaches. The uncertainty propagation method as defined in the Guide to the expression of Uncertainty in Measurement or GUM (JCGM, 2008b) can be applied to any single measurement (Section 1.2), while the repeated measures experiments (Section 1.3), also known as the inter-laboratory method (ISO, 1994a; ISO, 2010), provides the average uncertainty of the gauging technique in given conditions (cf. Figure 2). The uncertainty propagation method requires that the measurement process is fully described through a model, the data reduction equation (DRE). By contrast, the inter-laboratory method does not require a measurement model as it empirically estimates the uncertainty due to one (lumped) or several error sources from the variance of successive measurements done in repeatability conditions (exact same conditions of measurement) and in reproducibility conditions (changing a least one factor, e.g., instrument, operator, site). The uncertainty of input quantities (elemental measurements) must be estimated or modeled to feed the propagation method. The error model often has to be simplified, to lump difficult-to-estimate uncertainty components. In turn, the inter-laboratory method can provide realistic values for these uncertainty components and help validate the assumptions made in the propagation method. Both approaches are integrated into the same conceptual framework described in ISO 21748 (ISO, 2017), allowing the determination of the uncertainty through the GUM approach.

1.2. Propagation Methods for Estimating the Uncertainty

The main steps of the GUM (JCGM, 2008a, 2008b), which is the accepted general framework for measurement of UA, are summarized in Figure 2. The Hydrometric Uncertainty Guidance or HUG (ISO, 2007) applies GUM's recommendations for most streamgauging techniques. However, the computation proposed for moving-boat ADCP is problematic as the estimation of some uncertainty terms is missing or difficult. In addition, a software package is lacking to implement the equations and apply them to real ADCP data files. Uncertainty and errors are categorized according to different terminologies that are explained in Appendix B. Type A and type B evaluations distinguish the method used to evaluate the standard uncertainties: direct computation of the standard deviation estimate using data (type A) or assumptions on the error distribution based on other sources of information to compute the standard deviation (type B). The nature of errors can be classified as random or systematic. The way the standard uncertainty is assessed (type A or type B) has no relation with the systematic or random nature of the errors. The uncertainty propagation method is based on a DRE, which expresses the value taken by the measurement Y as a function of several input variables $X_1, \ldots, X_p$ viewed as elemental uncertainty sources: $Y = f(X_1, \ldots, X_p)$. The propagation to the final results uses either a first-order Taylor approximation of

the DRE (JCGM, 2008a) or Monte Carlo simulation (JCGM, 2008b). Mathematical detail relevant for this article is provided in Appendix A.

The general application of the GUM uncertainty propagation method to ADCP discharge measurements has been proposed by Muste et al. (2004), González-Castro and Muste (2007), Kim and Yu (2010), and the HUG (ISO, 2007). Some tools have been developed for estimating the uncertainty of ADCP measurements in stationary (Huang, 2011; Lee et al., 2014) or moving-boat deployment modes: RiverFlowUA (González-Castro et al., 2016), QUant (Moore et al., 2016), the uncertainty computation originally implemented in QRev (Mueller, 2016, 2020), which will be named QRev-UA in this article, and the original version of OURSIN (Naudet et al., 2019), which has been further modified as described in this article. Most of the tools were developed by hydrometric agencies. QRev-UA seems to be the only one that is released externally, is compatible with any type of Teledyne RDI and SonTek ADCP data, and is applied routinely. The South Florida Water Management District uses its software RiverFlowUA for estimating the measurement uncertainty of their ADCP discharge measurements, especially for rating hydraulic structures or resolving litigious conflicts of water use.

RiverFlowUA and QUant were developed to compute the uncertainty in Teledyne RDI ADCP discharge measurements only. RiverFlowUA uses a first-order Taylor approximation of the DRE and accounts for the correlation between the velocities measured in contiguous cells. The method combines the uncertainties estimated from multiple transects and calibration uncertainties in the measured portion. The main limitation of the RiverFlowUA method lies in the fact that it does not account for the uncertainty of the discharges extrapolated in unmeasured areas, which often bring a substantial part of the total uncertainty (Moore et al., 2016).

QUant uses Monte Carlo simulations for assessing the uncertainty, combining both Type A and Type B evaluations of uncertainty and accounting for both random and systematic errors. The simulations account for the uncertainty of each input quantity. At each iteration, the input quantities are randomly sampled from their respective probability distributions and the discharge is computed using these values. The calculation of 1,000 iterations makes the method time-consuming (around 30 min per measurement on a conventional laptop).

The QRev-UA method (Mueller, 2016, 2021) is a simplified approach to GUM. It is based on the combination of uncertainty components, some estimated by expert judgment, others by simple calculations. Assuming that errors are independent across error sources, the relative standard uncertainty $u'^2\left(\overline{Q}\right)$ of total discharge, that is, the standard uncertainty divided by the observed value of $\overline{Q}$, is obtained from the quadratic sum of the uncertainty components:

$$u'^2\left(\overline{Q}\right) = u'^2_{\text{syst}} + u'^2_{\text{cov}} + u'^2_{\text{mb}} + u'^2_{\text{comp}} + u'^2_{\text{invalid}} + u'^2_{\text{edges}} + u'^2_{\text{extrap}} \tag{4}$$

with $u'_{\text{syst}} = 1.5\%$ the uncertainty due to systematic errors related to instrumentation, u'_{cov} the uncertainty due to random measurement errors, u'_{mb} the uncertainty due to moving bed, u'_{comp} the uncertainty due to compass bias, u'_{invalid} the uncertainty due to invalid data (10% of the interpolated discharge), u'_{edges} the uncertainty of edge discharge (15% of the edge discharge), and u'_{extrap} the uncertainty of top and bottom extrapolated discharges (based on the percent difference in discharge among possible extrapolation methods).

The uncertainty due to random measurement errors is computed as $u'_{\text{cov}} = t^{P-1}_{2.5\%} CV / \sqrt{P}$, where CV is the coefficient of variation of the P transect discharges and $t^{P-1}_{2.5\%}$ is the quantile of order 97.5% of the Student's t-distribution with $P-1$ degrees of freedom. The latter coefficient accounts for the limited number P of observations. However, a practical problem arises when only two transects are measured ($P = 2$): then, there is only one degree of freedom and the coverage factor would be 12.7. For a minimal coefficient of variation of 0.01, which used to be the resolution limit in commercial ADCP software (now it is 0.0001 in WinRiverII), the expanded uncertainty would be $U'_{\text{CV}} = 12.7 \times 0.01/\sqrt{2} = 9\%$, which systematically leads to measurements rated as "poor" (uncertainty >8%). Based on an empirical assessment of the quality ratings of two-transect ADCP measurements, the USGS (Mueller, 2012) sets the multiplier to 3.3 instead of 9 with a minimum uncertainty of 3% if the coefficient of variation is equal to 0.01.

If the reference is bottom-track and a valid moving-bed test is performed, $u'_{\text{mb}} = 0.5\%$ if no moving bed is detected and 0.75% otherwise. If the moving-bed test is not performed or is invalid, $u'_{\text{mb}} = 1.5\%$. If the reference is not bottom-track, $u'_{\text{mb}} = 0\%$. The uncertainty due to compass bias assumes a one degree standard uncertainty, cf.

Mueller (2018) for details. The vertical extrapolation uncertainty u'_{extrap} is the average of the percent difference in discharge from the selected vertical velocity profile and the closest four options in a list of seven options.

Due to the complexity of the ADCP data workflow, the above mentioned tools do not account for all relevant error sources, in particular errors related to the operator or some of the measurement conditions (Despax et al., 2019), and for the possible correlations among the various uncertainty sources, resulting in simplifications. The input elemental uncertainties are also poorly known. For instance, the ADCP manufacturers do not disclose complete information about the uncertainty of instrumental errors, due to proprietary technologies and operational difficulties in instrument calibrations. The uncertainties due to the measuring environment and the uncertainties due to the operators are even more difficult to model than instrumental errors. Therefore, the input uncertainties are generally derived from site-specific experiments, based on expert judgment or based on available information.

1.3. Repeated Measures Experiments

The uncertainty estimates provided by the propagation methods cannot be validated for in situ conditions because certified, accurate discharge references are lacking in rivers and canals (Despax, Favre, et al., 2016). To solve this issue, a complementary approach to uncertainty propagation methods is the repeated measures experiments, also known as inter-laboratory comparisons (Le Coz et al., 2016). The main steps for conducting inter-laboratory experiments are presented in Figure 2. A repeated measures experiment consists of repeated measurements of the same variable (the discharge) by several participants, or "laboratories," using the same measurement procedure. A "laboratory" is the combination of one or several operator(s) (including their field procedure and settings), their equipment (and associated software), and their measurement site. Each participant a provides k_a discharge values $Q_{a,k}$ from repeated transects, where k denotes the index of the transect.

In compliance with the ISO-5725 standard (ISO, 1994b), the uncertainty of a streamgauging technique can be empirically deduced from the repeated measures experiment in given measurement conditions provided that the discharge is constant (Le Coz et al., 2016). Such experiments have been conducted in France during the last decade (Despax et al., 2017; Hauet et al., 2012; Le Coz et al., 2009; Pobanz et al., 2015, 2011). Final uncertainties are deduced from the variability of all the repeated measurements. For a six-transect discharge average, 95% uncertainty estimates ranged from 4% to 12% typically, depending on the site and measuring conditions (Despax et al., 2019; Le Coz et al., 2016).

1.4. Objectives of This Study

This study presents a propagation method named OURSIN, first developed by Dramais (2011) and then extended and implemented into a software program (Naudet et al., 2019) for computing the uncertainty of measurements acquired with some types of Teledyne RDI ADCPs. The OURSIN method has been modified to be integrated in the QRevInt software (coded in Python) so as to benefit from QA/QC process and QRevInt functionalities for both Teledyne RDI and SonTek ADCP measurements (Section 2). This revised OURSIN method follows the main steps proposed by the GUM (JCGM, 2008a): from the DRE of the moving-boat ADCP discharge measurement (Section 1.1), identify, categorize, and combine error sources to determine the uncertainty of single-transect and transect-averaged discharge measurements (Section 2.1), and evaluate the standard uncertainty of each input quantity (Section 2.2). Several modifications have been made to the original version of OURSIN published by Naudet et al. (2019): the uncertainty due to the projection angle between boat velocity and relative water velocity (cf. Equation 1) is neglected; the uncertainty of the measured discharge now relies on the bottom-track and water-track error velocities and the correlation of velocity errors in the cells of the same ensemble is neglected; less relevant discharge extrapolation scenarios are discarded; the combined uncertainty of multiple transect averaged discharge is modified to account for random and systematic errors; similar to QRev-UA, the moving-bed uncertainty is included and the transect-to-transect discharge variability is reflected through an estimate of the coefficient of variation of single-transect discharges; last, the reference discharge is that computed by QRevInt instead of WinRiverII.

For validation, the method is then applied to two sets of ADCP discharge measurements performed during large-scale repeated measures experiments conducted in 2010 (Le Coz et al., 2016; Pobanz et al., 2011) and in 2016 (Despax et al., 2017, 2019). The uncertainties and uncertainty budgets computed with the OURSIN and the QRev-UA methods are compared with the results of the repeated measures experiments (Section 3). Finally, operational and research perspectives are discussed in Section 4.

Table 1
List of Error Sources in Acoustic Doppler Current Profiler (ADCP) Discharge Measurements Covered by the OURSIN Method, With Their Nature (Systematic, Random, or Both), Type of Uncertainty (Type A or B, See Appendix B), and the Method Used for Their Quantification (See Text for Detail, Configurations Are Described in Table 2)

Error sources	Notation (standard uncertainty)	Nature	Type	Estimation method
Systematic errors of the instrumentation	u_{syst}	Systematic	B	Fixed value (1.31%)
Transect-to-transect variability	u_{CV}	Random	A	Coefficient of variation
Moving-bed	u_{mb}	Systematic	B	Fixed value (0%, 1.5%, or 3%)
Limited number of ensembles	u_{ens}	Systematic	B	Formula (Equation 11 (ISO, 2009))
Measured discharge	u_{meas}	Random	B	Formula (Equation 10) combining the three next components
BT error velocity	$u_{ev}(v_{BT,i})$	Random	A	Standard deviation (included in the u_{meas} component)
WT error velocity	$u_{ev}(w_{WT,i})$	Random	A	Standard deviation (included in the u_{meas} component)
Cell depth	$u(dz_{i,j})$	Random	B	Fixed value $u'(dz_{i,j}) = 0.5\%$ (included in the u_{meas} component)
Interpolation of invalid cells	$u_{invcell}$	Systematic	B	Configurations # 18, 19, 20, 22, 23
Interpolation of invalid ensembles	u_{invens}	Systematic	B	Configurations # 13, 14
Top discharge extrapolation	u_{top}	Both	B	Configurations # 3, 3min, 3max, 5, 5min, 5max, 7, 11, 12
Bottom discharge extrapolation	u_{bot}	Both	B	configurations # 3, 3min, 3max, 5, 5min, 5max, 7
Right discharge extrapolation	u_{right}	Both	B	configurations # 9, 10, 11, 12
Left discharge extrapolation	u_{left}	Both	B	configurations # 9, 10, 11, 12

In this study, the following notations are used.

- u is the absolute standard uncertainty, that is, the standard deviation of the probability distribution of errors, "absolute" meaning expressed in the physical unit of the measurement result (e.g., in m^3/s for discharge);
- u' is the relative standard uncertainty, "relative" meaning expressed in % of the measurement result;
- $U = ku$ is the absolute expanded uncertainty, with k a coverage factor. As recommended by the HUG (ISO, 2007) for hydrometry, we take $k = 2$, which corresponds to a 95% probability interval if the distribution of errors is Gaussian;
- U' is the relative expanded uncertainty expressed in % of the measurement result.

