

Response to the editor's comments.

The authors thanks the editor for their constructive and helpful comments, which have helped improve the manuscript. Below provides details of changes made as a result of these comments, shown in red text.

Ln 16: rephrase? water quality parameters? if this also relates to turbidity than nutrient parameters is not the correct term

Agreed, changed to water quality parameters.

Ln 22: what do you mean here? different water quality parameters?

Yes text changed to water quality parameters.

Ln 31: what do you mean here? do you mean: "... where the pattern is dependent on the time lag between the two variables" and Ln 33: what do you mean by nature of the lag?

Sentence has been reworded to read "When discharge-concentration data are plotted a cyclic pattern is often observed, the size and shape of the loop is dependent on the lag in response between the discharge and water quality variables".

Ln 34: what do you mean by wide range of hydrological parameters? which other parameters than discharge? Maybe also mention range of system states/intial conditions/antecedent wetness (whatever terminology you prefer)

Sentence has been amended to read "Quantification of hysteresis allows multiple storm behaviours to be examined between and within catchments as well as under varying antecedent conditions, for discharge and a wide range of hydrochemical parameters."

Ln 37: hypotheses?

Agreed text has been changed to hypotheses.

Ln 46: Similar to what?

Text amended t: "Hysteresis patterns with similar characteristics can be observed for a variety of different reasons...."

Ln 47: As I understand it clockwise hysteresis occurs when concentrations increase first faster than Q at low flows and then slower than Q at higher flows for the rising limb of the hydrograph (the peak of concentration can occur either before or at the same time as the peak of Q) and then concentrations usually decrease faster than Q on the first stretch of the falling limb. Small or no lag between discharge and concentration does not necessarily mean that hysteresis occurs.

Yes you are correct, the text has therefore been corrected to read "however it is generally assumed that clockwise hysteresis, caused by concentrations increasing more rapidly than discharge during the rising limb, suggests a source close to the monitoring point"

Ln 51: not clear - do you mean "differently shaped hysteresis plots and the possible underlying mechanisms?"

Text has been clarified to read "Williams (1989) provides a detailed summary of different shape hysteresis plots and the possible underlying mechanisms."

Ln 55: two uses for effective in the same sentence

The second effective has been changed to robust.

Ln 96: is this MS still in revision?

The reference has been updated, this manuscript is now published in Science of the Total Environment:

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

Ln 122: clarify if you are referring to the hydrograph or of the concentration curve

Text clarified to "on the rising limb of the hydrograph".

Ln 127: usually only?

Agreed, text amended.

Ln 158: not entirely clear - do you mean identified?

Yes, text amended to "identified".

Ln 162: not clear, please rephrase

Text elaborated to read "Table 1 summarises the number (and percentage) of storms tested which can be adequately represented by calculating the HI values using each of the different discharge interval frequencies stated in Section 2.2."

Ln 184: should be difference and not range as stated in the methods and Ln 185: should probably be "normalized turbidity values"

Agreed, sentence amended to read "The new index uses the difference between the normalised turbidity values on the rising and falling limb at each increment of discharge rather than the ratio, thereby directly quantifying the width of the loop".

Ln 224: Suggestion: Move this sentence to line 223 and then end the following sentence with a colon: "... a number of recommendations...between storms and field sites:" Then delete the sentence "These recommendations are:"

Thanks, this is a good suggestion and the text has been amended accordingly.

L229: difference?

Correct, the text has been updated.

Ln 234: do you really mean measured in the field or are you referring to the calculation of the difference index at each 10% section?

Thanks, this has been clarified by changing measured to calculated

Ln 238: all other recommendations are phrased as "... should..." recommendations - are you breaking this pattern up here in recommendation 4 on purpose?

This is a good point, the last bullet point has been reworded so that it matches the structure of the other three recommendations.

Ln 248: water quality? if this also refers to turbidity then nutrients isn't the correct term.

Agreed, this has been changed to water quality parameters.

Ln 248: which study are you talking about here? The cited one or the one presented in this manuscript? If it is the one you are citing it might become clearer if you say "The study..." instead of "This study..."

