Response to reviewer 1

The authors thank the reviewer for their constructive comments, our responses to each individual point can be found below:

Changes to the manuscript in line with the responses are outlined in blue:

This technical note compares different methods to quantify hysteresis patterns and introduces a new, more robust way to do so. The manuscript is well-organized, clearly written and potentially of interest to quite some of the readers. From my point of view, it can be considered for publication after addressing a few minor comments:

(1) although being widely used in hydrology, the term "hysteresis" used here is formally incorrect. Hysteresis is defined as the dependence of a system output on its history of inputs (and thus on its internal state). Although discharge is a manifestation of the system state, the discharge-concentration relationships are technically no hysteresis loops but rather closed loops of a functional relationship. In addition, actual hysteresis is characterized by unique input-output relationships below and above given threshold values (e.g. Schmitt-triggers from electronic circuits as examples for sharp hysteresis). I would therefore suggest to qualify the terminology here, for example by stating:"[...] closed loops, thereafter referred to as hysteresis loops".

As the reviewer says, the term hysteresis is widely used in hydrology and is generally understood by the hydrological community. As a result the authors do not think it is necessary to add and additional explanation.

(2) p.7883, l.9ff: I could not quite follow this explanation. In other words, I am not sure if the new method is capable of a more robust representation of figure-of-eight shapes. Even if using the normalized ranges, wouldn't a regular 8-shape (for the sake of the argument say for example horizontally aligned at an angle of 0 degrees) result in a HI of 0 in spite of exhibiting "hysteresis"? It would be great if the authors elaborated a bit on that and clarified this question.

Yes, we agree that this point needs clarifying in the manuscript. The reviewer is correct in saying that a regular symmetrical figure-of-8 loop (i.e. equal size loops on either side) would result in an HI index of 0. But that is a simple fact of an unbiased loop. The new index however, does allow the method of quantification to consider the portions of the loop which are in clockwise or anti-clockwise phase and this information could be extracted for further evaluation. This is an improvement on the previous published hysteresis indices. If the new index is used in conjunction with other existing hysteresis measures such as loop area, it is easy for the user to see that a loop which has a HI of 0 but a loop area which is larger than 0 has to exhibit figure-of-8 behaviour. In addition to this, because the calculation of the new index uses multiple sections across the loop, which will encompass the clockwise and anti-clockwise sections, it is possible to examine the distribution of values gained for the index before they are averaged, thus allowing the user to see the value of the index in each section of the loop. Text will be added to the technical note to clarify that users who are examining figure-of-eight loops may find it helpful to use the new index in conjunction with other loop measures and/or visual examination of the loop shape to ensure an effective interpretation of the results. So to be clear, here we focus on the basic output that can be generated. Once implemented other summary results can be gained that can be used to highlight different aspects of the loop characteristics (one could summarise separately the +/- aspects of the loops for more complex behaviour). This is beyond our technical note, but we shall briefly note that other characteristics can be quantified as this is a strength of the new methodology.

Text has been added to the manuscript lines 193-203

(3) is there a particular reason not to show the box plots in figure two with equal y-axis scales (at least for panel ii and iv of each storm). this could more clearly illustrate that HInew is somewhat more robust.

Yes, we shall modify the plots so that the y-axes match in each of the ii and iv panels for ease of comparison.

Plots have been modified, however, due to the fact that the index can only produce a number between -1 and 0 (unlike the original index), so all the y-axes in panel iv have been made consistently between -1 and 1.

Response to reviewer 2

Thank you for these useful comments, our response to each of the points raised can be found below:

This technical note gives a review of some indices that are used to describe direction and magnitude of hysteretic relationships between discharge and concentration and proposes a new hysteresis index. Hysteretic relationships between concentration (geochemical tracers, nutrients) and discharge or also between storage (i.e. moisture contained within a control volume) and discharge have been used to describe catchment functioning and to compare catchments or different time periods. Observed hysteretic behavior could help to infer flow processes and better understand runoff generation. In that respect, this technical note, although more geared in its current scope towards nutrient and sediment export from (agricultural) catchments, could be interesting for many readers dealing with hillslope and catchment hydrological processes. This technical note is well-written and mostly clear in its explanations and structure

I understand that a technical note has to be brief. Still, I would recommend to provide a short explanation in the introduction of what is meant by hysteresis in this context and to elaborate a bit on the value of a hysteresis index (HI). Why can it be a useful descriptor of catchment functioning? Has the examination of hysteresis patterns advanced process understanding? How can it help to pinpoint release mechanisms for nutrients or sediments beyond a mere comparison of numbers between catchments? What does it mean if a hysteretic loop is clockwise or anti-clockwise in terms of processes? This also refers to the conclusions section where authors state that the new HI could "become a standardized analytical technique to be used by the water quality research community".

The authors appreciate that some of this background information could be useful to the reader and can help support the value of using a HI, however they are also conscious of the need for brevity in a technical note. Therefore the authors propose that a sentence can be added to the technical note which refers the reader to an additional paper which is currently in press which uses the new hysteresis index as a tool for quantifying hysteresis loops across different parameters and field sites. This paper covers in detail all of the issues you highlight here in your comment and would allow the reader to see how the hysteresis index can be used. We ask for advice from the Editor on the basis this is a technical note paper and such discussion should be limited.

Text has been added to provide additional background and a definition of hysteresis in the introduction (lines 29-32), also details of what different hysteresis pattern mean in terms of processes (lines 44-50). An extra recommendation has been added to the conclusions section, along with an expanded explanation of why the index is of wider significance for the hydrological community (lines 241-244).

P. 7879, L 3: Please explain TNU

NTU is a standard unit of measurement of turbidity which stands for “Nephelometric Turbidity Units”. This could be added to the manuscript, however we would argue that this abbreviation is widely accepted and commonly used in the hydro-chemical literature. We are happy to clarify this though.

Has been added to the manuscript (line 96)

P. 7879, L 19-21: Please make the explanation of the calculation of the adapted HI clearer. What exactly does it mean to calculate HI “at every 25, 10% etc of the discharge” and to calculate for different “sections” (e.g. p. 7884, L 15-19) or use different “increments”. This remained somewhat unclear to me throughout the text.