2. The OURSIN Method

2.1. Combination of Uncertainty Components

2.1.1. Individual Transect Combined Uncertainty

The DRE of an ADCP single-transect measurement (Equation 2) is a sum of partial discharges. Treating the corresponding errors as uncorrelated, and applying the first-order Taylor approximation of the GUM (JCGM, 2008a), the combined squared uncertainty $u^2(Q_k)$ of a single-transect discharge measurement Q_k is the quadratic sum of absolute uncertainty components (cf. Appendix A):

$$u^2(Q_k) = u_{syst}^2 + u_{ens}^2 + u_{mb}^2 + u_{meas}^2 + u_{invcell}^2 + u_{invens}^2 + u_{top}^2 + u_{bot}^2 + u_{right}^2 + u_{left}^2 + u_{CV}^2 \quad (5)$$

Table 1 provides the definitions of these uncertainty components and Section 2.2 presents how they are estimated. In addition to the uncertainty components that relate to partial discharges summed in Equation 2, other components (u_{syst}, u_{ens}, u_{mb}, and u_{CV}) are included to reflect error sources that are not directly apparent in the DRE but should be added in the underlying error model.

The relative discharge uncertainty $u'(Q_k)$ is obtained by dividing all the absolute uncertainty terms in Equation 5 by Q_k:

$$u'^2(Q_k) = u'^2_{syst} + u'^2_{ens} + u'^2_{mb} + u'^2_{meas} + u'^2_{invcell} + u'^2_{invens} + u'^2_{top} + u'^2_{bot} + u'^2_{right} + u'^2_{left} + u'^2_{CV} \quad (6)$$

In this equation therefore, the relative uncertainty of a partial discharge associated with an error source is expressed as a percentage of the total discharge, not of the partial discharge.

Table 2
Summary of Parameter Configurations Applied to Estimate the Uncertainty of Unmeasured Discharges (cf. Table 1 For the Correspondence With Uncertainty Components)

Configuration #	Uncertainty components quantified	Tested models	Parameters/extrapolation options
1	All: u'_{top}, u'_{bot}, u'_{right}, u'_{left}, u'_{invens}, $u'_{invcell}$	QRevInt discharge (reference)	See text for detail.
3	u'_{top}, u'_{bot}	Power-power model	QRevInt optimized exponent
3min	u'_{top}, u'_{bot}	Power-power model	Min exponent
3max	u'_{top}, u'_{bot}	Power-power model	Max exponent
5	u'_{top}, u'_{bot}	Constant no-slip model	QRevInt optimized exponent
5min	u'_{top}, u'_{bot}	Constant no-slip model	Min exponent
5max	u'_{top}, u'_{bot}	Constant no-slip model	Max exponent
7	u'_{top}, u'_{bot}	3-point no-slip model	QRevInt optimized exponent
9	u'_{right}, u'_{left}	Edge	Triangular shape and min edge distance
10	u'_{right}, u'_{left}	Edge	Rectangular shape and max edge distance
11	u'_{top}, u'_{right}, u'_{left}	Draft (immersion depth)	Min draft
12	u'_{top}, u'_{right}, u'_{left}	Draft (immersion depth)	Max draft
13	u'_{invens}	Missing ensembles	Use the next valid ensemble
14	u'_{invens}	Missing ensembles	Hold the last valid ensemble
18	$u'_{invcell}$	Missing cells	1/6 power-power model
19	$u'_{invcell}$	Missing cells	Use above valid cell
20	$u'_{invcell}$	Missing cells	Use below valid cell
22	$u'_{invcell}$	Missing cells	Use before valid cell
23	$u'_{invcell}$	Missing cells	Use after valid cell

Table 1 summarizes the classification (type A or B, systematic or random, see Appendix B for definitions) of each uncertainty component and the methods used for their quantification, which will be explained and quantified in Section 2.2 and in Appendix C. The OURSIN method covers systematic and random errors. Each error term is generally a mixture of systematic and random effects which are often difficult to separate. As an approximation, an error term is considered purely systematic or random if the dominant effects are systematic or random, respectively. The uncertainty of discharge in the measured area is estimated based on explicit uncertainty propagation and the discharge variance of successive transects. The discharge uncertainties in unmeasured areas are estimated through sensitivity analysis. By varying some of the parameters, possible discharges are computed. These scenarios are used as an alternative to the Monte Carlo approach and to evaluate the standard uncertainty of discharge in each of the unmeasured areas.

2.1.2. Multiple Transect Averaged Combined Uncertainty

When averaging discharge over successive transects, transect-to-transect random (or uncorrelated) errors are averaged out while systematic (or correlated) errors remain. The uncertainty due to random errors, typically the measured discharge uncertainty (u'_{meas}), is divided by $\sqrt{P}$. Indeed, applying the GUM (JCGM, 2008a) uncertainty propagation method to Equation 3 as the DRE of an ADCP multiple-transect measurement yields for measured discharge errors (cf. Appendix A):

$$u'^2_{meas}\left(\overline{Q}\right) = \frac{1}{P^2}\sum_{k=1}^{P}\frac{Q_k^2}{\overline{Q}^2}u'^2_{meas} \approx \frac{1}{P}\overline{u'^2_{meas}} \tag{7}$$

where the over bar represents the simple average of values from successive transects. To simplify the expression of uncertainty, we will make the approximation which is only exact if the transect discharges are equal. For a good ADCP measurement, transect discharges Q_k are close, and otherwise, the impact would be limited as the contribution of u'_{meas} to the combined discharge uncertainty is usually small. The only other uncertainty component related to transect-to-transect random error is the transect-to-transect discharge variability u'_{CV} which is directly

estimated from the multiple transects. Like u'^2_{meas} in Equation 7, u'^2_{CV} will be divided by P in the expression of $u'^2\left(\overline{Q}\right)$.

The other uncertainty components are due to biases (systematic errors) of the instruments and the discharge extrapolation parameters and options that are constant or at least highly correlated from transect to transect. For instance, the draft bias will produce similar relative top discharge errors for all the transects. The corresponding squared uncertainty components in the expression of $u'^2\left(\overline{Q}\right)$ will not be divided by $\sqrt{P}$. The single-transect estimates of each squared uncertainty due to systematic errors are averaged to compute the multiple-transect mean. Since relative uncertainty is considered here, potential differences in transect discharges Q_k are accounted for in the average.

This leads to the following expression for the uncertainty of a multiple-transect averaged discharge $\overline{Q}$:

$$u'^2\left(\overline{Q}\right) = \overline{u'^2_{syst}} + \frac{1}{P}\overline{u'^2_{meas}} + \overline{u'^2_{ens}} + \overline{u'^2_{mb}} + \overline{u'^2_{invcell}} + \overline{u'^2_{invens}} + \overline{u'^2_{top}} + \overline{u'^2_{bot}} + \overline{u'^2_{right}} + \overline{u'^2_{left}} + \frac{1}{P}u'^2_{CV} \tag{8}$$

See Table 1 and Equation 6 for the definitions of the uncertainty components.

2.2. Discharge Uncertainty Components

2.2.1. Uncertainty Due To Systematic Errors of the Instrumentation

The discharge uncertainty due to systematic errors of the instrumentation, u'_{syst}, accounts for all the residual errors that remain after the ADCP calibration. It corresponds to the minimum uncertainty of an ADCP discharge measurement under ideal conditions. The systematic errors are being treated in aggregate across the transect (Equation 6).

The discharge measured by a moving-boat ADCP is basically the product of boat velocity (bottom-track) and relative water velocity (water-track) (cf. Equation 1 for the discharge through a cell), and extrapolated discharges are proportional to measured discharges. Discharge is also proportional to flow depth used to define the extent of the measured area and the bottom unmeasured area. Therefore, the bottom-track, water-track, and depth biases induce multiplicative errors on the total discharge Q_k of an ADCP transect. These systematic errors can be assumed independent since different acoustic pulses are used for bottom-track and water-track, and different signal processing is used for velocity and depth determination. Then (cf. Appendix A), u'_{syst} can be computed as the combination of the bottom-track bias uncertainty $u'_{syst}(v_{BT})$, the water-track bias uncertainty $u'_{syst}(w_{WT})$, and the depth bias uncertainty $u'_{syst}(D)$:

$$u'_{syst} = \sqrt{u'^2_{syst}(v_{BT}) + u'^2_{syst}(w_{WT}) + u'^2_{syst}(D)} \tag{9}$$

Note that the boat velocity is referenced either by bottom-tracking or by GPS. Only bottom-track reference is considered here. The potential uncertainty associated with the use of GPS is discussed in Section 4.3.

Based on mean differences between tow cart velocity and ADCP bottom-track and water-track velocities observed by Oberg and Mueller (2007) in a calibration tow-tank, the following values are used: $u'(v_{BT}) = 0.51\%$ and $u'(w_{WT}) = 1.1\%$. Based on tests conducted in canals and locks (Naudet et al., 2019), the depth bias uncertainty is assumed to be $u'(D) = 0.5\%$, by default. Those tests were conducted using the four-beam, inverse-depth-weighted average depth which is the default method in QRevInt and which is always used across all instruments in this study. The value of $u'(D)$ should be revised if better information is available or if the vertical beam depth measured by some ADCP models is used instead of the four-beam average depth, for instance. With such values, Equation 9 yields $u'_{syst} = 1.31\%$. This value is similar to the default value of 1.5% proposed in the QRev-UA method (Mueller, 2016) and to the bias uncertainty ($u'(\delta) = 1.25\%$) quantified from the Génissiat 2010 repeated-measures experiments described by Le Coz et al. (2016). Across the successive experiments, the empirical estimates ranged from 1.1%–1.2% (Pyrimont site) to 1.4%–1.9% (Génissiat site).

2.2.2. Uncertainty of the Measured Discharge

The uncertainty of the measured discharge is computed by applying variance propagation (cf. Appendix A) to Equation 1 with a number of approximations aiming at simplifying the derivation. We indeed assume that cell

discharges, water track velocity uncertainty, depth cell size uncertainty are the same for all cells in the same ensemble and that the bottom track velocity error is constant (or perfectly correlated). We neglect the measurement uncertainty of the time interval dt_i and assume that discharge is constant over the interval, typically 1 s or less. We also assume that errors in adjacent ensembles and cells are uncorrelated. The water and bottom track pings are emitted, received, and processed independently and thus from the standpoint of the ADCP signal processing, there is no correlation between the successive water and bottom track measurements. As presented in the previous section, the uncertainty due to systematic errors (biases) of the water and bottom track are accounted for in the u'_{syst} component. There is however a correlation of measurements between cells within an ensemble because the center weighted averaging of the cells includes some portion of the acoustic signal measured in the cells above and below (15% from each adjacent cell, according to Simpson and Oltmann (1991)). Due to the large number of cells in an ADCP transect, the measured discharge uncertainty would remain minor even if such partial correlation was considered. Actually, all these assumptions have no practical consequences because the uncertainty results are much more sensitive to the evaluation of the input terms of the uncertainty equation than its approximation. And fortunately, the discharge uncertainty u'_{meas} due to random errors in the measured area is very small in the discharge uncertainty budget of nearly all ADCP measurements.

With the aforementioned assumptions, this uncertainty term can be computed as:

$$u'_{meas} = \sqrt{\frac{1}{Q_k^2}\sum_{i=1}^{n} q_i^2 \left\{ u'^2_{ev}(v_{BT,i}) + \frac{1}{m_i}\left[u'^2_{ev}(w_{WT,i}) + u'^2(dz_{i,j})\right]\right\}} \quad (10)$$

where q_i is the discharge of each ensemble, $u'_{ev}(v_{BT,i})$ and $u'_{ev}(w_{WT,i})$ are the relative uncertainty components of the boat and water-track velocity, respectively, m_i is the number of cells of ensemble i and $u'(dz_{i,j})$ is the relative uncertainty of the depth cell size. The $u'(dz_{i,j})$ component is assumed to be the same as the uncertainty of the water depth: $u'(dz_{i,j}) = u'(D) = 0.5\%$, since both are measured from the transit-time of sound waves.

The $u'_{ev}(v_{BT,i})$ and $u'_{ev}(w_{WT,i})$ components are estimated as the coefficients of variation of the corresponding error velocity terms across the ensembles of a transect. The computation of the error velocity is explained in Appendix C. Boat velocity and water velocity errors are affected by ADCP motion (erratic deployment) and measuring conditions (turbulence, flow instability, waves). The error velocity terms reflect the homogeneity of the velocity field across the four ADCP beams and provide a meaningful estimator of the horizontal velocity measurement uncertainty (Gilcoto et al., 2009; Moore et al., 2016; Teledyne RDI, 1998, 2007).

In Equation 10, the relative uncertainties of single-cell measurements ($u'_{ev}(w_{WT,i})$, $u'(dz_{i,j})$) are weighted by $1/m_i$ and q_i while the relative uncertainties of single-ensemble measurements ($u'_{ev}(v_{BT,i})$) are only weighted by q_i. This means that the former uncertainties average out relative to the total number of valid cells in the transect whereas the latter uncertainties average out relative to the number of ensembles in the transect.

2.2.3. Uncertainty Due To the Limited Number of Ensembles

The limited number n of ensembles induces possible errors in the transverse integration of discharge (González-Castro & Muste, 2007), similar to the finite summation error in measurements with current-meters. Therefore, the associated uncertainty is computed following the equation proposed by Le Coz et al. (2012) as an approximation of tabulated values proposed by ISO (2009):

$$u'_{ens} = 32 \times n^{-0.88} \quad (11)$$

If more than 100 ensembles are collected in an ADCP transect, as is typical, then standard uncertainty u'_{ens} would be less than 0.5%.

2.2.4. Uncertainty of Unmeasured Discharges

Discharges estimated in unmeasured areas depend on the equations selected to extrapolate or interpolate the discharge. Some equations have parameters specified by the user or fitted by the QRevInt software. Thus, the uncertainty of unmeasured discharges includes structural uncertainty and parametric uncertainty. It is estimated by varying the equations and their parameters in order to simulate the realistic range of discharge variation.

