Response to Reviewer 1

General Comments

Reviewer Point P 1.1 — The manuscript deals with image processing aimed at obtaining reliable flow discharge estimates under continuous unsupervised operation. The utility of this approach is acknowledged by the WMO, recently issuing an update to WMO-No. 168, section 5.3.7.3 – Image-based methods for discharge measurements. The authors add to the general discussion on the applicability, accuracy and associated uncertainties by presenting the results from nearly one-year continuous observations on the River Dart, UK. The data were collected at Austins Bridge, where available hydrometric records span back to 1958, and treated in compliance with the WMO-No. 168 requirements.

The River Dart at Austins Bridge yields the longest hydrometric record on the River Dart, but its hydrologic presentation in subsection 2.1 is apparently scarce. It seems however important to put the experimentation in the proper context, in terms of both average and extreme flows. The notion on 'reference observations' appears as early as in Subsection 2.5.1, where they are briefly mentioned in L172 as a source of the conversion factor, but the details are lacking. What is their distribution across the velocity/discharge range? What are the standard data collection protocols? Even the number of observations is uncertain. The total number of reference observations (303) is presented in the Abstract, but never discussed elsewhere. Were some of the, or all aDcp observations (at least some were, as stated in L182, but clearly not all, since the record starts is 1981)? Subsection 2.5.1 is vague on the subject, L177 has no notions on the Ua protocol, and in the rest of the subsection only aDcp data are discussed.

The manuscript is generally well-written, and most issues that can be reported are related to an overall quality of presentation and text flow. While the manuscript grosso modo has most of the data needed, some of these data are scattered across the manuscript making it hard to follow.

Reply: Thankyou for the constructive comments. We will seek to implement each of these suggestions in the revised manuscript, and below outline how this will be achieved.

With reference to the point concerning the lack of context, information about the location where the experiment takes place, and information about the reference measurements used, we have sought to address these concerns. Specifically, in the revised document, *Section 2.1 Experimental Site* will deal with a more complete description of the geographical and hydrological background. Furthermore, an additional section is provided, *Section 2.2 Reference Data*, which provides detailed summaries of the reference gauging data and provides information on how this is used to establish Ua. In the supplementary information we will also provide summary information documenting the instruments and methods used to acquire the reference gauging data along with digitised copies of the flow gauging measurements obtained between 1980 to 2018.

Specific Comments

Reviewer Point P 1.2 — L31: are these mentioned below algorithms, in the sense of Kanade-Lukas algorithm, or rather methods and approaches, as per WMO-No. 168?

Reply: These are perhaps better described as classifications of methods (per WMO-No. 168). This has therefore been revised to read: *Measurements of surface velocity can be achieved through the application of methods that may be broadly categorised as: particle tracking velocimetry (PTV) (Brevis et al., 2011; Tauro et al., 2017); large scale particle image velocimetry (LSPIV) (Fujita et al., 1998; Muste et al., 2008); or space-time image velocimetry (STIV) (Fujita et al., 2007).*

Reviewer Point P 1.3 — L41: reference the Lukas-Tomasi algorithm, and probably mention Kanade here also, since KLT in KLT-IV might presumably stand for 'Kanade-Lukas-Tomasi' (indeed, references from L132 can be moved up here).

Reply: This suggestions, in addition to Reviewer 2's comments on this sentence, have been implemented in the revised document to read: *These computer vision algorithms e.g. Kande-Lucas-Tomasi algorithm (Lucas and Kanade, 1981; Tomasi and Kanade, 1991; Shi and Tomasi, 1994), automatically identify pixels that are distinct from their neighbours, and these distinct features can be iteratively tracked through a sequence of images.*

Reviewer Point P 1.4 — L43: PIV is not presented above, but is apparently the same as "PIV" in LSPIV, for the latter is the same PIV in the narrower context (WMO-No.168).