The original index proposed by Lawler et al. 2006 used the mid-point in discharge to determine the measurement point for the index (50% of the discharge range). Our adapted and new method instead determines multiple locations across the loop at which to measure the strength of the hysteresis. Therefore we tested the impact of using different numbers of measuring points or increments, including every 25% of the discharge range i.e. 3 equally spaced measurement increments across the loop, 10% of the discharge range (9 increments) etc… If helpful a visual aid could be produced and added to the methodology section to clarify this difference but we believed this was clear in the current text and presentation.

A visual aid has been added in the form of figure 3.

P. 7881, L 25 – p. 7882, L 4: redundant, as it is explained in the figure caption

Agreed, this can be removed as it is repeated in the figure caption.

Has been removed.

P. 7883, L 7-14: this description belongs to methods (section 2.3), not results

The description of the new index is covered in the methodology section, however, in the results the details are reiterated in order to clearly explain what the reader is observing in the figure. Therefore the authors would like the sentence to remain.

P. 7884, L 6-7: meaning of sentence unclear

Agreed, this sentence will be amended to read: “This technique is useful when the user’s interest is in the relative characteristics of the loop geometries”.

Amended as above

p. 7884, L 7: “These” means these recommendations?

Yes, should be these recommendations, text can be modified to clarify.

Has been modified.

Response to reviewer 3

The technical note from Lloyd et al. ("Testing an improved index for analysing storm nutrient hysteresis") compares methods for calculating a hysteresis index for concentration - discharge relationships during storm events. The note is appropriate for HESS and will be of interest to researchers seeking metrics to interpret C-Q relationships. My only major recommendation is that

the authors remove "nutrient" from the title since the paper does not discuss nutrient data (but rather turbidity - discharge relationships).

Thank you for your comment. We agree with your recommendation; we therefore prose to amend the title to “Technical Note: Testing an improved index for analysing storm hysteresis dynamics”. We use turbidity as an example in the technical note, as hysteresis in turbidity is prevalent in the literature and we had a large number of storms displaying a wide range of hysteretic behaviours for which we could test our methodology (explained P7879 ln3-6), however the technique is more widely applicable to any quantifiable water quality parameter. We would like therefore to represent this in a broader title.

Following additional constructive advice from the editor, the title has been amended to “ Technical Note: Testing an improved index for analysing storm discharge- concentration hysteresis.”

Response to reviewer: Remi Dupas

Thank you for your comments, our response to the individual points you raise can be found below:

This technical note reviews some of the hysteresis-descriptor variables used to analyse high frequency storm concentration time series. Two major shortcomings of the widely used hysteresis index (Lawler et al., 2006) are highlighted: the influence of initial concentration and of initial discharge in the case of 8-shaped hysteresis. A new hysteresis index is presented to overcome these two shortcomings. It worth noting that this is one of the rare studies where uncertainty in the data is accounted for in classifying hysteresis loops. This technical note is well-written, logically organized, and the figures are clear. This technical note would benefit from two major improvements

(1) An alternative method already exist to deal with the problems of changing baseline value and 8-shaped hysteresis loops. See Rossi et al. (2005) and also Stutter et al. (2008) and Dupas et al. (2015) for examples of application. Here is an extract from Stutter et al. (2008): “Further analyses were undertaken using the ‘pollutogram’ approach developed by Rossi et al. (2005) approximated by the relationship: $F(x)=x^{\beta}$ where F(x) is the fraction of the total mass of the determinant during the storm event and x is the total mass of water during the event. The parameter β is a coefficient representing the relationship between the mass and water volume over time which may be plotted as the cumulative proportion of the total mass transported against the cumulative proportion of water transported. Values of β of 1 indicate that the determinant mass arrived predominantly towards the start, or end of the event, respectively. A value of β = 1 denotes either that the pollutant mass and water volumes are proportional, or that the pollutant concentrations stay constant over the event.” Maybe mention this method.

Thank you for this suggestion. There are a number of different methods which can be used to examine storm behaviours, some of which we have discussed in this technical note. The method you describe is another viable method for examining storm behaviour, however the pollutogram is designed to examine discharge-load relationship, which are subtly different to discharge-concentration relationships. This is important in our work as we consider variables such as turbidity from which a load cannot be directly calculated unless converted to suspended sediment. With this in mind, the authors would prefer not to add this method to our discussion as we only wish to include methods which directly examine discharge-concentration relationships as we have indeed identified in the introduction to the paper.

(2) Maybe mention the fact that the new HI gives a description the size and direction of the biggest loop in the case of a 8-shaped loop but the information that it is a 'figure-ofeight' is lost. See also comment (2) Anonymous Referee #1. The method mentioned in (1) leads to the same information loss.

Please see the response provided to Reviewer 1 (comment 2). In brief, the new index provides a useful method for quantification which reflects the proportion of the loop which is in clockwise and anti-clockwise phase in the case of figure-of-8 loops. If the index is coupled with a visual inspection of the loops, then no information is lost, or indeed if the information is extracted separately from each loop as noted here: 1) If the value obtained for the HI is small but other metrics such as loop area are large in comparison, then it can quickly be determined that the loop is a figure-of-eight. 2) In addition, the multiple sections of the loop which are measured as part of the index calculation can be examined before they are averaged, and therefore a switching between positive and negative values indicates the switching from clockwise to anti-clockwise behaviour, resulting in a figure-of-8. The amendments proposed in response to comments made by reviewer 1 should also help to clarify the point raised here and will we add text to note these additional behaviours that can be quantified as another positive aspect of the new approach.

See earlier response and modifications

Minor comments:

P 7876 l2: "in extreme flow events" -> why not all storm events?

Agreed, this could apply to any storm events, text will be modified.

Text in the abstract has been amended to storm events.

P 7877 l14: a major interest of hysteresis-descriptor variables is that they enable statistical analysis of near continuous high-frequency measurements, when the amount of data exceeds the capacity of manual analysis.

Agreed, the hysteresis index therefore is a useful tool, and if it is used along-side other metrics such as loop-area it can provide detailed information about the loop shape without having to visually examine each loop. See comments above and in response to reviewer 1.

See earlier response and modifications

P 7881 l20-22: the hysteresis shapes are already described before using the method presented in the paper. Maybe specify that this is based on preliminary visual observation of discharge-concentration plots.