Table 2 summarizes the parameter configurations applied by OURSIN to compute alternative discharges. These configurations are designed to explore the broader range of discharge estimates in each of the unmeasured areas,

as presented in Table 1, with reasonable sets of parameters and options. The reference configuration #1 corresponds to the default options of QRevInt (Mueller, 2016, 2021): optimized exponent and model for top/bottom discharge extrapolation, user-defined values of edge parameters and ADCP draft, and ABBA interpolation of missing cells (i.e., using measurements available above, or below, or before, or after the missing cell). To save computational time, not all the possible situations are sampled (as opposed to Monte Carlo simulation), and sampling is limited to a small number of configurations that were designed based on expert judgment. The maximum and minimum discharges obtained from the set of configurations specified in Table 1 for each error source are assumed to be the bounds of a rectangular distribution from which the standard uncertainty is estimated (see Equation B2). Even when the uniform distribution is not centered on the measured value, the measurement result is not replaced by the expected value $(a_+ + a_-)/2$ of the uniform distribution because the measured value remains the best estimate of the true value.

2.2.4.1. Top and Bottom Discharge Errors

The top and bottom discharge uncertainties are estimated through configurations (#1, #3, #3min, #3max, #5, #5min, #5max, #7 and #11, #12 specifically for top discharge uncertainty) presented in Table 2, which cover the draft error and alternative models to extrapolate the discharge. Three alternative models are available in QRevInt for computing vertical velocity distribution and extrapolating top/bottom discharges.

The possible extrapolation models for the top layer are power fit, constant fit, or three-point linear extrapolation. The bottom fit can be based on either a power or a no-slip model. For top and bottom power fit, a power model is applied to extrapolate the unmeasured areas. The velocity profile $w(z)$ at elevation z above the bed is expressed as (Chen, 1989):

$$w(z) = w(h)\left(\frac{z}{h}\right)^{1/m'} \tag{12}$$

where $w(z)$ is the flow velocity measured at elevation z, h is the total flow depth, and $1/m'$ is the power law exponent. In rivers, the exponent usually ranges from 1/3 to 1/10 (Hauet et al., 2018).

The no-slip model fits a power curve through zero at the bottom (solid boundary) and through depth cells in the lower 20% of the flow depth. A constant fit for estimating the top discharge assumes that the velocity in the topmost valid depth cell is the mean velocity until the water surface.

The three-point model is a linear fit of the velocities in the three uppermost cells. It fits situations where wind or other effects significantly affects the velocity at the water surface, causing the velocity at the surface to deviate substantially from either a constant or power fit.

For all models, the coefficients used are the optimized exponents computed by the Extrap module in QRevInt (Mueller, 2016). In addition, power-power and constant no-slip exponents may range between a lower and an upper bound for the optimized exponent. Note that a similar approach was conducted by Gordon (1989) as an early attempt to quantify the uncertainty of ADCP measurement made in River Elbe and by Mueller (2016) in the QRevInt software. In the OURSIN method, lower and upper bounds for the optimized exponent (min-max exponents) are estimated as follows.

- **Power-power exponent** (configurations #3min and #3max): If QRevInt does not select the power-power model for any of the transects, only the optimized power-power exponent is used. The minimum-maximum power-power configurations are not tested since the power-power method is not relevant for any of the transects. Otherwise, the minimum and the maximum exponents are the average of the lower and the upper bound 95% confidence intervals provided by QRevInt, respectively. If the flow is bi-directional the power-power method is not considered at all.
- **Constant no-slip exponent** (configurations #5min and #5max): If QRevInt does not select the constant no-slip model for any of the transects, the optimized no-slip exponent is only used. The minimum-maximum no-slip configurations are not tested since it has not been selected as a relevant model for any of the transects. Otherwise, the minimum and the maximum optimized exponents (only from transects for which constant no-slip model is proposed) constitutes the lower and the upper bound of $1/m'$, respectively.

The maximum difference between the lower/upper bounds of the exponent interval and the average of all optimized coefficients is limited at 0.2 for both power-power and constant no-slip models.

The transducer draft, that is, the immersion depth of the ADCP probe, is also varied as it affects the top discharge estimation. The user specifies the maximal draft error further added or subtracted to the draft value to define the minimum and maximum values for configurations #11 and #12, respectively. By default, the draft error is 0.02 m if the transect depth does not exceed 2.5 m, 0.05 m otherwise. The minimum draft used in the configuration #11 is at least 0.01 m.

2.2.4.2. Edges

The edge (left or right) discharge is expressed as:

$$Q_{\text{edge}} = C_e \, V_m \, L \, d_m \tag{13}$$

where C_e is the edge-shape coefficient (0.3535 for triangular edges and 0.91 for rectangular edges), V_m is the mean depth-averaged velocity over a fixed number of valid ensembles at the start or end of the measured area, L is the distance from the first or last valid ensemble to the edge of water, and d_m is the depth at first or last valid ensemble.

The edge (left or right) discharge uncertainty is estimated by varying C_e, L, and the transducer draft as also done for the top discharge uncertainty estimation. The edge discharge will be minimized when a triangular shape is selected and the edge distance is decreased concurrently, while the discharge will be maximized with a rectangular shape and an increased edge distance, provided the near edge velocities are positive. The default edge distance error used in the OURSIN method is 20% of the edge distance. This relatively conservative default setting corresponds to a vague edge estimation, as often practiced in France for instance when edge discharges are assumed to be relatively small: edge distances are eye-balled and the edge shape is not measured. If the edges are determined more carefully in the field procedure, for example, using a ruler or laser range finder to measure edge distances and knowing edge shapes precisely, then the maximum edge distance and edge-shape coefficient errors should be reduced accordingly.

2.2.4.3. Invalid Data

The discharge of a transect with invalid data is computed from interpolated values of boat velocity, depth, and water velocity.

To evaluate the uncertainty due to invalid ensembles, configurations #13 and #14 replace the missing variables by those of the next or the previous valid ensembles, respectively.

To evaluate the uncertainty due to invalid cells, configurations #18, #19, #20, #22, and #23 interpolate missing discharges based on 1/6 power-power model, from the nearest valid cell above, below, before, or after, respectively.

2.2.5. Discharge Uncertainty Due To Moving Bed

Since bottom tracking assumes that the streambed is motionless, moving bed conditions result in the boat-velocity measurement being biased in the upstream direction. Moving bed conditions occur due to sediment transport near and along the streambed, especially during high flows. The uncertainty due to moving bed $\left(u'_{\text{mb}}\right)$ in OURSIN is assumed to be the same as in QRev-UA (cf. Section 1.2).

2.2.6. Discharge Uncertainty Due To Transect-To-Transect Variability

ADCP discharge measurements are usually the average of a number P of individual discharge measurements from successive ADCP transects. Typically, the USGS recommends performing at least two reciprocal transects with a minimum exposure time of 720 s (Mueller et al., 2013; Oberg & Mueller, 2007), while in France, a minimum of six transects acquired during at least 900 s is recommended (Le Coz et al., 2008). Of course, acquiring fewer transects is advisable for very large rivers to keep the measurement duration reasonable, or for transient flows. The repeatability of successive discharge results, through their coefficient of variation, is a valuable measure of the overall uncertainty of the transect-averaged discharge measurement. The coefficient of variation CV is the standard deviation of the P single-transect discharge divided by the mean discharge $\overline{Q}$.

These combined random errors that appear among successive transects certainly overlap with some errors already included in the discharge error model. However, the coefficient of variation reflects other errors that are ignored despite being potentially large, especially velocity and discharge variations due to turbulence or flow unsteadiness during a measurement, or other site-related effects not captured by the uncertainty components presented

above. It therefore looks conservative to include an additional uncertainty component u'_{CV} empirically estimated from the coefficient of variation of successive transects (Type A uncertainty):

$$u'_{CV} = \frac{CV}{\sqrt{P}} \tag{14}$$

Since underlying errors are assumed independent among successive transects, CV is divided by the square root of the number P of single-transect discharges included in the mean discharge $\overline{Q}$.

As presented in Section 1, QRev-UA (Mueller, 2016) includes a similar term u'_{cov} (cf. Equation 4) with a correction factor suggesting that some prior knowledge is needed to guide the estimation of the coefficient of variation when there are too few transects available. We propose a more formal and more general solution to this problem by applying Bayesian inference to the estimation of the coefficient of variation. The Bayesian approach allows the combination of information from the successive transects and a prior estimation of CV, which will dominate the inference when few, or even one or two transects only are available. Successive discharge measurements are assumed to follow a Gaussian distribution:

$$\tilde{Q}_i \sim \mathcal{N}(Q_{\text{true}}, CV \times Q_{\text{true}}) \tag{15}$$

where $\mathcal{N}(a, b)$ is the Gaussian distribution with mean a and standard deviation b. Both the true discharge Q_{true} and the coefficient of variation CV are unknown parameters to be estimated.

The prior distributions of Q_{true} and CV are assumed to be flat (no prior knowledge) and log-normal, respectively. The prior pdf therefore is:

$$p(Q_{\text{true}}, CV) = 1 \times p_{LN}\left(m_{\text{CVprior}}, s_{\text{CVprior}}\right) \tag{16}$$

where $p_{LN}(z \mid m, s)$ is the pdf of a Lognormal random variable whose logarithm has mean equal to m and standard deviation equal to s, evaluated at value z.

The likelihood of observed discharge values $\tilde{\mathbf{Q}}$, given discharge parameters (Q_{true}, CV) is given by:

$$p\left(\tilde{\mathbf{Q}} \mid Q_{\text{true}}, CV\right) = \prod_{i=1}^{P} p_N\left(\tilde{Q}_i \mid Q_{\text{true}}, CV \times Q_{\text{true}}\right) \tag{17}$$

where $p_N(z \mid m, s)$ is the pdf of a Gaussian distribution with mean m and standard deviation s, evaluated at value z.

Using Bayes theorem, the posterior distribution of the discharge parameters can be computed as follows:

$$\underbrace{p\left(Q_{\text{true}}, CV \mid \tilde{\mathbf{Q}}\right)}_{\text{posterior}} \propto \underbrace{p\left(\tilde{\mathbf{Q}} \mid Q_{\text{true}}, CV\right)}_{\text{likelihood}} \underbrace{p(Q_{\text{true}}, CV)}_{\text{prior}} \tag{18}$$

The posterior distribution is sampled (20,000 realizations) using an adaptive block Metropolis Markov Chain Monte Carlo (MCMC) sampler described by Metropolis et al. (1953). The computation time is a few seconds at most so the MCMC sampling does not slow down QRevInt significantly. The default simulation parameters ensure that, in the vast majority of situations, convergence is reached after the first half of the simulations have been discarded. Convergence can be checked visually: the MCMC sample traces are not stationary and explore the full posterior distribution. The sample that maximizes the posterior density is taken as the best estimates of Q_{true} and CV. The resulting estimate of CV is used in the computation of u'_{CV} (Equation 14). The estimate of Q_{true} is not used.

In the analysis presented in this paper, the priors of the coefficient of variation are taken as $m_{\text{CVprior}} = \log(0.03)$ and $s_{\text{CVprior}} = 0.6$. As the prior distribution is lognormal, the mean is $\exp\left(m_{\text{CVprior}} + s^2_{\text{CVprior}}/2\right) = 3.6\%$ with an uncertainty of about 60%. Through Equation 14, the expected CV of 3.6% leads to $u'_{CV} = 3.6\%$ and 1.5% for a single-transect measurements and for a six-transect-averaged discharge measurement, respectively, that is, 95% uncertainty intervals of half-width 7.2% and 3.0%.

These default prior values of CV have been specified to be consistent with the repeatability standard deviations (s_r) obtained from interlaboratory experiments in various site conditions (Despax et al., 2019; Le Coz

et al., 2016). They are also consistent with the experimental values reported by Huang (2018) (see his Equation 25 and Table 1, his "Field RSUm" being similar to our *CV*). Based on sensitivity tests (cf. Supporting Information S1), $s_{CVprior}$ was set low enough so the *CV* estimate is influenced by the priors when the number of transects is very low ($P = 2$, typically), and high enough so the *CV* estimate converges toward the empirical *CV* estimates as soon as the number of transects gets high enough ($P = 6$, typically). Of course, the prior values of *CV* ($m_{CVprior}$ and $s_{CVprior}$) can be adapted by the user to reflect his/her expert judgment of the actual measuring conditions and the corresponding values of *CV* that may be expected.

3. Comparison With Repeated Measures Experiments

3.1. Overview of 2010 and 2016 Experiments

Two large repeated measures experiments were conducted in France in 2010 and 2016. The first one was conducted in 2010 at two sites (named GE: Génissiat dam and PY: Pyrimont bridge), approximately 50–200 m and 3.5 km, respectively, downstream of Génissiat hydropower plant in the Rhône River. During the 2-day experiment, six time periods with different steady discharge released from the hydropower plant were made available (sessions #1, #2, #2b, #3, #4, and #4b). Six models of ADCP, mounted on boats, were involved in 2010: 1 BroadBand 600 kHz, 6 RioGrande 600 kHz, 6 RiverRay 600 kHz, 11 RioGrande 1,200 kHz, and 3 StreamPro 2,400 kHz ADCPs all made by Teledyne RDI and 2 M9 1,000–3,000 kHz ADCPs made by SonTek. The cross-section width ranged from 50 to 70 m at GE, and from 75 to 80 m at PY. The depth ranged from 3 to 5 m and from 5 to 9 m, at PY and GE, respectively. The cross-section average velocity ranged from 0.5 m/s to 1.6 m/s (the maximum velocity was about 2.5 m/s).

The second large experiment was conducted on the Taurion River, downstream of the Chauvan Dam in southwest France in 2016. A single constant discharge was released by the dam during three half-day measurement sessions. Four models of ADCPs were involved: 18 M9 and 1 S5 (3,000 kHz) ADCPs made by SonTek, and 25 StreamPro and 4 RiverPro (1,200 kHz) ADCPs made by Teledyne RDI. A total of 24 cross-sections (from A to X, upstream to downstream) with various shapes and flow conditions were distributed over 500 m along the Taurion River. The 24 cross-sections were equipped with ropes and pulleys. Most cross-sections were about 35 m wide. The depth ranged from 0.6 to 1.2 m. The cross-section average velocity ranged from 0.5 m/s to 0.8 m/s (the maximum velocity was about 1.5 m/s).