Reply: This will be clarified in the revision. We replace mention of PIV with the more specific LSPIV term. The revised sentence reads: *These approaches are computationally very efficient and capable of performing analysis up to two orders of magnitude faster than traditional LSPIV and PTV approaches (Tauro et al.,2018b)*

Reviewer Point P 1.5 — L43-44: in referencing (Tauro et al., 2018b) it can be noted that at least some of the results of that paper were obtained using "woodchips ... continuously deployed" from the upstream so not fully in the context of a "continuous, automated, and unsupervised workflow".

Reply: This is correct. We were using reference to this particular paper to illustrate that the results generated by optical flow methods are comparable to standard technologies, rather than suggesting that Tauro et al (2018) tested the algorithms in a continuous and automated workflow.

Reviewer Point P 1.6 — L94-95: this phrase evokes the question on what were the hydrological conditions during the experiment, of which there is no notion in Section 2.1. How they are compared to long-term averages and extremes? It is only later in the manuscript (L178) that we will know that the reference flow measurements range from 1.7 to 145 m3s-1; no info is provided on the flow discharge ranges during the experiment.

Reply: We will address this through the addition of new material provided in Sections 1.1 and 1.2.

Reviewer Point P 1.7 — L170: "20 segments of equal size"; from Figure 2, these are segments of equal width, not equal area, though for polygons the latter is most commonly assumed. This needs to be stated explicitly. Am I right to understand this so that while the flow width increases with the increasing discharge, the width of each segment increases accordingly?

Reply: Very true. These segments are of equal width, not equal size and this will be modified in the revised document. Yes, as the flow width increases, the width of each cell will increase to maintain the condition of 20 segments of equal width.

Reviewer Point P 1.8 — Also, this first paragraph of the Subsection 2.5.1 is somewhat confusing. First, while the basic references are provided, I would be interested in having more details on how exactly the observed data were extrapolated over the regions out of the camera sight using these three techniques; why there are 20 segments, and not more/less – how is this justified? In L170, was it so that "average velocities for 20 segments of equal sizes are established" stands for the water surface velocities averaged across the segment width? What is the number of pixels with data in each segment surface? How exactly was alpha derived? The (Perks, 2024b) reference is misguiding in that might presumably lead to the supportive information while it simply references a figure derived from unexposed data. Were the alpha values derived segment-wise for all 60 aDcp measurements? Were the segments similar or close in size (width?) in aDcp and extrapolation exercises? Further on, is the "cell" in L173 equivalent to "segment" in the lines above? If not, the difference must be explained otherwise the text is ambiguous.

Reply: In terms of the selection of 20 cells, we select this value so that our discharge calculation may meet the guidelines set out by ISO 748 which state that the discharge of any verticals should not be more than 10 per cent of the total discharge and ideally no more than 5 percent of total expected discharge, and that in fixing verticals, equal spacing should be preferred wherever possible.

The number of velocity measurements present within each cell may be as low as zero (requiring interpolation/extrapolation), single digit numbers, up to several hundred. With an increasing number of successfully tracked tracers within a particular cell, the greater the level of confidence in the reported average values. However, a cell with a single velocity reconstruction may also be valid.

Alpha is derived by producing a power-law fit between the normalised river velocity and normalised distance from the river bed using all aDcp pings for an individual transect. Using a unique function developed for each transect, we calculate the ratio between the (extrapolated) velocity at the water surface with the predicted velocity at 0.6D (which corresponds to the theoretical depth-averaged velocity). This ratio provides us with the alpha value for each individual transect. Finally, we take the median of all these alpha values and employ this in the scaling of surface velocities to depth-averaged velocities. In addition to the summary plot that we initially provided, we will also provide plots showing the alpha results for each individual transect and provide reference to this is made in the main text.

In the original submission we used the terms 'segment' and 'cell' interchangeably and acknowledge this as an error. In the revised version of the manuscript we will consistently use the term 'cell' to describe the 20 units that make up the cross-section area.

Reviewer Point P 1.9 — L180: In "calculate the (1-D) velocities derived from reference observations", can these be replaced simply by Ua, as given in L177? Similar to this, can "1-D" be omitted from most instances where its presence is misleading or redundant, e.g., L184, L187, L191-192, L198 and almost each and every instance below. Once the definition is given, there is no need to reiterate, and it makes sense to substitute. As in L180, same in L196-197, same in L198-199, and so on.