This is a good point, text has been amended to "The cited study" to clarify.

Ln 252: It seems like this has already been stated in line 245-246.

Agreed, the repetition has therefore been deleted.

Ln 256: what do you mean by changing? Changing in which context?

The sentence has been clarified to now read "Our ongoing research is exploring the use of this new index in understanding differences in catchment dynamics associated with storm behaviours."

Table 1: To reduce confusion I would only put HI_LA in the header and then add "... = HI_L" in the first line.

This is a good suggestion, thanks. The table has been edited as recommended.

Reference: what does this mean?

This can now be updated, had been in Early View online in Hydrological Processes since 29th July, and has only just been given a volume and issue number: Vol 30, Issue 1 (1st Jan 2016)

Figure 2: how can you only have one boxplot in the grey area if the grey area identifies distributions that do not differ from each other?

Agreed, there were no distributions which did not differ in that case, therefore the grey area has been removed.

Figure 3: Please explain in the caption what the red, blue and green arrows and the dashed lines mean. It might also be helpful to mark the corresponding points on the rising and falling limb which are used to calculate the difference.

The diagram has been amended with dots added on the rising and falling limb to show the points used to calculate the HI as suggested. The figure caption has also been updated to: Figure 3: diagram showing examples of how the sampling intervals for the calculation of the HI_{LA} and HI_{new} are determined. The coloured arrows and dashed lines illustrate the position of sections used for the calculation of the HI, where 50%, 25% or 10% intervals are used. The coloured dots show the positions on the rising and falling limbs used to calculate the HI.

1 **Technical Note: Testing an improved index for analysing storm discharge-concentration hysteresis.**

2 Lloyd, C.E.M.^{1,2}, Freer, J.E.², Johnes, P.J.² and Collins, A.L.³

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

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

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

7 * Corresponding author: Charlotte Lloyd

8 Address: School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.

9 e-mail: charlotte.lloyd@bristol.ac.uk

10 Telephone: +44 (0)117 3316795 Fax: +44 (0)117 927 7985

11

12 *Abstract*

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

24

25 1. Introduction

26 The analysis of hysteresis patterns is a key tool for the interrogation of in-stream physical and chemical
27 responses to storm events, which have been shown to be important periods for the transport of
28 nutrients and sediment within catchments (Bowes et al., 2003;Jarvie et al., 2002;Jordan et al.,
29 2007;Burt et al., 2015;Evans and Johnes, 2004). In the context of this paper, hysteresis is defined as
30 the nonlinear relationship between discharge and concentration of nutrients or sediment. When
31 discharge-concentration data are plotted a cyclic pattern is often observed, the size and shape of the
32 strength of the relationship loop is dependent on the ~~nature of~~ the lag in response between the
33 discharge and water quality ~~two~~ variables. Quantification of hysteresis allows multiple storm
34 behaviours to be examined between and within catchments as well as under varying antecedent
35 conditions, for for a wide range of hydrological discharge and a wide range of hydrochemical
36 parameters. This can provide insight into catchment function, allowing the development and testing
37 of process-based understanding hypotheses. This type of analysis has been used in recent years by
38 many authors investigating nutrient concentration-discharge relationships in catchments of differing
39 environmental character (e.g. Bowes et al., 2015;Darwiche-Criado et al., 2015;Cerro et al.,
40 2014;Rodriguez-Blanco et al., 2013;Oeurng et al., 2010;Eder et al., 2010;Evans and Johnes, 2004) but,
41 traditionally, has been used for the examination of turbidity or suspended sediment data (e.g. Ziegler
42 et al., 2014;House and Warwick, 1998;Williams, 1989;Tena et al., 2014;Klein, 1984;Whiting et al.,
43 1999). Hysteresis analysis has been used to support the investigation of the temporal variations in
44 nutrient transport to streams as a means of characterising the likely contributing source areas and
45 flow pathways linking source to stream in complex landscapes (Outram et al., 2014;Bowes et al.,
46 2015;Lloyd et al., 2016a). ~~Similar h~~Hysteresis patterns with similar characteristics can be observed for
47 a variety of different reasons, however it is generally assumed that clockwise hysteresis, caused by
48 concentrations increasing more rapidly than discharge during the rising limb, a small or no lag between
49 discharge and concentration suggests a source close to the monitoring point. Conversely, anti-
50 clockwise hysteresis generally signifies a longer lag between the discharge and concentration peak,
51 suggesting that the source was located further from the monitoring point. Williams (1989) provides a
52 detailed summary of different shape hysteresis plots and the possible underlying mechanisms.