Bottom-track was used as the reference to compute the boat speed for both Génissiat 2010 and Chauvan 2016 experiments. A total of 634 four-transect and 574 six-transect discharge measurements were performed in 2010 and in 2016, respectively. For the Génissiat 2010 experiments, a power-power model was proposed for extrapolation of top and bottom discharges consistently by QRevInt. At GE, the mean exponent ranges from 0.15 during session #2b (440 m^3/s) to 0.21 during session #3 (120 m^3/s). At PY, the mean exponent is about 0.20 during all the sessions. A power-power model was imposed for all measurements of Chauvan 2016 experiments, with an average exponent fitted at each cross section. The exponent used ranges from 0.15 at cross-section X to 0.37 at cross-section T. The mean exponent over all the cross-section is 0.29. Further details on site description and experimental design can be found in Pobanz et al. (2011), Le Coz et al. (2016) and Despax et al. (2017, 2019), respectively.

3.2. Empirical Uncertainty Estimates

The empirical uncertainty results for the two repeated measurements experiments are the same as previously published by Le Coz et al. (2016) and Despax et al. (2019). The six time periods with steady discharge and the two sites (GE and PY) of the Génissiat 2010 experiments lead to 12 individual interlaboratory experiments, and the 24 cross-sections (A to X) of the Chauvan 2016 experiments yield 24 separate interlaboratory experiments (Table 3). The repeatability standard deviation (s_r), interlaboratory standard deviation (s_L), and expanded relative uncertainty $U'\left(\overline{Q}\right)$ are computed for each separate experiment. The meanings and the computation of these uncertainty components are explained by Le Coz et al. (2016). Remember that in OURSIN, the prior values of the coefficient of variation are defaulted based on typical repeatability

Table 3
Data of the Génissiat 2010 and the Chauvan 2016 Acoustic Doppler Current Profiler (ADCP) Repeated Measures Experiments Copied From Le Coz et al. (2016) and Despax et al. (2019), Respectively: Number of Laboratories (Lab.), Transects (Tr.), Transect-Averaged Measurements (Meas.), Mean and Standard Deviation (Std.) of the Discharge, Mean Measured to Total Discharge Ratio (η), and Empirical Uncertainty Estimates (ADCP Bias Uncertainty $u'(\hat{\delta})$, Repeatability (s_r), Interlaboratory s_L Standard Deviations, and 95% Expanded Discharge Uncertainty $U'(\overline{Q})$)

	Number of			Discharge [m³/s]		Ratio, η [%]	Empirical uncertainty [%]			
Site name	Lab.	Tr.	Meas.	Mean	Std.	Mean	$u'(\hat{\delta})$	s_r	s_L	$U'(\overline{Q})$
GE_1	11	388	97	221.7	11.9	70	1.25	4.1	3.5	8.5
GE_2	9	184	46	333.8	18.5	73	1.25	4.4	3.8	9.2
GE_2b	11	352	88	438.9	20.5	73	1.25	3.6	3.8	8.7
GE_3	8	236	59	114.4	9.0	67	1.25	5.4	5.3	12.2
GE_4	5	104	26	224.5	12.2	68	1.25	4.5	4.3	10.0
GE_4b	8	216	54	330.6	18.0	72	1.25	4.9	3.9	9.5
PY_1	10	188	47	220.7	5.4	68	1.25	2.1	2.2	5.5
PY_2	12	172	43	331.7	9.4	73	1.25	2.9	2.0	5.5
PY_2b	10	212	53	432.8	12.6	74	1.25	2.1	1.8	4.9
PY_3	12	184	46	117.9	2.6	67	1.25	2.6	1.8	5.1
PY_4	12	148	37	227.8	6.8	67	1.25	2.8	2.3	5.9
PY_4b	9	152	38	332.3	7.2	72	1.25	1.9	1.8	4.8
A	24	144	24	14.76	0.31	56	1.25	2.2	1.9	4.6
B	24	144	24	14.86	0.34	55	1.25	1.7	2.2	5.0
C	24	144	24	14.71	0.41	56	1.25	1.9	2.7	6.0
D	24	144	24	14.87	0.53	54	1.25	1.9	3.5	7.5
E	24	144	24	14.64	0.53	35	1.25	3.6	2.3	7.5
F	24	144	24	14.59	0.54	38	1.25	2.7	3.6	7.8
G	24	144	24	14.93	0.40	47	1.25	2.2	2.6	5.8
H	24	144	24	15.00	0.37	50	1.25	1.9	2.4	5.4
I	24	144	24	14.83	0.40	53	1.25	2.2	2.6	5.8
J	24	144	24	14.94	0.48	40	1.25	2.1	3.1	6.7
K	24	144	24	14.70	0.45	41	1.25	2.4	2.9	6.5
L	24	144	24	14.62	0.41	49	1.25	1.7	2.7	6.0
M	24	144	24	14.83	0.25	57	1.25	1.9	1.5	4.0
N	24	144	24	14.58	0.34	49	1.25	1.9	2.2	5.1
O	23	138	23	14.57	0.36	52	1.25	2.4	2.3	5.4
P	24	144	24	14.72	0.31	55	1.25	1.9	1.9	4.6
Q	24	144	24	14.73	0.31	58	1.25	2.1	1.9	4.7
R	24	144	24	14.74	0.31	58	1.25	2.0	1.9	4.6
S	24	144	24	14.65	0.38	54	1.25	2.0	2.4	5.5
T	23	138	23	14.73	0.35	53	1.25	1.8	2.3	5.2
U	24	144	24	14.62	0.28	61	1.25	1.8	1.7	4.3
V	24	144	24	14.97	0.26	54	1.25	1.6	1.6	3.9
W	24	144	24	14.58	0.26	59	1.25	1.6	1.7	4.1
X	24	144	24	14.99	0.29	56	1.25	1.6	1.8	4.3

standard deviations obtained in repeated measures experiments. The empirical expanded uncertainty $U'\left(\overline{Q}\right)$ is computed as:

$$U'\left(\overline{Q}\right) = k\sqrt{\frac{s_r^2}{PN_p} + \frac{s_L^2}{N_p} + u'^2\left(\hat{\delta}\right)} \tag{19}$$

where N_p is the number of participants in an experiment and $u'\left(\hat{\delta}\right)$ is the uncertainty related to the ADCP method bias, assumed to be 1.25% as proposed by Le Coz et al. (2016) for Génissiat 2010 experiments. In accordance with the procedures followed by the participants, $U'\left(\overline{Q}\right)$ is computed for four-transect ($P = 4$) and six-transect ($P = 6$) averaged discharge measurements for Génissiat 2010 and Chauvan 2016 experiments, respectively.

The expanded uncertainty for a four-transect measurement at PY (4.8%–5.5%) is half of the uncertainty at GE (8.5%–12.2%), across the range of discharges from 120 to 440 m^3/s. The difference is explained by more favorable measurement conditions at PY compared to GE, which presents an unstable flow field with marked secondary currents and intermittent macroturbulent structures (Le Coz et al., 2016).

In the Chauvan 2016 experiments, the expanded uncertainty for a six-transect measurement ranges between 3.9% and 7.8% at cross-sections V and F, respectively. Despax et al. (2019) report that the difference is mostly explained by the measured discharge ratio, which suggests that larger uncertainty is due to discharge extrapolation errors. The cross-section V presents a highest measured discharge ratio (54%) compared to cross-sections E and F (35% and 38%, respectively). A relatively high uncertainty (7.5%) is also observed at cross-section D where the near-bank unmeasured area is important (between 3% and 12% of the total discharge is extrapolated at the left edge, among the 24 measurements).

3.3. Comparison of Computed and Observed Uncertainties

Prior to uncertainty calculations, all the ADCP data were post-processed and reviewed using QRevInt 1.18 (Mueller, 2021). Such post-processing using QRevInt ensures that all the ADCP data and discharges are processed homogeneously, irrespective of their manufacturers. The uncertainty of each discharge measurement was computed following the QRev-UA and OURSIN methods using QRevInt 1.18 and the default parameter values, except $s_{CVprior} = 0.6$ instead of 0.2.

The 95% expanded discharge uncertainty estimates $U'\left(\overline{Q}\right)$ of QRev-UA and OURSIN are compared with the empirical estimation of uncertainty from the repeated measures experiments (Figure 3), and the cross-section by cross-section detail is presented in Figure 4. Four-transect and six-transect average discharge measurements are considered for Génissiat 2010 and Chauvan 2016 experiments, respectively. The 95% expanded uncertainty of the empirical uncertainty estimates computed by Le Coz et al. (2016) and Despax et al. (2019) is reported as error bars in the two figures. This uncertainty of the uncertainty reflects the limited number of instruments and repeated measurements included in an experiment. In QRev-UA and OURSIN, the uncertainty due to moving-bed conditions is assumed to be negligible ($u_{mb} = 0$), since some moving-bed tests were conducted and no discernible moving bed was observed for both experiments.

Despite the substantial variability in the computed uncertainty of individual measurements, significant trends in the responses of QRev-UA and OURSIN uncertainty results appear clearly (cf. Figure 4). Overall, the two methods produce realistic uncertainty levels and the increase of uncertainty with more difficult measuring conditions is generally captured, even though the larger uncertainty of the measurements conducted at GE site is underestimated. This is less true for low-uncertainty measurements (PY cross-sections and Chauvan 2016 experiment) for which QRev-UA seems to yield nearly constant uncertainty (about 4% at PY cross-sections and 5%–6% at Chauvan). Both methods seem to lack "dynamic range" for the Chauvan experiments, though QRev-UA is worse and the mean for OURSIN is better.

For Génissiat 2010 experiments, the fit of the ordinary least squares linear regression is good (coefficient of determination $R^2 > 0.9$) for QRev-UA and OURSIN (cf. Figure 3). However, QRev-UA underestimates the empirical uncertainty for both GE and PY cross-sections more substantially than OURSIN. The mean and the 95% quantiles of the uncertainty differences from all the empirical results presented in this study are −1.5% [−5.3%; +1.7%] and −3.1% [−7.6%; +0.6%] for OURSIN and QRev-UA, respectively.

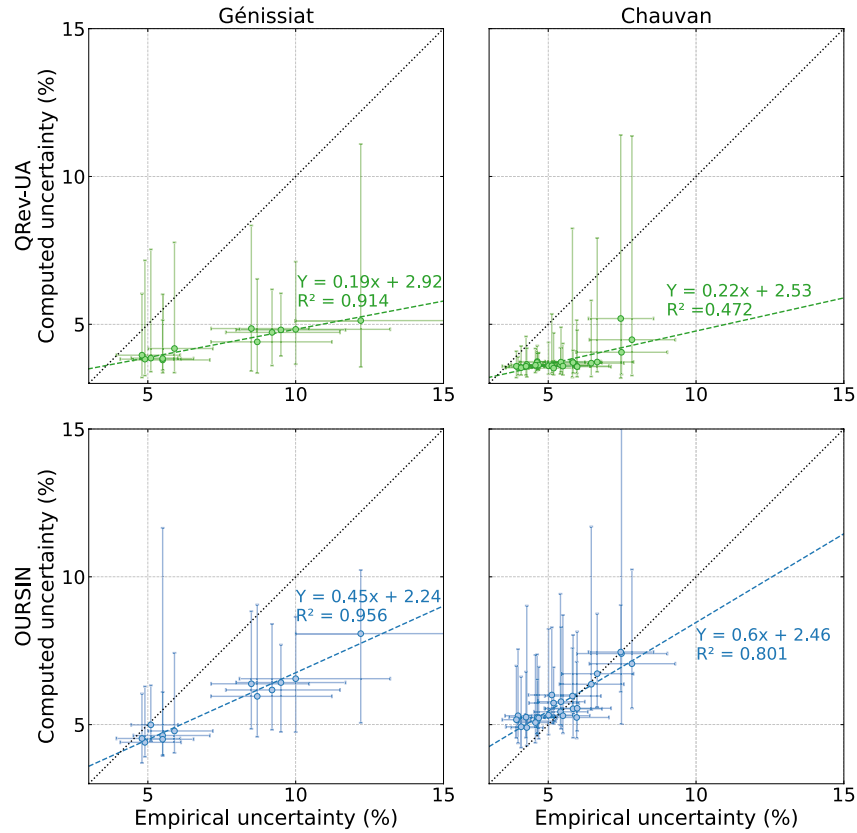

Figure 3. Comparison of the median of the 95% expanded discharge uncertainty $U'\left(\overline{Q}\right)$ estimated using QRev-UA and OURSIN methods and the empirical uncertainty estimates from the repeated measures experiments of Génissiat 2010 and Chauvan 2016. Horizontal and vertical error bars show the 95% uncertainty of empirical uncertainty estimates and the 95% quantiles of uncertainties computed for individual ADCP measurements, respectively. The equation and coefficient of determination of the ordinary least squares linear regression are displayed.

As for Chauvan 2016 experiments, linear regression is more difficult to interpret because of the smaller range of uncertainty explored (cf. Figure 3). But again, QRev-UA uncertainty appears to be underestimated whereas OURSIN uncertainty deviations are more balanced around the line of perfect agreement. The mean and the 95% quantiles of the uncertainty differences from all the empirical results presented in this study are +0.5% [−1.4%; +3.1%] and −1.5% [−3.9%; +0.2%] for OURSIN and QRev-UA, respectively.

For both experiments, the empirical confidence intervals most often overlap the distributions of the OURSIN uncertainty estimates, which is not true for QRev-UA (cf. Figure 4). The mean and the 95% quantiles of the uncertainty differences from all the empirical results presented in this study are −0.4% [−4.4%; +2.5%] and −2.2% [−6.9%; +0.3%] for OURSIN and QRev-UA, respectively. As mentioned earlier, in OURSIN some random errors may be accounted for twice as the same effects may contribute to both the measured discharge uncertainty u'_{meas} and the coefficient of variation term u'_{CV}. The comparison with empirical results suggests that the possible double-counting effect does not create significant overestimation of the final uncertainty, at least in the measuring conditions of the reported experiments, and unless other terms are significantly underestimated.