Yes, this is based on visual inspection, this was done to ensure that a large range of loop shape and sizes were available to thoroughly test the proposed new method. Text will be added to clarify this point.

Text has been added lines 153-154

Technical Note: Testing an improved index for analysing storm ~~nutrient hysteresis~~discharge-concentration hysteresis.

Lloyd, C.E.M.[1,2], Freer, J.E.[2], Johnes, P.J.[2] and Collins, A.L.[3]

[1] School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.

[2] School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS, UK.

[3]Department of Sustainable Soils and Grassland Systems, Rothamsted Research, North Wyke, Okehampton, EX20 2SB, UK

* Corresponding author: Charlotte Lloyd
Address: School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.
e-mail: charlotte.lloyd@bristol.ac.uk
Telephone: +44 (0)117 3316795 Fax: +44 (0)117 927 7985

Abstract

Analysis of hydrochemical behaviour ~~in extreme flow~~during storm events can provide new insights into the process controls on nutrient transport in catchments. The examination of storm behaviours using hysteresis analysis has increased in recent years, partly due to the increased availability of high temporal resolution datasets for discharge and nutrient parameters. A number of these analyses involve the use of an index to describe the characteristics of a hysteresis loop in order to compare ~~different~~ storm behaviours both within and between catchments. This technical note reviews the methods for calculation of the hysteresis index (HI) and explores a new more effective methodology. Each method is systematically tested and the impact of the chosen calculation on the results is examined. Recommendations are made regarding the most effective method of calculating a HI which can be used for comparing data between storms and between different parameters and catchments.

1. Introduction

The analysis of hysteresis patterns is a key tool for the interrogation of in-stream physical and chemical responses to storm events, which have been shown to be important periods for the transport of nutrients and sediment within catchments (Bowes et al., 2003;Jarvie et al., 2002;Jordan et al., 2007;Burt et al., 2015;Evans and Johnes, 2004). In the context of this paper, hysteresis is defined as the nonlinear relationship between discharge and concentration of nutrients or sediment. When discharge-concentration data are plotted a cyclic pattern is often observed, the strength of the relationship is dependent on the nature of the lag in response between the two variables. Quantification of hysteresis allows multiple storm behaviours to be examined between and within catchments, for a wide range of hydrological and hydrochemical parameters. This can provide insight into catchment function, allowing the development and testing of process-based understanding. This type of analysis has been used in recent years by many authors investigating nutrient concentration-discharge relationships in catchments of differing environmental character (e.g. Bowes et al., 2015;Darwiche-Criado et al., 2015;Cerro et al., 2014;Rodriguez-Blanco et al., 2013;Oeurng et al., 2010;Eder et al., 2010;Evans and Johnes, 2004) but, traditionally, has been used for the examination of turbidity or suspended sediment data (e.g. Ziegler et al., 2014;House and Warwick, 1998;Williams, 1989;Tena et al., 2014;Klein, 1984;Whiting et al., 1999). Hysteresis analysis has been used to support the investigation of the temporal variations in nutrient transport to streams as a means of characterising the likely contributing source areas and flow pathways linking source to stream in complex landscapes (Outram et al., 2014;Bowes et al., 2015;Lloyd et al., 2016). Similar hysteresis patterns can be observed for a variety of different reasons, however it is generally assumed that clockwise hysteresis, caused by a small or no lag between discharge and concentration suggests a source close to the monitoring point. Conversely, anti-clockwise hysteresis generally signifies a longer lag between the discharge and concentration peak, suggesting that the source was located further from the monitoring point. (Williams, 1989) provides a detailed summary of different shape hysteresis plot and the possible mechanisms.

For hysteresis analysis to be effective and easy to interpret there is a need to develop an effective method of classifying storms according to their hysteretic behaviour. Many papers have classified storms into clockwise or anticlockwise responses, and described the strength of the hysteresis as small or large (Bowes et al., 2015;Evans and Davies, 1998;Butturini et al., 2008). Other authors have used an index approach, which allows a dimensionless quantification of the hysteresis, and thus, comparison of hysteresis indices between catchments of differing size, morphology and hydrological function. An index approach is also useful as it provides information about both the direction and strength of the hysteresis. Hysteretic indices proposed by Butturini et al. (2008) provide semi-

quantitative methods to describe whether the measured parameter is enriched or diluted during a storm event and to assess the area inside the hysteresis loop, along with its direction. Langlois et al. (2005) propose a quantitative method which involves splitting the discharge hydrograph into the rising and falling limb and fitting regression lines to each dataset. The hysteresis index is calculated as the ratio (rising:falling) of the areas under the regression curves. Whilst this index provides a quantitative solution, the authors suggest that the method should only be applied to simple uni-directional loops, i.e. not those which exhibit figure-of-eight or more complex behaviours. A quantitative index was also proposed by Lawler et al. (2006), which uses the ratio of the turbidity (or other parameter) concentration on the rising and falling limb, at the mid-point in the discharge. The mid-point in discharge is defined as 50% of the range in discharge during the storm event. This index has been used by a number of other authors (McDonald and Lamoureux, 2009;Outram et al., 2014), as it is flexible and can be applied to hysteresis loops of all shapes. However it is not without limitations. In a recent paper, Aich et al. (2014) highlight that the index of Lawler et al. (2006) in its current form becomes skewed at higher concentrations, with a smaller index calculated for loops of the same shape and area in the case of storms commencing at a higher concentration (Figure 1a). In addition, the calculation of the index using only the mid-point (50%) in discharge can be problematic. Lawler et al. (2006) state that the mid-point was used as it avoids the often noisy sections at the beginning and end of the loops. However, the result of the calculated index may be misleading in many figure-of-eight scenarios, especially those which cross close to the mid-point in discharge (see Figure 1b). The example shown in Figure 1b illustrates that a hysteresis index (HI) calculated at the mid-point in discharge would suggest that there was very little hysteresis, even though there is a strong effect but in different directions during different periods of the storm event. As suggested by Lawler et al. (2006), the HI can be calculated at multiple increments through the flow range and an average HI value gained. Against the above background, this technical note reports the impact of the chosen method on the index values generated from a series of storms of varying size and hysteretic shapes, using an adapted version of the Lawler et al. (2006) index (HI_{LA}). The paper also introduces a new method for calculating the hysteresis index (HI_{new}) and, as a result of this analysis, suggests a recommendation for the most appropriate calculation for a HI for storm-driven nutrient transport in catchments.