53 For hysteresis analysis to be effective and easy to interpret there is a need to develop an effective
54 robust method of classifying storms according to their hysteretic behaviour. Many papers have
55 classified storms into clockwise or anticlockwise responses, and described the strength of the
56 hysteresis as small or large (Bowes et al., 2015;Evans and Davies, 1998;Butturini et al., 2008). Other
57 authors have used an index approach, which allows a dimensionless quantification of the hysteresis,
58 and thus, comparison of hysteresis indices between catchments of differing size, morphology and

59 hydrological function. An index approach is also useful as it provides information about both the
60 direction and strength of the hysteresis. Hysteretic indices proposed by Butturini et al. (2008) provide
61 semi-quantitative methods to describe whether the measured parameter is enriched or diluted during
62 a storm event and to assess the area inside the hysteresis loop, along with its direction. Langlois et al.
63 (2005) propose a quantitative method which involves splitting the discharge hydrograph into the rising
64 and falling limb and fitting regression lines to each dataset. The hysteresis index is calculated as the
65 ratio (rising:falling) of the areas under the regression curves. Whilst this index provides a quantitative
66 solution, the authors suggest that the method should only be applied to simple uni-directional loops,
67 i.e. not those which exhibit figure-of-eight or more complex behaviours. A quantitative index was also
68 proposed by Lawler et al. (2006), which uses the ratio of the turbidity (or other parameter)
69 concentration on the rising and falling limb, at the mid-point in the discharge. The mid-point in
70 discharge is defined as 50% of the range in discharge during the storm event. This index has been used
71 by a number of other authors (McDonald and Lamoureux, 2009; Outram et al., 2014), as it is flexible
72 and can be applied to hysteresis loops of all shapes. However it is not without limitations. In a recent
73 paper, Aich et al. (2014) highlight that the index of Lawler et al. (2006) in its current form becomes
74 skewed at higher concentrations, with a smaller index calculated for loops of the same shape and area
75 in the case of storms commencing at a higher concentration (Figure 1a). In addition, the calculation of
76 the index using only the mid-point (50%) in discharge can be problematic. Lawler et al. (2006) state
77 that the mid-point was used as it avoids the often noisy sections at the beginning and end of the loops.
78 However, the result of the calculated index may be misleading in many figure-of-eight scenarios,
79 especially those which cross close to the mid-point in discharge (see Figure 1b). The example shown
80 in Figure 1b illustrates that a hysteresis index (HI) calculated at the mid-point in discharge would
81 suggest that there was very little hysteresis, even though there is a strong effect but in different
82 directions during different periods of the storm event. As suggested by Lawler et al. (2006), the HI can
83 be calculated at multiple increments through the flow range and an average HI value gained. Against
84 the above background, this technical note reports the impact of the chosen method on the index
85 values generated from a series of storms of varying size and hysteretic shapes, using an adapted
86 version of the Lawler et al. (2006) index (HI_{LA}). The paper also introduces a new method for calculating
87 the hysteresis index (HI_{new}) and, as a result of this analysis, suggests a recommendation for the most
88 appropriate calculation for a HI for storm-driven nutrient transport in catchments.