3.4. Analysis of the Variance Decomposition

From Equations 6 and 8, the contribution of each uncertainty component u to the uncertainty budget can be assessed as the ratio between the variance u^2 and the combined variance $u^2(Q_k)$ or $u^2\left(\overline{Q}\right)$. The expanded

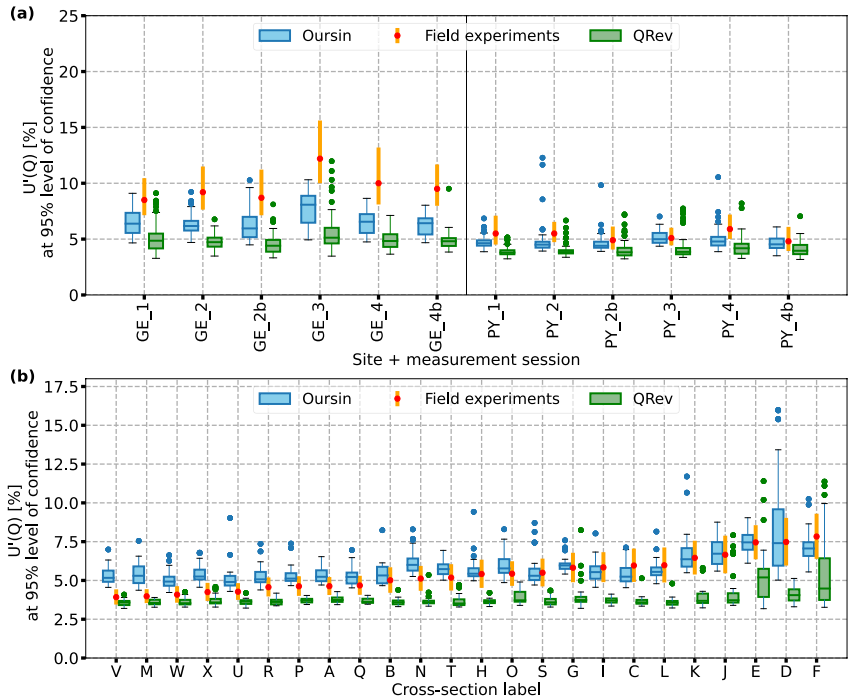

Figure 4. Boxplot of the 95% expanded discharge uncertainty $U'\left(\overline{Q}\right)$ estimated using QRev-UA and OURSIN methods and the empirical uncertainty estimates (red dots) with their 95% interval of confidence (orange solid line) at each cross-section: (a) Génissiat 2010 experiment and (b) Chauvan 2016 experiment. The boxes show the median, first, and third quartiles of the uncertainty computed over the discharge measurements, while the whiskers show 1.5 times the inter-quartile distance. Outliers are represented by dots. Note that the Chauvan cross-sections are sorted based on the magnitude of empirical uncertainty, from the least (V) to the most uncertain cross-section (F).

uncertainty U (at a 95% probability level) is the combined relative uncertainty u' multiplied by the coverage factor $k = 2$.

Figure 5 and Table 4 show the mean decomposition of the total variance into each uncertainty source for QRev-UA and OURSIN methods (mean results of all the participants at each cross-section) as follows:

- Systematic errors related to instrumentation: $u'^2_{\text{syst}}/u'^2\left(\overline{Q}\right)$ with $u'_{\text{syst}} = 1.5\%$ and 1.3% in QRev-UA and OURSIN, respectively.
- Measured discharge: $(1/P)\,\overline{u'^2_{\text{meas}}}/u'^2\left(\overline{Q}\right)$ (only for OURSIN).
- Limited number of ensembles: $\overline{u'^2_{\text{ens}}}/u'^2\left(\overline{Q}\right)$ (only for OURSIN).
- Moving bed: $\overline{u'^2_{\text{mb}}}/u'^2\left(\overline{Q}\right)$ (neglected in this study).
- Invalid data: $\left(\overline{u'^2_{\text{invcell}}} + \overline{u'^2_{\text{invens}}}\right)/u'^2\left(\overline{Q}\right)$ (QRev-UA does not separate invalid cells from invalid ensembles).
- Top/bottom discharge extrapolation: $\left(\overline{u'^2_{\text{top}}} + \overline{u'^2_{\text{bot}}}\right)/u'^2\left(\overline{Q}\right)$ (QRev-UA does not separate top and bottom extrapolated discharges).
- Edge discharge extrapolation: $\left(\overline{u'^2_{\text{right}}} + \overline{u'^2_{\text{left}}}\right)/u'^2\left(\overline{Q}\right)$ (QRev-UA does not separate right and left extrapolated discharges).
- Transect-to-transect variability: $(1/P)\,\overline{u'^2_{\text{CV}}}/u'^2\left(\overline{Q}\right)$ (coefficient of variation).

As already seen, for the two experiments and at the different cross-sections, OURSIN uncertainty results are in closer agreement than QRev-UA uncertainty results with the empirical uncertainty estimates derived from the intercomparison experiment results. By construction, the variance decomposition proposed by OURSIN is more detailed than that proposed by QRev-UA but some components can be compared across the two methods. Substantial differences appear in the variance decomposition results of the two methods.

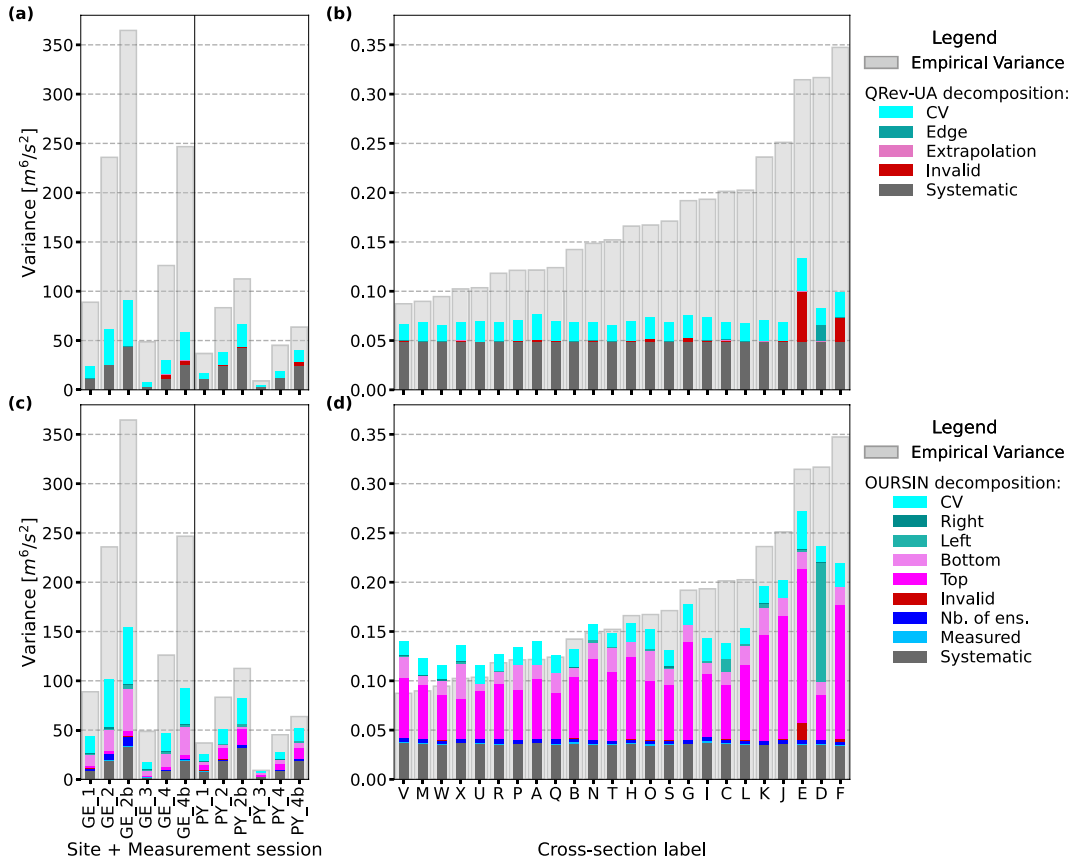

Figure 5. Decomposition of the mean variance computed by QRev-UA (a and b) and OURSIN (c and d) into variance due to each error source for Génissiat 2010 (left) and Chauvan 2016 (right) repeated measures experiments. The mean variance results at each cross-section are presented. In the background, light gray bars show the empirical variance as shown in Figure 3.

At GE cross-sections for instance, QRev-UA uncertainty is dominated by the transect-to-transect discharge variability (through the coefficient of variation). The OURSIN uncertainty results are dominated by the same component and by the bottom discharge uncertainty. At Chauvan, the contributions of the transect-to-transect discharge variability are similar in magnitude. While this term is formally similar between the two methods, the methods of estimating the coefficient of variation are different (see Supporting Information S1 for a sensitivity analysis). As expected, the differences are larger for measurements with fewer transects (four at GE/PY compared to six at Chauvan): the Bayesian estimator of OURSIN is influenced by the prior values, potentially underestimated for such difficult measuring conditions, whereas the Student's coverage factor of QRev-UA may overestimate the coefficient of variation for such a small number of transects. In OURSIN results, the bottom discharge uncertainty fairly reproduces the uncertainty differences across the GE cross-sections. OURSIN provides a realistic explanation for larger discharge uncertainty at GE cross-sections since bottom discharge accounts for about 18% of total discharge and the unstable flow field (Le Coz et al., 2016) may increase the variability of vertical velocity profiles.

At PY, the total uncertainty estimated by QRev-UA is small and dominated by the uncertainty due to systematic errors, which is fixed to a similar value in OURSIN and QRev-UA. Again, the supplementary variance comes from the coefficient of variation in QRev-UA, and from the coefficient of variation and various discharge extrapolation terms in OURSIN. The OURSIN variance decomposition is consistent with the relative weights of extrapolated discharges in the total discharge. Discharge extrapolation uncertainty terms are generally smaller in QRev-UA because the parameter configurations explored by OURSIN include more modeling options, and more error sources, such as the draft error or edge parameter errors for instance.

For Chauvan experiments, the better uncertainty results from OURSIN as compared to QRev-UA are due to larger bottom and especially top discharge extrapolation terms, in relation with large extrapolated discharges and

Table 4

Variance Decomposition of Uncertainty Results From OURSIN and QRev-UA for the Génissiat 2010 and the Chauvan 2016 Acoustic Doppler Current Profiler (ADCP) Repeated Measures Experiments (Data of Figure 5, See Text for Definition of Uncertainty Terms)

	OURSIN										QRev-UA						
	Variance decomposition (%)										Variance decomposition (%)						
Site name	Syst.	Meas.	Nb. ens.	Inv.	Top	Bottom	Left	Right	*CV*	Total (m^6/s^2)	Syst.	Inv.	Extrap.	Edges	*CV*	Total (m^6/s^2)	Empirical total variance (m^6/s^2)
GE_1	19.8	0.6	5.7	0.3	4.6	26.5	3.1	1.3	38.2	43.4	46.8	1.4	<0.1	1.2	50.6	24.0	89
GE_2	18.9	0.5	6.0	<0.1	3.1	20.7	1.5	1.7	47.6	101.9	41.2	0.1	0.2	0.6	58.1	61.3	236
GE_2b	21.6	0.5	6.3	0.1	3.2	28.3	2.1	0.9	37.0	154.2	48.1	0.1	<0.1	0.5	51.3	90.7	365
GE_3	13.3	0.4	0.5	0.2	6.0	26.5	10.1	5.0	38.0	17.4	37.2	2.6	<0.1	2.8	57.3	8.2	49
GE_4	19.0	0.5	1.3	0.4	5.5	28.6	4.7	3.2	36.8	47.0	39.3	14.0	0.2	1.1	45.4	29.9	126
GE_4b	20.9	0.5	1.3	0.3	3.7	30.9	1.9	1.8	38.6	92.5	43.1	8.2	<0.1	0.7	48.0	58.9	247
PY_1	33.1	0.5	2.4	0.1	23.9	11.7	1.0	0.4	26.9	25.2	64.9	0.4	<0.1	0.5	34.2	16.9	37
PY_2	36.8	0.5	2.6	0.2	21.6	7.4	1.0	0.2	29.6	51.4	64.2	1.4	0.1	0.3	34.0	38.7	83
PY_2b	39.2	0.5	2.7	0.5	18.9	3.3	3.6	0.2	31.0	82.3	63.7	2.3	<0.1	0.4	33.6	66.7	112
PY_3	28.5	0.5	1.6	0.1	28.8	17.3	1.9	0.3	21.0	8.4	64.8	3.4	<0.1	0.6	31.2	4.9	9
PY_4	32.6	0.6	1.8	0.3	24.1	13.4	2.2	0.3	24.7	27.4	62.3	4.5	0.1	0.3	32.8	18.9	45
PY_4b	36.3	0.8	2.4	0.5	21.4	9.7	3.7	0.4	24.9	52.2	61.3	10.3	<0.1	0.4	28.0	40.4	63
A	26.1	0.5	2.8	0.2	43.1	10.3	0.1	0.1	16.9	0.14	65.1	2.6	<0.1	<0.1	32.2	0.07	0.12
B	27.6	1.2	2.8	0.2	47.7	7.1	0.1	<0.1	13.3	0.13	71.4	0.8	<0.1	0.1	27.8	0.07	0.14
C	26.3	0.8	2.4	0.2	39.7	9.9	9.6	<0.1	11.1	0.14	71.0	2.3	<0.1	1.9	24.8	0.07	0.20
D	14.9	0.4	1.6	0.1	19.3	5.6	50.8	0.5	6.8	0.24	57.5	0.5	<0.1	20.3	21.7	0.08	0.32
E	12.9	0.3	1.5	6.6	57.2	6.3	1.1	0.1	14.0	0.27	34.9	39.9	<0.1	0.2	25.0	0.13	0.31
F	15.7	0.1	1.6	1.4	62.2	8.0	<0.1	0.1	10.8	0.22	48.8	26.9	<0.1	0.1	24.2	0.09	0.35
G	20.1	0.4	2.1	0.3	55.4	9.9	0.1	0.1	11.7	0.18	64.7	4.6	<0.1	0.1	30.5	0.07	0.19
H	22.8	0.7	2.1	0.1	52.5	9.7	0.2	<0.1	11.9	0.16	70.1	1.4	<0.1	0.1	28.4	0.07	0.17
I	26.3	0.9	3.0	0.3	44.4	7.9	1.4	0.1	15.7	0.14	67.1	1.0	<0.1	0.5	31.5	0.07	0.09
J	17.7	0.3	1.9	0.4	62.2	8.6	0.2	0.1	8.6	0.20	71.0	1.5	<0.1	0.3	27.2	0.07	0.25
K	17.8	0.5	1.6	0.2	55.3	13.7	2.2	0.2	8.4	0.20	70.6	0.9	<0.1	0.7	27.8	0.06	0.24
L	22.8	0.7	2.4	0.1	49.5	12.8	1.2	<0.1	10.5	0.15	73.1	0.4	<0.1	0.2	26.3	0.06	0.20
M	29.1	1.0	3.2	0.1	44.5	7.9	0.1	0.1	14.0	0.12	71.6	0.6	<0.1	<0.1	27.7	0.07	0.09
N	22.4	0.7	2.6	0.2	51.7	10.7	2.0	0.1	9.8	0.16	70.6	1.8	<0.1	0.9	26.8	0.07	0.15
O	22.8	0.7	2.3	0.7	39.1	20.1	1.5	0.1	12.7	0.15	66.7	3.5	<0.1	0.4	29.5	0.07	0.17
P	26.7	0.7	2.3	0.5	38.1	18.3	0.1	<0.1	13.5	0.13	69.8	1.0	<0.1	<0.1	29.1	0.07	0.12
Q	28.1	0.9	3.5	0.2	36.7	16.7	<0.1	0.1	13.8	0.13	71.4	1.1	<0.1	0.1	27.3	0.07	0.12
R	28.0	0.8	3.3	0.2	44.0	9.8	0.2	0.4	13.4	0.13	71.0	0.7	<0.1	0.2	28.1	0.07	0.12
S	26.9	0.8	2.3	0.4	42.9	12.5	0.3	1.6	12.3	0.13	72.3	0.6	<0.1	0.3	26.9	0.06	0.17
T	23.8	0.5	2.1	0.2	46.7	16.9	0.2	0.1	9.4	0.15	74.9	0.6	<0.1	0.1	24.4	0.06	0.15
U	31.0	1.0	3.6	0.1	41.9	6.0	0.1	0.3	16.1	0.12	69.9	0.3	<0.1	0.1	29.7	0.07	0.10
V	26.5	0.6	3.1	0.2	42.9	15.9	0.8	0.2	9.8	0.14	74.3	1.7	<0.1	0.8	23.2	0.07	0.09
W	30.2	0.7	3.1	0.2	39.8	12.0	0.7	1.0	12.1	0.12	73.5	0.4	<0.1	0.6	25.5	0.06	0.09
X	26.9	0.6	2.5	0.5	29.8	25.9	2.1	0.1	11.7	0.14	71.2	2.3	<0.1	0.4	26.1	0.07	0.10