2. Methodology

2.1 Datasets

The example uses a series of storms extracted from high-temporal resolution (15-min) data collected on the River Wylye at Brixton Deverill (Wiltshire, UK) as part of the Defra Demonstration Test

Catchment project (McGonigle et al., 2014) from March 2012 to March 2014. Detailed descriptions of the field site and the datasets are available in previously published work (Lloyd et al., 2015, in revision). For the purposes of this study, discharge data were obtained from the Environment Agency gauge (Gauge Number 43806) and turbidity data were collected using a YSI 6-series sonde, which was cleaned and calibrated once a month over the monitoring period. Turbidity (measured in Nephelometric Turbidity Units (NTU)) was chosen for this study as it is the most widely examined parameter in terms of hysteresis and the storms selected from the data set exhibit a wide range of turbidity values and hysteretic shapes. A total of 66 storms were extracted for this analysis from the two year observational data. A storm was classified as an increase in discharge of more than 20% above baseflow and the end of the storm was determined by either a return to baseflow conditions or when discharge began to rise again if another storm occurred before the system had returned to baseflow conditions. Previous work had quantified the uncertainty associated with the discharge and turbidity measurements ~~(Lloyd et al., 2015; Lloyd et al., submitted)~~(Lloyd et al., 2015;Lloyd et al., 2016) and this provided 100 resampled iterations of each measured parameter for every storm, accounting for observational uncertainties, for this analysis. Figure 2a-f(I) shows some example storms, where the boxes represent the 5^{th}- 95^{th} percentile uncertainty range for each data point.

2.2 Lawler et al. (2006) method and modification

The HI was then calculated according to the standard method of Lawler et al. (2006) (HI_L) for combinations of all 100 iterations of each of the storms to provide a distribution of HI when the mid-point in discharge was calculated (50%). The Lawler et al. (2006) method was also adapted (HI_{LA}), where HI was calculated at every 25%, 10%, 5% and 1% increments of the discharge (see Figure 3 for visualisation) as shown below:

if $T_{RL} > T_{FL}$ (clockwise hysteresis):

$$HI_L = \left(\frac{T_{RL}}{T_{FL}}\right) - 1 \qquad (1)$$

Or, if $T_{RL} < T_{FL}$ (anti-clockwise hysteresis):

$$HI_L = \left(-1/\frac{T_{RL}}{T_{FL}}\right) + 1 \qquad (2)$$

Where: T_{RL} is the value of turbidity at a given point in flow on the rising limb and T_{FL} is the value on the falling limb.

When multiple sections per storm were calculated, the average value was taken to represent the HI of the complete storm event. In some cases there were not corresponding values on both the falling and rising limbs, when this occurs the maximum number of available pairs of data were used to calculate the index. This only usually occurred at lowest discharges and when a large number of intervals were being analysed. This meant that the number of missing pairs was small compared with the available pairs (<5%) and as a result had little impact on the overall calculation. The analyses were completed for both the raw data and for normalised storms to assess the impact of the different analysis methods on the HI values obtained. The data were normalised using the following equations:

$$Normalised\ Q_i = \frac{Q_i - Q_{min}}{Q_{max} - Q_{min}} \quad (3)$$

$$Normalised\ T_i = \frac{T_i - T_{min}}{T_{max} - T_{min}} \quad (4)$$

Where: Q_i/T_i is the discharge/turbidity at timestep *i*, Q_{min}/T_{min} is the minimum storm parameter value and Q_{max}/T_{max} is the maximum storm parameter value.

2.3 Proposed new Hysteresis Index method (HI_{new})

A new method of calculating a HI was also tested (HI_{new}) with the aim of eliminating the impact of a changing baseline value on the ratio as multiple measurements are taken from the same storm. The new index uses the difference between the turbidity values on the rising and falling limbs of the normalised storms, rather than a ratio, and effectively normalises the rising limb at every measurement point, thereby resulting in an index between -1 and 1.

$$HI_{new} = T_{RL\_norm} - T_{FL\_norm} \quad (5)$$

As with the other methods, the analysis was carried out using different intervals of discharge (25%, 10%, 5% and 1%) and the mean was used as the final HI value for the storm. The impact of this number of chosen intervals of discharge on the magnitude of the resulting HI was tested.

The resulting distributions of HI values for each method were then scrutinised using boxplots. Differences between the distributions of data for each storm were analysed statistically using ANOVA where normality and variance assumptions were met, and the non-parametric alternative Kruskal-Wallis-H on ranked data where the ANOVA assumptions did not hold. When a significant difference between the groups was detected, a pairwise Tukey test was used to establish which of the groups were contributing to the effect. The main aim of the analysis was to determine the point at which sufficient intervals of discharge were used so that there was no statistically significant difference between the different datasets for each storm.

153 *3. Results and discussion*

154 A total of 66 storms were analysed using the three methods for calculating the HI, which included 35
155 anti-clockwise loops, 11 clockwise loops, 12 figure-of-eight loops which were mainly anti-clockwise
156 and, 8 figure-of-eight loops which were mainly clockwise (loop shapes were examined by visual
157 inspection). The peak turbidity during the storms ranged between 10 and 392 NTU (mean = 91 NTU)
158 and the starting values were between 2 and 31 NTU (mean = 8 NTU). Figure 2 shows six example
159 storms (a-f, panel I) from the range of behaviour identified above, each with varying shape and size.
160 ~~Figure 2a-f (panel II) shows the distibution of values for the calculated H-index using HI_L (measured~~
161 ~~at 50% of discharge range) and the HI_{LA} (measured at varying percentile increments of discharge).~~
162 ~~The grey areas on the plots show the boxplots which were not statistically different from one~~
163 ~~another, that is, there is no gain by increasing the number of intervals of discharge measured for~~
164 ~~that storm.~~ Table 1 summarises the number (and percentage) of storms tested which can be
165 adequately represented by the different discharge interval frequencies tested.