Reply: Thank-you for this comment. In our initial submission we intentionally reiterate the definition throughout based on feedback on an earlier draft. Previous feedback suggested that use of Ua and Us may make the text more difficult to follow. However, given your feedback we will remove the redundant messaging as outlined in your comment.

Reviewer Point P 1.10 — In this Subsection 2.5.2, again, the statistics for reference stage observation can be highly utile. Were these reference observations mentioned in L196 the same as presented elsewhere including the Abstract? Can the authors consider providing the details either here or in the Section 3?

Reply: In Sections 2.6.1 and 2.6.2 we now add additional information about the flow range of the reference measurements that are used in the calibration and validation phases of the research.

Reviewer Point P 1.11 — L210: It is probably more correct to not start the "Velocity reconstruction" section with the overview of the EA data. Also, in Figure 3, velocity reconstruction is applicable to both procedures employed, see Subsections 2.5.1 and 2.5.2 accordingly. Thus Section 3.1 must be renamed accordingly.

Reply: Whilst we acknowledge this point, we feel that description of the Ua data distribution provides useful context for the subsequent information we present on the relationships between Ua and Us. This information could perhaps be placed in the newly produced Section 2.2 (Reference Data), however, this level of detail is perhaps more than would be required at such an early stage of the article.

Reviewer Point P 1.12 — L220: more justification is needed here or above on claiming the constant Froude as a best-performing model, since the Figure 4 is indecisive on this matter.

Reply: Additional detail is provided in this section to illustrate that at flow velocities in excess of $0.59 m\ s^{-1}$ the Froude approach produces velocity estimates in closest agreement with the reference velocity.

Reviewer Point P 1.13 — Figure 5: is it correct that the distances are given from the left bank while the camera is on the right bank?

Reply: Yes this is labelled correctly.

Reviewer Point P 1.14 — L244-245: how exactly the conclusion that "as few as eight flow gauging measurements" suffice to successfully quantify the cross-section averaged velocity was reached? Does it imply that the uncertainty (variability) level is acceptable at this n.

Reply: In the modified version of the manuscript, this is rephrased to highlight that the levels of uncertainty do not appreciably improve when the number of measurements used in the calibration is increased beyond eight.

Reviewer Point P 1.15 — L266: "autonomously analysed in an unsupervised workflow". This said, it must be noted however that while the imagery was indeed collected in that way, its further treatment and the final output weren't possible without the prior work on-site, including aDcp surveys under challenging conditions of the below 1% flood. Without this, based solely on the a priori knowledge of the site, what the potential errors of the standalone application could have been?

Reply: This is a fair point. In order to adopt one of the theoretical flow distribution approaches in the autonomous workflow we needed to have cross-section information and an estimate of the α value to convert the surface velocities to a depth-averaged. To adopt the index-velocity approach in the autonomous workflow we would need to have a number reference flow gauging's to calibrate the surface velocities to (as Discussed in Section 3.2). We will add this information to this section.

Reviewer Point P 1.16 — L278-279: Top and bottom blind zones are present in aDcp measurements, along with right and left margins, so the reported Q is corrected for these in RiverSurveyor software. The wetted cross-section area has some effects on the aDcp-derived flows, but it is not to be overestimated.

Reply: The statement that this comment refers to will be qualified in the revised document.

Reviewer Point P 1.17 — Overall, I concur with the authors in that the major utility of the described approach is in the "development or refinement of stage-discharge rating curves", where the OTV and OTV-like approaches can be highly utile in increasing the accuracy of high flow estimations. For the River Dart at Austins Bridge, the highest observed peak flow is 550 m3s-1 (1979-1980), almost 400 m3s-1 over the observed highest peak flow. At this part of the flow duration curve, as far as I understand, the OTV can substantially improve the flow estimates in the straight single-thread non-deformable channel.

Reply: Thank-you for the constructive comments made within this review. We believe that the manuscript will be significantly improved by addressing these as outlined in the responses above.