Note. Abbreviations mean: systematic errors (Syst.), measured discharge (Meas.), limited number of ensembles (Nb. Ens.), invalid cells and ensembles (Inv. cells, Inv. ens., Inv. when combined), top/bottom discharge extrapolation (Top, Bottom, Extrap. when combined), left/right edge discharge extrapolation (Left, Right, Edges when combined), and transect-to-transect variability (*CV* for coefficient of variation). Moving bed uncertainty does not appear as it is taken as zero for these experiments. Variance terms expressed in % are divided by the total variance. Total variances from OURSIN, QRev-UA, and the experiments (empirical variance) are displayed.

relatively uncertain vertical velocity profiles and exponents. At cross-section D, a substantial amount of discharge (6%) had to be extrapolated based on uncertain parameters, since almost none of the participants measured the edge distances and shapes precisely. The default settings result in a large edge discharge uncertainty, which is also visible in QRev-UA results but likely underestimated. At cross-sections E and F with shallower flows disturbed by a weir and a fallen tree, QRev-UA detects significant uncertainty due to invalid data, unlike OURSIN due to different uncertainty estimation procedures. In QRev-UA, the standard uncertainty is simply taken as 15% of invalid discharge while OURSIN compares the discharges computed with various interpolation options. As the invalid cells are distributed throughout the cross-section, differences in total discharge are small, hence a small uncertainty component.

4. Discussion

4.1. Empirical Validation of the Uncertainty Computation

As proposed by Le Coz et al. (2016) and applied by Despax, Favre, et al. (2016) to current-meter discharge measurements, repeated-measures (inter-laboratory) experiments are useful to understand and improve uncertainty propagation methods in hydrometry. Empirical uncertainty estimates from two large-scale moving-boat ADCP experiments compare relatively well with the uncertainty results from OURSIN: the mean and the 95% quantiles of the uncertainty differences from expanded empirical uncertainties ranging from 4% to 16% are −0.4% [−4.4%; +2.5%]. The OURSIN uncertainty estimates have only a small average bias across the measurements analyzed here. Large-scale experiments are useful for such validation purpose as the large numbers of participants and repeated measurements reduce the uncertainty of the uncertainty estimates.

Obviously, the empirical uncertainty estimates are specific to the study site and to the instruments and operators involved. The Génissiat 2010 and Chauvan 2016 experiments cover a range of cross-sectional characteristics (geometry, depth, width, and flow velocities), from shallow (0.6 m) to deep (9 m) and from relatively narrow (30 m) to wide (80 m) cross-sections. Different brands and models of ADCP from Teledyne RDI and SonTek manufacturers were involved and deployed either from powered vessels (Génissiat 2010) or tethered from the bank (Chauvan 2016). Operators were nearly all skilled professionals but field procedures, and software parameters and options could slightly differ across the operators. During the same experiment (Génissiat 2010) and at the same studied site, repeated measures have been conducted in steady flow conditions for various flow rates (from 120 to 440 m^3/s).

The uncertainty of a discharge measurement depends on the measurement conditions. Accordingly, the dominant contribution to the OURSIN uncertainty depends on the studied site. The dominant contributions are the coefficient of variation and the bottom discharge extrapolation at GE (Génissiat 2010), the systematic errors at PY (Génissiat 2010), and the top discharge extrapolation in the Chauvan 2016 experiments, except for cross-section D for which the edge discharge extrapolation dominates the uncertainty.

All the error sources related to ADCP discharge measurements found in Muste et al. (2004), González-Castro and Muste (2007), Kim and Yu (2010) or Despax et al. (2019) were not active simultaneously in the two experiments (cf. Table 5). Would OURSIN work as well in other measurement conditions? Additional comparisons with empirical uncertainty estimates obtained in other conditions are necessary for further validation. If different sources of errors that were small if not negligible in the Génissiat 2010 and Chauvan 2016 experiments are covered, their modeling in OURSIN could be evaluated. Challenging conditions for ADCP measurements include aquatic vegetation, low velocities, or high turbulence conditions. For instance, slow velocity may increase the impact of the measured discharge uncertainty; a limited number of ensembles would increase the contribution of the u_{ens} component, which could be quantified by alternative equations such as those developed by Despax, Perret, et al. (2016). Also, other experiments in different conditions will be useful to confirm that the possible double-counting of errors included in both the transect-to-transect variability and other discharge uncertainty components is not problematic.

4.2. Default Settings and Assumptions

The default values or options used in OURSIN to compute the uncertainty (Section 2) must be questioned and tested. They can be changed by the user if additional information or expert knowledge is available. Currently, the

Table 5
List of Error Sources That Are Active or Not During the Inter-Laboratory Experiments and Covered or Not by the Propagation Methods (QRev-UA and OURSIN Methods)

		Inter-laboratory experiments			Propagation method		
Uncertainty source	Variable	GE site	PY site	Chauvan	QRev-UA	OURSIN	Comments
Site characteristics							
Cross-section effect (geometry)	$\overline{Q}$	Active	Active	Active	Partially	Partially	
Moving-bed	$\overline{Q}$	Not active	Not active	Not active	Frozen	Frozen	
Temporal flow variation	$\overline{Q}$	Not active	Not active	Not active	u_{CV}	u_{CV}	
Flow field variation	$\overline{Q}$	Active	Not active	Not active	u_{CV}	u_{CV}	
Instrument							
Settings (water mode, cell size, etc.)	$\overline{Q}$	Active	Active	Active	Not covered	Not covered	[a]
Temperature, salinity	$\overline{Q}$	Active	Active	Active	Not covered	Not covered	[a]
Near transducer disturbance	$\overline{Q}$, Q_{top}	Active	Active	Not active	Partially	Partially	
ADCP model, frequency, float	$\overline{Q}$	Active	Active	Active	Not covered	u_{meas}	Active with respect to ADCP range limitations
Methods (algorithms)							
Top/bottom extrapolation	Q_{top}, Q_{bot}	Active	Active	Active	Active	Active	Chauvan: PP model used at each cross-section
Edge extrapolation	Q_{left}, Q_{right}	Active	Active	Active	Active	Active	
Interpolation (invalid data)	$Q_{invcell}$, Q_{invens}	Active	Active	Active	Active	Active	
Navigational reference (GPS or BT)	$\overline{Q}$	Not active	Not active	Not active	Not covered	Not covered	
Depth reference (vertical beam, 4-beam average)	$\overline{Q}$	Not active	Not active	Not active	Not covered	Not covered	
Operator							
Sensor draft	Q_{top}	Active	Active	Active	Not covered	$u(Q_{top})$	[a]
Edge distance	Q_{left}, Q_{right}	Active	Active	Active	$u(Q_{edge})$	$u(Q_{left})$, $u(Q_{right})$	
Edge sampling	Q_{left}, Q_{right}	Active	Active	Not active	$u(Q_{edge})$	$u(Q_{left})$, $u(Q_{right})$	Chauvan: 10 pings in stationary position requested
Boat operation	$\overline{Q}$	Active	Active	Active	Not covered	u_{meas}	
Sampling time	$\overline{Q}$	Active	Active	Active	u_{CV}	u_{CV}, u_{ens}	
Operator skills	$\overline{Q}$	Active	Active	Active	Not covered	Not covered	

[a]Propagation methods assume that the general guidance for making Acoustic Doppler Current Profiler (ADCP) discharge measurements are followed and all possible biases should be corrected prior the uncertainty analysis (UA).

default options in OURSIN are usually conservative, in order to cover the worst situations with minimal added information from the operator. For instance, the default settings for estimating edge discharge uncertainty correspond to the minimal (but common) edge estimation procedure with edge distances eye-balled and edge shape unchecked. This setting makes sense for the Chauvan 2016 experiment as the vast majority of the participants did not measure the edge distances and shapes precisely. As a result, the edge discharges were either negligible in the total discharge or significantly contributed to the total variance (at cross-section D). The resulting edge discharge uncertainty may look overestimated for measurements with more precise edge distance measurements and more precise knowledge of the edge shapes. Then, other settings should be proposed to reflect different quality levels of field procedures and information from the operators.

Some user-defined parameters have a small impact on the expanded uncertainty since they affect the input standard uncertainties in the measured area, which accounts for a limited contribution to the combined uncertainty. Those input standard uncertainties are averaged out by the number of cells, the number of ensembles, and the

number of transects. On the other hand, the combined uncertainty is sensitive to other parameters, for example: the draft error which affects the uncertainty of edge and top discharges, the models and options selected by the operator to extrapolate discharges in unmeasured areas of the cross-section as also reported by Moore et al. (2016), the uncertainty due to the ADCP method bias (assumed to be 1.31% by default), and the Bayesian priors used to estimate the coefficient of variation *CV*.

The bottom-track bias uncertainty $u'_{\text{syst}}(v_{BT})$ and the water-track bias uncertainty $u'_{\text{syst}}(w_{WT})$ could be re-evaluated using additional results from tow-tank calibration tests. Indeed, many ADCP calibration tests have been conducted since those reported by Oberg and Mueller (2007). In particular, it would be worth exploring whether the bottom-track and water-track systematic errors increase when velocity decreases, that is, whether the bias uncertainty should be constant or variable with velocity.

The default prior uncertainty of *CV* (see Section 2.2.6) has been defined small enough to stick to a credible value when the number of transects is too small, and large enough to tend toward the empirical value of *CV* when more transects are available. The default prior median of *CV* (3%) reflects the typical repeatability observed in repeated-measures experiments conducted in average conditions. The default prior values proposed in this work can be modified by the user when better information is available or more realistic assumptions can be defended. This better information may come from the results of repeated-measures experiments conducted in more similar conditions. Another perspective for defining generic prior values of *CV* may be to model them as a function of the time of exposure and/or the turbulent scales of the flow, as suggested by Oberg and Mueller (2007) and García et al. (2012). However, further investigation is needed to establish predictive equations and validate them for a broad range of river conditions.

4.3. Uncovered Error Sources

A number of perspectives for developing and improving the uncertainty computation have been identified. Table 5 shows the moving-boat ADCP errors that are covered or not by the QRev-UA and the OURSIN propagation methods.

First, the uncertainty related to using satellite navigation instead of bottom-tracking as a boat velocity reference is not covered in this paper, as all the validation data were measured with bottom-track reference. Rennie and Rainville (2006) and Wagner and Mueller (2011) provide rough estimates of the associated uncertainty. An uncertainty component related to compass errors and the use of GPS has been prepared for implementation in OURSIN. The potential bias in the measurement due to dynamic compass errors when using satellite navigation as the boat-velocity reference is computed using the method proposed by Mueller (2018) (Equation 41). The default value for the compass standard uncertainty is 1°, since most compasses used in ADCP are ±2°. The effect of an error in magnetic variation is still not taken into account however.

The uncertainty due to moving-bed was not considered either in this paper (frozen for both OURSIN and QRev-UA methods) since no discernible moving bed was observed during the experiments. A few moving-bed tests conducted at some cross-sections confirmed the absence of moving-bed effects. As it stands, the OURSIN method proposes to use the same expert values as proposed in QRev-UA. Further investigations must be conducted to get a more reliable assessment of the associated uncertainty (Mueller & Wagner, 2007). Other factors including the impact of sound speed (salinity, temperature) are assumed to be corrected prior to UA analysis and are not computed in OURSIN.