166 Figure 2a-f(II) shows the distributions of HI values (using HI_L) measured at only 50% of discharge are
167 often very different from the analyses which measure multiple sections across the loop (HI_{LA}). The
168 more complex the shape of the loop, the more measured sections are needed to represent it
169 adequately. The analysis shows that by using 5% increments of discharge (19 sections), 98% of the
170 storms analysed showed stable distributions and therefore no significant changes were observed
171 when additional increments were included. While including more increments of the loop in the
172 analysis does improve the HI results, it does not solve all of the issues highlighted earlier. Both HI_L
173 and HI_{LA} are sensitive to the size of the storm and, as a result, for a similar pattern in hysteresis but a
174 larger magnitude of storm, a comparatively smaller value would be calculated for the index, as
175 shown in Figure 1a. This means that the results generated for a series of storms are very difficult to
176 interpret and it is difficult to compare between individual storms and catchments. By normalising
177 the storms as described above and continuing to use the HI_{LA} method, the comparability of the
178 outputs between storms is improved as they are all assessed on the same scale. However, if multiple
179 increments of discharge are included, which has been shown to be beneficial, then effectively each
180 of the individual measured sections of the storm need to be normalised, otherwise the problem is
181 reduced but not eradicated. This problem is illustrated in Figure 1c, which shows an example of an
182 idealised and normalised storm where the width of the loop remains constant through most of the
183 storm. However at different quantiles of flow, HI value varies due to the loop gradient, the HI is
184 inflated towards the lower and reduced at higher quantiles of discharge. The HI_{new} was designed to
185 overcome this problem. The new index uses the range of turbidity values between the rising and

falling limb at each increment of discharge rather than the ratio, thereby directly quantifying the width of the loop.

Figure 4~~3~~ shows how the new index effectively normalises the rising limb and examines the relative behaviour of the falling limb, thereby identifing the proportion of the storm occurring in a clockwise or anti-clockwise phase. For this new method to be robust, it is necessary to normalise the data as described earler before the analysis. Figure 2a-f(III) show the example storms in their normalised forms. The new index produces a value between -1 and 1, where 0 represents no hysteretic pattern and positive values clockwise and negative values, anti-clockwise hysteresis. A figure-of-eight storm will be represented as a weighted average of the intervals of discharge measured when the storm was in a clockwise phase and when it was in an anticlockwise phase. Therefore, for example, if the storm exhibits anti-clockwise behaviour for a large proportion of the storm event the average HI_{new} will produce a negative number. It should be noted that in the unusal case that an exactly symmetrical figure-of-eight storm is presented the index would produce a value of 0, suggesting no hysteresis. Using the HI value in conjuction with loop area will however provide clarification as a storm which has an HI of 0 but a positive loop area has to be a complex loop shape. The advantage with our new technique is that the user can choose to interrogate other output metrics within these results, such as the quantified loop area and the distribution of HI values calculated for each section of the loop in addition to the averaged HI value. By looking at the distribution of values it is simple to identify complex loop shapes such as figure-of-eight (due to both positive and negative values calculated for the various loop sections) and ensures correct interpretation of the HI values. Although we do not explore the advantage of these further analyses here, we suggest they potentially provide a richer analyses of hysteresis dynamics that we aim to explore in future papers.

We suggest t~~T~~he~~is~~ new index provides a consistent approach to the core loop characteristics and therefore is more easily interpretable by the user when comparing behaviour between storms or field sites. Figure 2a-f(IV) show the resulting distributions of HI_{new} generated using varying increments of discharge. The analysis shows that the distribution of calculated values was generally more stable compared with the HI_{LA} method and, in many cases, fewer increments of discharge were necessary to produce a statistically stable representation of the storm loop shape (Table 1). The results demonstrate that increasing the increments to every 10% of discharge allowed 95% of storms and using 5% increments allows 100% of storms to be robustly characterised in terms of their loop shape, meaning that the addition of more sections did not significantly alter the distribution of HI results.

4. Conclusions and recommendations

The concept of using an index to aid the quantification of storm hysteresis has been established for over two decades. However few papers have chosen to use them, perhaps due to the limitations associated with the most common methods. This technical note was designed to test systematically, for the first time, the way that the HI is calculated and to quantify the impact of the chosen method on the results. The analysis has led to a number of recommendations concerning how the HI should be calculated in order to produce results which are both statistically robust and comparable between storms and field sites. This technique is useful when the user's interest is in the relative characteristics of the loop geometries~~This technique is useful when the interest and interpretation is in the core relative characteristics of the loop geometries themselves~~. These recommendations are:

1. Storms should be normalised before analysis so that multiple storms can be robustly compared.
2. A range method, such as the new index (HI_{new}) proposed here, should be used in preference to a ratio method as it produces results which are easier to interpret, allowing quantification of the extent of the hysteresis effect that can be directly compared between contrasting catchments even when the magnitude of the storms varies greatly.
3. Multiple sections of each loop should be analysed so that the extent and direction of the hysteresis can be accounted for throughout the flow range. Sections should be measured at least every 10% of the discharge range, although every 5% is recommended as it is likely, based on our analysis, to produce robust results for almost all storm sizes and shapes.
4. Examine the distribution of HI values calculated across the sections in addition to the averaged value, as this aids robust classification of complex loop shapes, including figure-of-eight loops.

Undertaking the analysis of hysteresis loops using these guidelines improves the clarity of the hysteresis index as a diagnostic tool for the analysis of storms and how discharge-concentration patterns vary. The new index (HI_{new}) is able to describe robustly the shape and direction of a hysteretic pattern in storms of any size, and can be used to compare storms from multiple catchments. This means that the index becomes more useful as it has the potential to become a standardised analytical technique that can be utilised by the water quality research community. Lloyd et al. (2016) illustrates the use of the new hysteresis index to investigate storm behaviours across different nutrient parameters and between contrasting catchments. This study exemplifies the power of having such a summary statistic, as different parameters and field sites can be rapidly and robustly compared. The information provided by the HI_{new} can be used in conjuction with other common metrics such as storm maximum concentration to produce a useful and robust quantitative representation of storm hydrochemical behaviour. Standardising approaches for the calculation of

253 HI would provide a useful tool for assessing storm behaviour. This is timely given the marked
254 increase in the number of catchment scale water quality monitoring initiatives, which are now
255 employing high temporal resolution monitoring to improve understanding of pollution sources and
256 delivery pathways. Our ongoing research is exploring the use of this new index in understanding
257 changing catchment dynamics associated with storm behaviours.

258 *Acknowledgements*

259 The authors gratefully acknowledge the funding provided by Defra project WQ0211 (the Hampshire
260 Avon Demonstration Test Catchment project) and NERC Grant NE/1002200/1 (The Environmental
261 Virtual Observatory Pilot), and the access to the Brixton Deverill gauging site and flow data provided
262 by Geoff Hardwicke at the Environment Agency.

Table 1: showing the increments of discharge measured and the corresponding number of storms (out of 66 analysed) and the percentage of storms which can be robustly* characterised using different HI methods. *Where adding extra measurement sections does not statistically change the distribution of HI vales for a storm.

Percentile increments	Sections measured	Storms (HI_L/HI_{LA})	Storms (HI_{new})
50%	1	5 (8%)	1 (1.5%)
25%	3	34 (52%)	41 (62%)
10%	9	55 (83%)	63 (95%)
5%	19	65 (98%)	66 (100%)
1%	99	66 (100%)	66 (100%)

5. References

Aich, V., Zimmermann, A., and Elsenbeer, H.: Quantification and interpretation of suspended-sediment discharge hysteresis patterns: How much data do we need?, Catena, 122, 120-129, 10.1016/j.catena.2014.06.020, 2014.

Bowes, M. J., House, W. A., and Hodgkinson, R. A.: Phosphorus dynamics along a river continuum, Science of the Total Environment, 313, 199-212, 10.1016/s0048-9697(03)00260-2, 2003.

Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., Gozzard, E., Newman, J. R., and Palmer-Felgate, E. J.: Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration-flow relationships, The Science of the total environment, 511, 608-620, 10.1016/j.scitotenv.2014.12.086, 2015.

Burt, T. P., Worrall, F., Howden, N. J. K., and Anderson, M. G.: Shifts in discharge-concentration relationships as a small catchment recover from severe drought, Hydrological Processes, 29, 498-507, 10.1002/hyp.10169, 2015.

Butturini, A., Alvarez, M., Bernal, S., Vazquez, E., and Sabater, F.: Diversity and temporal sequences of forms of DOC and NO3-discharge responses in an intermittent stream: Predictable or random succession?, J. Geophys. Res.-Biogeosci., 113, 10.1029/2008jg000721, 2008.

Cerro, I., Sanchez-Perez, J. M., Ruiz-Romera, E., and Antiguedad, I.: Variability of particulate (SS, POC) and dissolved (DOC, NO3) matter during storm events in the Alegria agricultural watershed, Hydrological Processes, 28, 2855-2867, 10.1002/hyp.9850, 2014.

Darwiche-Criado, N., Comin, F. A., Sorando, R., and Sanchez-Perez, J. M.: Seasonal variability of NO3-mobilization during flood events in a Mediterranean catchment: The influence of intensive agricultural irrigation, Agriculture Ecosystems & Environment, 200, 208-218, 10.1016/j.agee.2014.11.002, 2015.

Eder, A., Strauss, P., Krueger, T., and Quinton, J. N.: Comparative calculation of suspended sediment loads with respect to hysteresis effects (in the Petzenkirchen catchment, Austria), Journal of Hydrology, 389, 168-176, 10.1016/j.jhydrol.2010.05.043, 2010.

Evans, C., and Davies, T. D.: Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry, Water Resources Research, 34, 129-137, 10.1029/97wr01881, 1998.

Evans, D. J., and Johnes, P.: Physico-chemical controls on phosphorus cycling in two lowland streams. Part 1 - the water column, Science of the Total Environment, 329, 145-163, 10.1016/j.scitotenv.2004.02.016, 2004.

House, W. A., and Warwick, M. S.: Hysteresis of the solute concentration/discharge relationship in rivers during storms, Water Research, 32, 2279-2290, 10.1016/s0043-1354(97)00473-9, 1998.

Jarvie, H. P., Neal, C., Williams, R. J., Neal, M., Wickham, H. D., Hill, L. K., Wade, A. J., Warwick, A., and White, J.: Phosphorus sources, speciation and dynamics in the lowland eutrophic River Kennet, UK, Science of the Total Environment, 282, 175-203, 10.1016/s0048-9697(01)00951-2, 2002.

Jordan, P., Arnscheidt, A., McGrogan, H., and McCormick, S.: Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser, Hydrology and Earth System Sciences, 11, 372-381, 2007.

Klein, M.: ANTI CLOCKWISE HYSTERESIS IN SUSPENDED SEDIMENT CONCENTRATION DURING INDIVIDUAL STORMS - HOLBECK CATCHMENT - YORKSHIRE, ENGLAND, Catena, 11, 251-257, 10.1016/s0341-8162(84)80024-7, 1984.

Langlois, J. L., Johnson, D. W., and Mehuys, G. R.: Suspended sediment dynamics associated with snowmelt runoff in a small mountain stream of Lake Tahoe (Nevada), Hydrological Processes, 19, 3569-3580, 10.1002/hyp.5844, 2005.

Lawler, D. M., Petts, G. E., Foster, I. D. L., and Harper, S.: Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK, Science of the Total Environment, 360, 109-126, 10.1016/j.scitotenv.2005.08.032, 2006.