On the other hand, some error sources might be double-counted in OURSIN. As already mentioned, including the transect-to-transect discharge variability through the u'_{CV} term is useful to reflect error sources that are not covered by other terms. Even if the error sources contributing to the measured discharge uncertainty term u'_{meas} may also contribute to u'_{CV}, the validation results so far available suggest that it is more conservative to keep both terms in the uncertainty computation. If future verification tests find that uncertainty is overestimated in some conditions due to error double-counting, dropping u'_{meas} systematically or conditionally might be considered.

4.4. Improving the Uncertainty Propagation Method?

In OURSIN equations, measurement errors are considered either perfectly correlated (systematic) or perfectly uncorrelated (random). For instance, in the expression of the measured discharge uncertainty (Equation 10),

water-track and depth cell size errors are assumed perfectly uncorrelated, and factor $1/m_i$ would be removed if they were assumed perfectly correlated. Also, correlation of errors across successive transects is accounted for through the separation of systematic versus random errors in the computation of the uncertainty of a multiple-transect averaged discharge (Equation 3). Clearly, the truth is in-between as some elemental measurements in an ADCP transect are partially correlated, for example, velocities measured in adjacent cells of an ensemble from the same acoustic pings, beam-averaged depths of adjacent ensembles across the same bed area, extrapolated discharges, and the measured discharges used for the extrapolation. Covariance factors are easy to write in the uncertainty propagation equations but quantifying intermediate degrees of correlations precisely seems out of reach without additional studies and experiments. Meanwhile we only can select between perfectly correlated or perfectly uncorrelated errors as done in the proposed OURSIN method.

OURSIN is based on simplifications of the DRE (JCGM, 2008a) and uncertainty components, compared to the full-fledged framework proposed by Kim and Yu (2010) for instance. The advantage of reducing the error model to the most important error sources is that the end-user can more easily handle the uncertainty inputs and understand the uncertainty outputs. It clarifies the assumptions and makes the method more applicable by hydrological services. Also, in OURSIN the uncertainty propagation equation derives from the first-order Taylor expansion approach of the GUM, and a few configurations are computed to estimate the extrapolated discharge uncertainty components, instead of Monte Carlo simulations (JCGM, 2008b) as used in QUant for instance (Moore et al., 2016). This results in simpler and much faster computations, but the errors due to such approximations should be evaluated by applying OURSIN and a Monte Carlo-based counterpart (QUant, for instance) to a set of representative ADCP measurements and comparing the results. It is not impossible that Monte Carlo simulation procedures fast enough for use with ADCP measurements in the field can be implemented in QRevInt in the future. Then, as the OURSIN configurations used to compute alternative unmeasured discharges are based on assumed distributions of parameters and inputs, a Monte Carlo-based version of the OURSIN method should be relatively easy to establish.

5. Conclusions

This study presents the adaptation of the OURSIN method for UA of moving-boat ADCP discharge measurements in rivers as implemented in QRevInt, the international fork of the ADCP measurement processing software, QRev. This method combines uncertainties due to systematic and random errors to estimate the uncertainty of single or multiple-transect ADCP discharge measurements. The discharge uncertainty in measured areas is estimated by propagating the uncertainty of elemental measurements (boat velocity, cell size, water velocity). The uncertainty of discharges extrapolated in unmeasured areas (top, bottom, right/left edges, invalid ensembles, and cells) is estimated through expert-designed configurations varying meaningful parameters and options, as an alternative to the Monte Carlo approach.

The uncertainties provided by the OURSIN method and the built-in QRev-UA have been compared with the empirical uncertainty from two large-scale repeated measures experiments. In the given conditions of Génissiat 2010 and Chauvan 2016 experiments, with expanded empirical uncertainties in the range of 4%–16%, the comparison tends to validate the propagation method OURSIN as a consistent method for modeling the uncertainty of ADCP discharge measurements. The QRev-UA method originally implemented in QRev and QRevInt performs well in most scenarios despite its fairly simple UA. However the results indicate that, in the given conditions of the experiments, the OURSIN method is more accurate than QRev-UA, with a smaller mean bias (−0.4% vs. −2.2%) and a smaller range of uncertainty differences from the empirical uncertainty estimates ([−4.4%; +2.5%] vs. [−6.9%; +0.3%] as 95% quantiles). Additionally, OURSIN provides a more detailed decomposition of error sources. Such an uncertainty budget is most useful to practitioners for improving their measurement procedures.

The main ADCP error sources are included in the OURSIN method. However, as some error sources are not covered or were negligible in the experiments used for its validation, further comparison with empirical uncertainty estimates obtained in different environments, especially in challenging conditions, would be valuable. Also, additional options and uncertainty components could be included in OURSIN through future development. The numerical and statistical approximations of the uncertainty propagation scheme could be evaluated by comparison with a Monte Carlo based method such as QUant.

The OURSIN method is designed as an original trade-off between statistical accuracy and practical efficiency. The results presented in this paper and other applications not shown suggest that it provides reliable uncertainty

estimates for various site and flow conditions. Together with the OURSIN method, the QRevInt software provides an operational decision-making tool that combines quality assurance/quality control and UA. Through the quantification of discharge uncertainty and its sources, the value of having and enforcing standard operating procedures, as well as the use of a software like QRevInt for reviewing the quality of ADCP data, can be explored quantitatively. In turn, a better understanding of how ADCP discharge uncertainty can be minimized will help promote and improve best operating practices.

Appendix A: Combined Uncertainty Computation

The establishment of the OURSIN uncertainty propagation model presented in this article relies on the computation of the combined uncertainty introduced in the Guide to the expression of Uncertainty in Measurement or GUM (JCGM, 2008a). Here is a reminder of the general equation and its simple expression for two specific situations encountered in the derivation of OURSIN.

The measurement process is modeled through the data reduction equation (DRE):

$$Y = f(X_1, X_2, \ldots X_n) \tag{A1}$$

where random variables Y and X_1, X_2, …, X_n are the output (the measurand) and the inputs, respectively, of the measurement model f.

In the OURSIN model derivation, the input quantities of the DRE are always assumed to have independent errors. In such a case, the combined uncertainty $u_c(y)$, where y is the output estimate (i.e., the measurement result), can be computed through a first-order Taylor expansion of the DRE:

$$u_c^2(y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) \tag{A2}$$

where x_1, x_2, …, x_n are the input estimates (i.e., the elementary measurements).

When the DRE is a sum $\left(Y = \sum_{i=1}^{n} X_i\right)$, then the sensitivity coefficients $\partial f/\partial x_i$ are equal to 1, and the squared uncertainty of the output is equal to the quadratic sum of the uncertainties of the inputs:

$$u_c^2(y) = \sum_{i=1}^{n} u^2(x_i) \tag{A3}$$

When the DRE is a product $\left(Y = \prod_{i=1}^{n} X_i\right)$, then the sensitivity coefficients are:

$$\frac{\partial f}{\partial x_i} = \frac{\prod_{j=1}^{n} x_j}{x_i} = \frac{y}{x_i} \tag{A4}$$

As a result, the squared relative (i.e., percentage) uncertainty of the output is equal to the quadratic sum of the relative uncertainties of the inputs:

$$u_c'^2(y) = \frac{u_c^2(y)}{y^2} = \sum_{i=1}^{n} \frac{u^2(x_i)}{x_i^2} = \sum_{i=1}^{n} u'^2(x_i) \tag{A5}$$

Appendix B: Uncertainty Classification and Error Terminology

B1. Type A Versus Type B Uncertainty Evaluation

Type A and type B evaluations distinguish the method used to evaluate the standard uncertainty, denoted u (JCGM, 2008a). Type A uncertainty is calculated by a statistical analysis of measured quantity values (series of repeated observations) obtained under defined measurement conditions. The estimated standard uncertainty u is simply taken as the standard deviation σ of N observations x_i of the measurand (quantity intended to be measured) X:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \left(x_i - \overline{x}\right)^2} \tag{B1}$$

where $\bar{x}$ is the average of the N observations of X.

Type B uncertainty denotes the evaluation of standard uncertainty determined by any other mean than a Type A evaluation of measurement uncertainty. It includes estimation based on expert (including subjective) knowledge, results from calibration reports, proficiency testing reports, theoretical models, or sensitivity analysis.

B2. Probability Distributions for Measurement Uncertainty

Probability distributions (pdf) are a function $f(x)$ that shows the relationship between the outcome of an event x and its frequency of occurrence. Several probability distributions can be used for uncertainty analysis (UA). The most common, and the ones used in the OURSIN method, are the normal distribution and the rectangular distribution.

The normal (or Gaussian) distribution is a function that represents the distribution of the random variables as a symmetrical bell-shaped graph where the peak is centered about the mean and is symmetrically distributed in accordance with the standard deviation σ (see Figure B1). Central Limit Theorem states that the distribution of the sum of a large number of random variables, such as errors, will tend toward a normal distribution if none of the sources of error is dominant over the others. As a consequence, the normal distribution is the most commonly used probability distribution for evaluating Type A uncertainty. When computing the standard deviation (Equation B1) to evaluate the standard uncertainty u, it is implied that the quantified error follows a normal distribution.

The rectangular (or uniform) distribution is a function that represents a continuous uniform distribution and constant probability (see Figure B1). In a rectangular distribution, all outcomes are equally likely to occur between a lower and an upper bounds (a_- and a_+). In the lack of knowledge regarding the shape of the distribution, the rectangular distribution is often a default option in UA as an alternative to normal distribution. For instance, the uncertainty induced by a parameter selected by an operator is often modeled based on a rectangular distribution of error. There is an underlying assumption that all scenarios have the same probability to occur in the possible range of parameters.

To get a standard deviation equivalent, the standard uncertainty is evaluated as follows:

$$u(x) = \frac{a_+ - a_-}{2\sqrt{3}} \tag{B2}$$

In the practical application of the GUM (JCGM, 2008a), the bias correction corresponding to the expected value $(a_+ + a_-)/2$ of the uniform distribution is sometimes neglected and the user-defined parameters are applied to compute the measurement result. However, this may be problematic when the uniform distribution is substantially not centered on the measurement result.

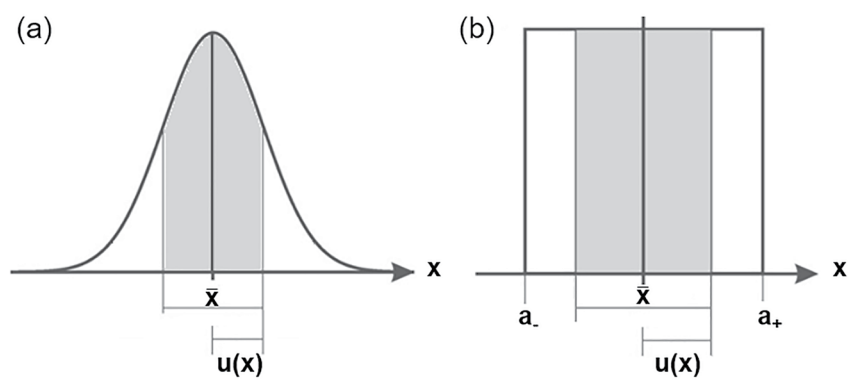

Figure B1. Distribution functions: (a) normal distribution and (b) rectangular distribution with mean $\bar{x}$ and standard uncertainty $u(x)$. The shaded areas of $\bar{x} \pm u(x)$ are the 68% and 58% confidence intervals around the means of the normal and rectangular distributions, respectively. Adapted from Meyer (2007).

B3. Random Error Versus Systematic Error

The uncertainty of a measurement may come from systematic or random errors. Random errors cannot be eliminated from an experiment, but most systematic errors can be reduced. Systematic errors induce a possible bias in the measurement. They can be detected by applying a QA/QC process. If the magnitude of systematic errors is known accurately enough, it must be corrected using adjustment factors or coefficients, before conducting the UA. However, unknown systematic errors can be treated as random variables, mostly because the correction factor is unknown and/or uncertain (JCGM, 2008a). On the other hand, random errors are unpredictable errors mostly due to environment sources of fluctuation. Random errors often follow a normal distribution. Note that random errors are reduced by data averaging, but systematic errors are not.

Appendix C: Error Velocity

Most of the ADCPs have four beams, disposed in a Janus configuration. Four-beam systems have a redundant beam that can be used to compute a quality check either on water-track velocity or on boat velocity (bottom-track referenced). For instance, the difference velocity dv can be calculated by averaging vertical velocities from opposing transducers pairs (pairs 1–2 and 3–4) as follows:

$$dv = \frac{(B_1 + B_2) - (B_3 + B_4)}{2\cos\theta} \tag{C1}$$

with B_1, B_2, B_3, and B_4 the radial velocities measured in beam 1, 2, 3, and 4, respectively. θ is the tilt angle of the beams referenced to vertical.

The so-called difference velocity is a good proxy to evaluate the validity of the homogeneous velocity field assumption (Gilcoto et al., 2009). If the flow field is homogeneous, the difference between these vertical velocities will average to zero. Teledyne RDI and QRev scale the difference velocity so that the error velocity is comparable to the horizontal velocity. The equation used to compute the scaled error velocity is:

$$e_v = \frac{(B_1 + B_2) - (B_3 + B_4)}{2\sqrt{2}\sin\theta} \tag{C2}$$

The error velocity can be computed based either on water-track velocities associated to each cell or on boat velocity (bottom-track referenced) associated to each ensemble following Equation C2.

QRev statistically filters the ensembles where the error velocity associated with the boat velocity is significantly different from the others. In the same way, water-track velocities are filtered when the error velocity associated to a cell is different from the others.

The dispersion of the remaining error velocities reflects the combination of errors due to ADCP noise and measuring conditions (bed slope, uneven bed configuration, turbulence, flow instability, waves). Under ideal conditions, the variance of the error velocity will indicate the part of the variance attributable to instrument noise (Teledyne RDI, 1998). On the other hand, errors induced by the measuring conditions may increase the error velocity.

Data Availability Statement

The raw ADCP data (binary files) from Génissiat 2010 and Chauvan 2016 experiments used in this paper are available in the following Zenodo repository: https://doi.org/10.5281/zenodo.7142646 (Despax et al., 2022). QRevInt Windows executables can be downloaded freely at https://www.genesishydrotech.com/qrevint.