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., Coxon, G., and Collins, A. L.: Discharge and nutrient uncertainty: implications for nutrient flux estimation in small streams, Hydrological Processes, n/a-n/a, 10.1002/hyp.10574, 2015.
Lloyd, C. E. M., Freer, J. E., Johnes, P. J., and Collins, A. L.: Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments, Science of The Total Environment, 543, Part A, 388-404, http://dx.doi.org/10.1016/j.scitotenv.2015.11.028, 2016.
McDonald, D. M., and Lamoureux, S. F.: Hydroclimatic and channel snowpack controls over suspended sediment and grain size transport in a High Arctic catchment, Earth Surface Processes and Landforms, 34, 424-436, 10.1002/esp.1751, 2009.
McGonigle, D. F., Burke, S. P., Collins, A. L., Gartner, R., Haft, M. R., Harris, R. C., Haygarth, P. M., Hedges, M. C., Hiscock, K. M., and Lovett, A. A.: Developing Demonstration Test Catchments as a platform for transdisciplinary land management research in England and Wales, Environmental Science: Processes & Impacts, 10.1039/C3EM00658A, 2014.
Oeurng, C., Sauvage, S., and Sanchez-Perez, J.-M.: Temporal variability of nitrate transport through hydrological response during flood events within a large agricultural catchment in south-west France, Science of the Total Environment, 409, 140-149, 10.1016/j.scitotenv.2010.09.006, 2010.
Outram, F. N., Lloyd, C. E. M., Jonczyk, J., Benskin, C. M. H., Grant, F., Perks, M. T., Deasy, C., Burke, S. P., Collins, A. L., Freer, J., Haygarth, P. M., Hiscock, K. M., Johnes, P. J., and Lovett, A. L.: High-frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of the 2011–2012 drought in England, Hydrol. Earth Syst. Sci., 18, 3429-3448, 10.5194/hess-18-3429-2014, 2014.
Rodriguez-Blanco, M. L., Taboada-Castro, M. M., and Taboada-Castro, M. T.: Phosphorus transport into a stream draining from a mixed land use catchment in Galicia (NW Spain): Significance of runoff events, Journal of Hydrology, 481, 12-21, 10.1016/j.jhydrol.2012.11.046, 2013.
Tena, A., Vericat, D., and Batalla, R. J.: Suspended sediment dynamics during flushing flows in a large impounded river (the lower River Ebro), Journal of Soils and Sediments, 14, 2057-2069, 10.1007/s11368-014-0987-0, 2014.
Whiting, P. J., Samm, J. F., Moog, D. B., and Orndorff, R. L.: Sediment-transporting flows in headwater streams, Geological Society of America Bulletin, 111, 450-466, 10.1130/0016-7606(1999)111<0450:stfihs>2.3.co;2, 1999.
Williams, G. P.: Sediment concentration versus water discharge during single hydrologic events in rivers, Journal of Hydrology, 111, 89-106, 10.1016/0022-1694(89)90254-0, 1989.
Ziegler, A. D., Benner, S. G., Tantasirin, C., Wood, S. H., Sutherland, R. A., Sidle, R. C., Jachowski, N., Nullet, M. A., Xi, L. X., Snidvongs, A., Giambelluca, T. W., and Fox, J. M.: Turbidity-based sediment monitoring in northern Thailand: Hysteresis, variability, and uncertainty, Journal of Hydrology, 519, 2020-2039, 10.1016/j.jhydrol.2014.09.010, 2014.

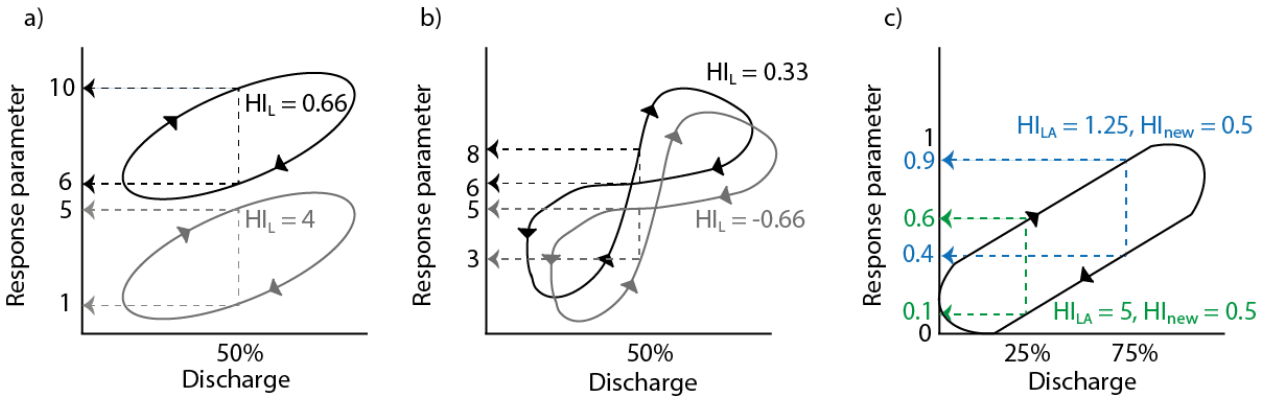

Figure 1: Plots showing a) impact of storm initial concentration, b) storm initial discharge on the value of the calculated HI when the mid-point in discharge and raw data is used and c) an idealised and normalised storm illustrating the impact of measuring different quantiles of flow on the HI calculated. Where HI_L and HI_{LA} are the original and adapted Lawler et al. (2006) methods, respectively and HI_{new}, the proposed new method. Colours represent different discharge intervals measured.

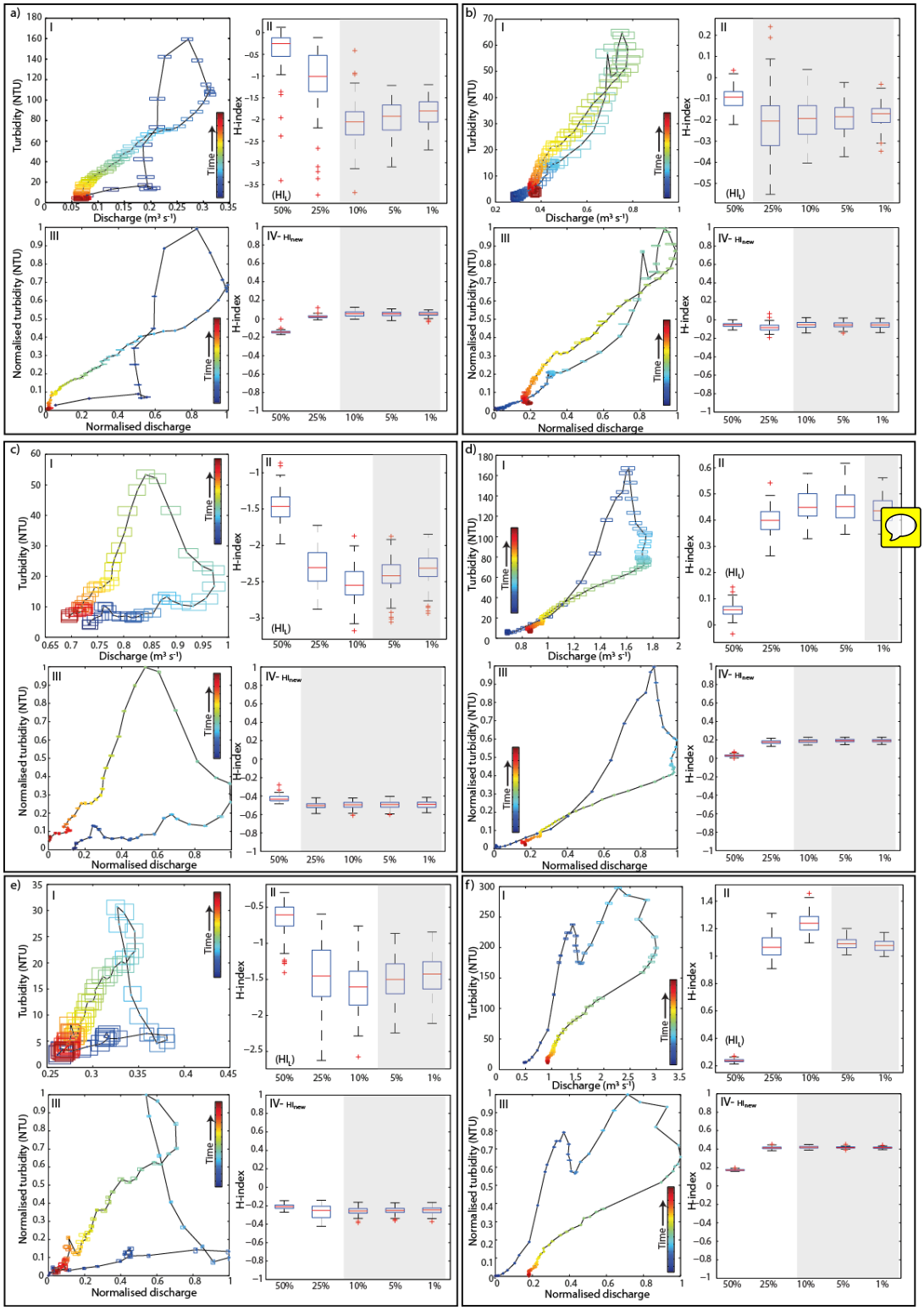

Figure 2: Plots showing six storms with varying loop shapes and sizes (a-f), where (I) is the hysteresis loop using the raw data, (II) is the distribution of HI values using the original and adapted Lawler et al. (2006) methods (HI_L/HI_{LA}) using varying percentiles of flow, (III) is the hysteresis loop plotted using normalised data, and (IV) is the distribution of HI values using the new method (HI_{new}) using varying percentiles of flow. The grey areas show the distributions which are not statistically different from each other. In panels I and III, the black line represents the median and the boxes represent the 5th-95th percentiles of the uncertainty range.

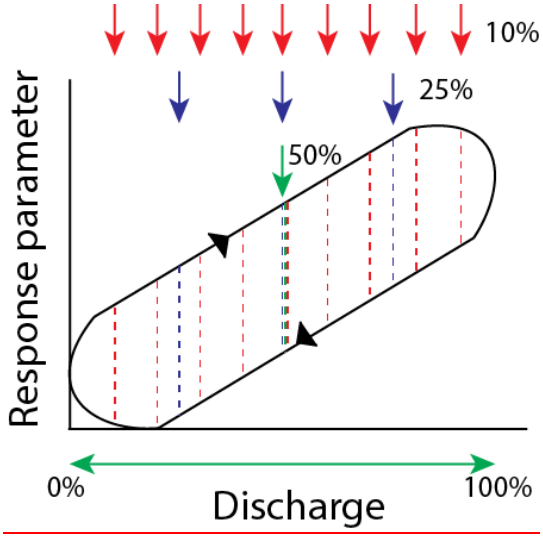

Figure 3: diagram showing examples of how the sampling intervals for the calculation of the HI_{LA} and HI_{new} are determined.

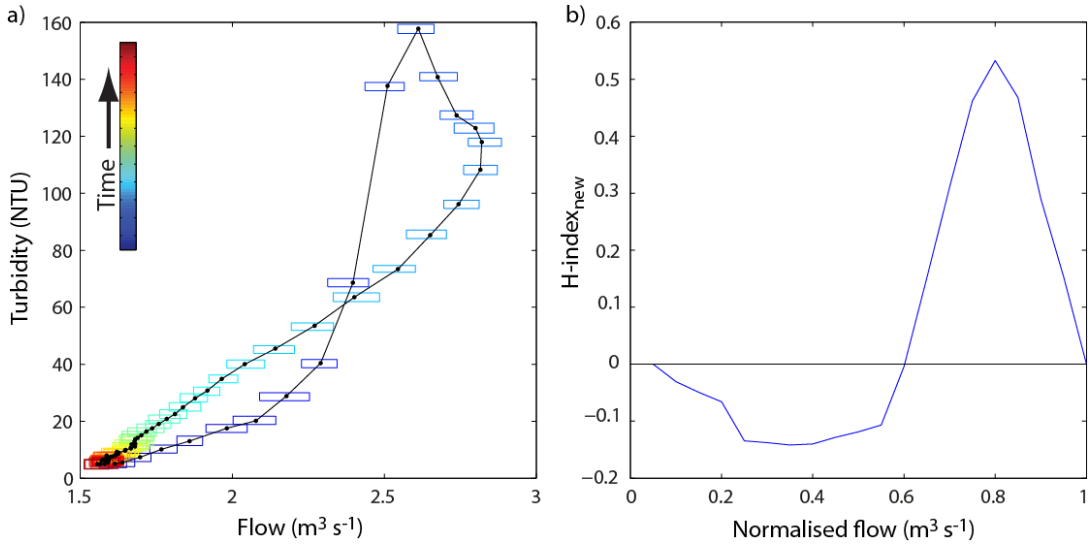

Figure 4~~3~~: showing a) the original storm, where the black line represents the median and the boxes the 5th-95th percentiles of the uncertainty around the line, and b) illustrates the HI_{new} of the normalised storm.