Acknowledgments
We acknowledge the meaningful and valued comments received from three anonymous reviewers on previous versions of the manuscript. The authors thank the USGS for sharing the QRev code in Python and allowing OURSIN to be implemented in QRev first, then in QRevInt, the internationally developed fork of QRev. The development of QRevInt is supported by the following agencies: The Norwegian Water Resources and Energy Directorate, Environment and Climate Change Canada, New Zealand Hydrological Society, Swedish Meteorological and Hydrological Institute, Groupe Doppler Hydrométrie (France), U.S. Geological Survey, UK Centre for Ecology & Hydrology, Queensland Government Department of Resources, and the Australian Hydrographers Association. This study was funded by SCHAPI (SRNH 2019 and 2020 contracts), CNR, and EDF-DTG. Special thanks to Benjamin Renard (INRAE) for his insightful thoughts and his help for the Bayesian estimation of the coefficient of variation.

References

Blanquart, B. (2013). Panorama des méthodes d'estimation des incertitudes de mesure. *La Houille Blanche*, *6*, 9–15. https://doi.org/10.1051/lhb/2013045

Boldt, J. A., & Oberg, K. A. (2015). Validation of streamflow measurements made with M9 and RiverRay acoustic Doppler current profilers. *Journal of Hydraulic Engineering*, *142*(2), 04015054. https://doi.org/10.1061/(asce)hy.1943-7900.0001087

Chen, C.-L. (1989). Power law of flow resistance in open channels, Manning's formula revisited. In *International conference on channel flow and catchment runoff, Charlottesville* (pp. 817–848).

Despax, A., Favre, A.-C., Belleville, A., Hauet, A., Le Coz, J., Dramais, G., & Blanquart, B. (2016). Field inter-laboratory experiments versus propagation methods for quantifying uncertainty in discharge measurements using the velocity-area method. In *Proceedings of river flow 2016, Saint-Louis, USA* (p. 7).

Despax, A., Hauet, A., Le Coz, J., Dramais, G., Blanquart, B., Besson, D., & Belleville, A. (2017). Inter-laboratory comparison of discharge measurements with acoustic Doppler current profilers Chauvan field experiments, 8, 9 and 10th November 2016. (Technical report). Groupe Doppler (p. 92).

Despax, A., Le Coz, J., Hauet, A., Mueller, D. S., Engel, F. L., Blanquart, B., et al. (2019). Decomposition of uncertainty sources in acoustic Doppler current profiler streamflow measurements using repeated measures experiments. *Water Resources Research*, *55*(9), 7520–7540. https://doi.org/10.1029/2019WR025296

Despax, A., Le Coz, J., Mueller, D. S., Hauet, A., Calmel, B., Pierrefeu, G., et al. (2022). Validation of an uncertainty propagation method for moving-boat ADCP discharge measurements [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7142646

Despax, A., Perret, C., Garçon, R., Hauet, A., Belleville, A., Le Coz, J., & Favre, A.-C. (2016). Considering sampling strategy and cross-section complexity for estimating the uncertainty of discharge measurements using the velocity-area method. *Journal of Hydrology*, *533*, 128–140. https://doi.org/10.1016/j.jhydrol.2015.11.048

Dramais, G. (2011). *Quantification des incertitudes d'un jaugeage par profileur acoustique à effet Doppler. (Mémoire de fin d'études. CNAM)* (p. 96). CNAM.

García, C., Tarrab, L., Oberg, K., Szupiany, R., & Cantero, M. (2012). Variance of discharge estimates sampled using acoustic Doppler current profilers from moving platforms. *Journal of Hydraulic Engineering*, *138*(8), 684–694. https://doi.org/10.1061/(asce)hy.1943-7900.0000558

Gilcoto, M., Jones, E., & Fariña-Busto, L. (2009). Robust estimations of current velocities with four-beam broadband ADCPs. *Journal of Atmospheric and Oceanic Technology*, *26*(12), 2642–2654. https://doi.org/10.1175/2009jtecho674.1

González-Castro, J., Buzard, J., & Mohamed, A. (2016). RiverFlowUA—A package to estimate total uncertainty in ADCP discharge measurements by FOTSE—With an application in hydrometry. In *River flow 2016* (pp. 715–723). CRC Press.

González-Castro, J. A., & Muste, M. (2007). Framework for estimating uncertainty of ADCP measurements from a moving boat by standardized uncertainty analysis. *Journal of Hydraulic Engineering*, *133*, (12), 1390–1410. https://doi.org/10.1061/(asce)0733-9429(2007)133:12(1390)

Gordon, R. L. (1989). Acoustic measurement of river discharge. *Journal of Hydraulic Engineering*, *115*(7), 925–936. https://doi.org/10.1061/(asce)0733-9429(1989)115:7(925)

Hamilton, A., & Moore, R. (2012). Quantifying uncertainty in streamflow records. *Canadian Water Resources Journa*l, *37*(1), 3–21. https://doi.org/10.4296/cwrj3701865

Hauet, A., Le Coz, J., Sevrez, D., Dramais, G., Hénault, F., Perret, C., et al. (2012). *ADCP intercomparison in the Gentille canal (2011/09/12–16). (Tech. Rep.)* (p. 62). Groupe Doppler Hydrométrie. (in French).

Hauet, A., Morlot, T., & Daubagnan, L. (2018). Velocity profile and depth-averaged to surface velocity in natural streams: A review over a large sample of rivers. In *E3s web of conferences* (Vol. 40, p. 06015).

Huang, H. (2011). Uncertainty model for in situ quality control of stationary ADCP open-channel discharge measurement. *Journal of Hydraulic Engineering*, *138*(1), 4–12. https://doi.org/10.1061/(asce)hy.1943-7900.0000492

Huang, H. (2018). Estimating uncertainty of streamflow measurements with moving-boat acoustic Doppler current profilers. *Hydrological Sciences Journal*, *63*(3), 353–368. https://doi.org/10.1080/02626667.2018.1433833

Huang, H. (2019). The importance of ADCP alignment with GPS in moving-boat streamflow measurements. *Flow Measurement and Instrumentation*, *67*, 33–40. https://doi.org/10.1016/j.flowmeasinst.2019.04.002

ISO. (1994b). ISO 5725:2 – Accuracy (trueness and precision) of measurement methods and results. Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method (p. 42).

ISO. (1994a). ISO 5725-2:1994 – Exactitude (justesse et fidélité) des résultats et méthodes de mesure – Partie 2: Méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée (p. 55).

ISO. (2007). ISO/TS 25377:2007 – Hydrometry – Hydrometric Uncertainty Guidance (HUG) (p. 59).

ISO. (2009). ISO 748:2009 – Hydrometry – measurement of liquid flow in open channels using current-meters or floats (p. 58).

ISO. (2010). *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation* (p. 38). International Organization for Standardization.

ISO. (2017). ISO/TS 21748: Guidance for the use of repeatability reproducibility and trueness estimates in measurement uncertainty estimation (p. 58).

JCGM. (2008a). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (p. 132) (Guide No. 100). BIPM.

JCGM. (2008b). *Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method* (p. 90) (Guide No. 1). BIPM.

Kim, D., & Yu, K. (2010). Uncertainty estimation of the ADCP velocity measurements from the moving vessel method,(I) development of the framework. *KSCE Journal of Civil Engineering*, *14*(5), 797–801. https://doi.org/10.1007/s12205-010-0950-6

Le Coz, J. (2017). *Quantifying discharges and fluxes of matters in rivers (Habilitation à diriger des recherches HDR)* (p. 92). Université Grenoble Alpes.

Le Coz, J., Blanquart, B., Pobanz, K., Dramais, G., Pierrefeu, G., Hauet, A., & Despax, A. (2016). Estimating the uncertainty of streamgauging techniques using in situ collaborative interlaboratory experiments. *Journal of Hydraulic Engineering*, *7*(142), 04016011. https://doi.org/10.1061/(asce)hy.1943-7900.0001109

Le Coz, J., Camenen, B., Peyrard, X., & Dramais, G. (2012). Uncertainty in open-channel discharges measured with the velocity–area method. *Flow Measurement and Instrumentation*, *26*, 18–29. https://doi.org/10.1016/j.flowmeasinst.2012.05.001

Le Coz, J., Pierrefeu, G., Saysset, G., Brochot, J., & Marchand, P. (2008). *Mesures hydrologiques par profileur Doppler* (p. 164). Quae.

Le Coz, J., Saysset, G., & Pierrefeu, G. (2009). *ADCP regatta – Vézère 2009. (Tech. Rep.)*. (p. 14). Groupe Doppler Hydrométrie (in French).

Lee, K., Ho, H.-C., Marian, M., & Wu, C.-H. (2014). Uncertainty in open channel discharge measurements acquired with StreamPro ADCP. *Journal of Hydrology*, *509*, 101–114. https://doi.org/10.1016/j.jhydrol.2013.11.031

McMillan, H., Seibert, J., Petersen-Overleir, A., Lang, M., White, P., Snelder, T., et al. (2017). How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications. *Water Resources Research*, *53*(7), 5220–5228. https://doi.org/10.1002/2016wr020328

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/10.1063/1.1699114

Meyer, V. R. (2007). Measurement uncertainty. *Journal of Chromatography A*, *1158*(1–2), 15–24. https://doi.org/10.1016/j.chroma.2007.02.082

Moore, S. A., Jamieson, E. C., Rainville, F., Rennie, C. D., & Mueller, D. S. (2016). Monte Carlo approach for uncertainty analysis of acoustic Doppler current profiler discharge measurement by moving boat. *Journal of Hydraulic Engineering*, *143*(3), 04016088.

Mueller, D. S. (2012). *Estimating random uncertainty for 2 transect discharge measurements*. (Tech. Rep.). U.S. Geological Survey (p. 8).
Mueller, D. S. (2016). *QRev—Software for computation and quality assurance of Acoustic Doppler Current Profiler moving-boat streamflow measurements—User's manual (Tech. Rep.)*. US Geological Survey.
Mueller, D. S. (2018). Assessment of acoustic Doppler current profiler heading errors on water velocity and discharge measurements. *Flow Measurement and Instrumentation*, *64*, 224–233. https://doi.org/10.1016/j.flowmeasinst.2018.10.004
Mueller, D. S. (2020). QRev. U.S. Geological Survey software release. https://doi.org/10.5066/P9OZ8QDL
Mueller, D. S. (2021). *QRevInt, version 1.16*. Genesis HydroTech LLC. Retrieved from https://bitbucket.org/genesishydrotech/qrevint/src/master/
Mueller, D. S., & Wagner, C. R. (2007). Correcting acoustic Doppler current profiler discharge measurements biased by sediment transport. *Journal of Hydraulic Engineering*, *133*(12), 1329–1336. https://doi.org/10.1061/(asce)0733-9429(2007)133:12(1329)
Mueller, D. S., Wagner, C. R., Rehmel, M. S., Oberg, K. A., & Rainville, F. (2013). Measuring discharge with acoustic Doppler current profilers from a moving boat (version 2). Techniques and Methods, book 3, chap. A22. US Geological Survey. (p. 95). https://doi.org/10.3133/tm3A22
Muste, M., Yu, K., Gonzalez-Castro, J., Ansar, M., & Startzman, R. (2004). Methodology for estimating ADCP measurement uncertainty in open-channel flows. In *Critical transitions in water and environmental resources management* (pp. 1–13).
Naudet, G., Pierrefeu, G., Berthet, T., Triol, T., Delamarre, K., & Blanquart, B. (2019). Oursin: Outil de répartition des incertitudes de mesure de débit par aDcp mobile. *La Houille Blanche*, *105*(3–4), 93–101. https://doi.org/10.1051/lhb/2019056
Oberg, K., Morlock, S. E., & Caldwell, W. S. (2005). *Quality-assurance plan for discharge measurements using acoustic Doppler current profilers (Tech. Rep.)*. (Report 2005-5183). US Geological Survey .
Oberg, K., & Mueller, D. (2007). Validation of streamflow measurements made with acoustic Doppler current profilers. *Journal of Hydraulic Engineering*, *133*(12), 1421–1432. https://doi.org/10.1061/(asce)0733-9429(2007)133:12(1421)
Pagano, T. C., Wood, A. W., Ramos, M.-H., Cloke, H. L., Pappenberger, F., Clark, M. P., et al. (2014). Challenges of operational river forecasting. *Journal of Hydrometeorology*, *15*(4), 1692–1707. https://doi.org/10.1175/jhm-d-13-0188.1
Pobanz, K., Le Coz, J., Hauet, T. A. F., Longefay, Y., & Pierrefeu, G. (2015). *ADCP/SVR intercomparison in the Rhône river downstream of Génissiat dam, 2012/09/25-28. (Tech. Rep.)*. Groupe Doppler Hydrométrie (in French) (40 p.).
Pobanz, K., Le Coz, J., & Pierrefeu, G. (2011). *ADCP intercomparison in the Rhône river downstream of Génissiat dam (2010/10/12–15). (Tech. Rep.)*. (p. 59). Groupe Doppler Hydrométrie (in French).
Rennie, C. D., & Rainville, F. (2006). Case study of precision of GPS differential correction strategies: Influence on ADCP velocity and discharge estimates. *Journal of Hydraulic Engineering*, *132*(3), 225–234. https://doi.org/10.1061/(asce)0733-9429(2006)132:3(225)
Simpson, M., & Oltmann, R. (1991). *Discharge-measurement system using an acoustic Doppler current profiler with applications to large rivers and estuaries*. US Department of the Interior, US Geological Survey.
Teledyne RDI. (1998). *ADCP coordinate transformation: Formulas and calculations (Tech. Rep.)*. RDI.
Teledyne RDI. (2007). *Workhorse ADCP technical manual (Tech. Rep.)*. RDI.
Wagner, C. R., & Mueller, D. S. (2011). Comparison of bottom-track to global positioning system referenced discharges measured using an acoustic Doppler current profiler. *Journal of Hydrology*, *401*(3–4), 250–258. https://doi.org/10.1016/j.jhydrol.2011.02.025
WMO. (2010). *Manual on stream gauging – Vol. I* (p. 252). World Meteorological Organization.