89 *2. Methodology*

90 *2.1 Datasets*

91 The example uses a series of storms extracted from high-temporal resolution (15-min) data collected
92 on the River Wylye at Brixton Deverill (Wiltshire, UK) as part of the Defra Demonstration Test
93 Catchment project (McGonigle et al., 2014) from March 2012 to March 2014. Detailed descriptions of
94 the field site and the datasets are available in previously published work (Lloyd et al., 2016a;Lloyd et
95 al., 2016b)(Lloyd et al., 2015, in revision). For the purposes of this study, discharge data were obtained
96 from the Environment Agency gauge (Gauge Number 43806) and turbidity data were collected using
97 a YSI 6-series sonde, which was cleaned and calibrated once a month over the monitoring period.
98 Turbidity (measured in Nephelometric Turbidity Units (NTU)) was chosen for this study as it is the most
99 widely examined parameter in terms of hysteresis and the storms selected from the data set exhibit
100 a wide range of turbidity values and hysteretic shapes. A total of 66 storms were extracted for this
101 analysis from the two year observational data. A storm was classified as an increase in discharge of
102 more than 20% above baseflow and the end of the storm was determined by either a return to
103 baseflow conditions or when discharge began to rise again if another storm occurred before the
104 system had returned to baseflow conditions. Previous work had quantified the uncertainty associated
105 with the discharge and turbidity measurements (Lloyd et al., 2016a;Lloyd et al., 2016b) and this
106 provided 100 resampled iterations of each measured parameter for every storm, accounting for
107 observational uncertainties, for this analysis. Figure 2a-f(l) shows some example storms, where the
108 boxes represent the 5th- 95th percentile uncertainty range for each data point.

109 2.2 Lawler et al. (2006) method and modification

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

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

$$116 \quad HI_L = \left(\frac{T_{RL}}{T_{FL}} \right) - 1 \quad (1)$$

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

$$118 \quad HI_L = \left(-1 / \frac{T_{RL}}{T_{FL}} \right) + 1 \quad (2)$$

119

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

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

$$130 \quad \text{Normalised } Q_i = \frac{Q_i - Q_{min}}{Q_{max} - Q_{min}} \quad (3)$$

$$131 \quad \text{Normalised } T_i = \frac{T_i - T_{min}}{T_{max} - T_{min}} \quad (4)$$

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

134 2.3 Proposed new Hysteresis Index method (HI_{new})

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

$$140 \quad HI_{new} = T_{RL_{norm}} - T_{FL_{norm}} \quad (5)$$

141 As with the other methods, the analysis was carried out using different intervals of discharge (25%,
 142 10%, 5% and 1%) and the mean was used as the final HI value for the storm. The impact of this number
 143 of chosen intervals of discharge on the magnitude of the resulting HI was tested.

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

152 3. Results and discussion

153 A total of 66 storms were analysed using the three methods for calculating the HI, which included 35
154 anti-clockwise loops, 11 clockwise loops, 12 figure-of-eight loops which were mainly anti-clockwise
155 and, 8 figure-of-eight loops which were mainly clockwise (loop shapes were ~~examined-identified~~ by
156 visual inspection). The peak turbidity during the storms ranged between 10 and 392 NTU (mean = 91
157 NTU) and the starting values were between 2 and 31 NTU (mean = 8 NTU). Figure 2 shows six example
158 storms (a-f, panel I) from the range of behaviour identified above, each with varying shape and size.
159 Table 1 summarises the number (and percentage) of storms tested which can be adequately
160 represented by calculating the HI values using each of the different discharge interval frequencies
161 stated in Section 2.2~~tested~~.

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

184 Figure 4 shows how the new index effectively normalises the rising limb and examines the relative
185 behaviour of the falling limb, thereby identifying the proportion of the storm occurring in a clockwise
186 or anti-clockwise phase. For this new method to be robust, it is necessary to normalise the data as

187 described earlier before the analysis. Figure 2a-f(III) show the example storms in their normalised
188 forms. The new index produces a value between -1 and 1, where 0 represents no hysteretic pattern
189 and positive values clockwise and negative values, anti-clockwise hysteresis. A figure-of-eight storm
190 will be represented as a weighted average of the intervals of discharge measured when the storm was
191 in a clockwise phase and when it was in an anticlockwise phase. Therefore, for example, if the storm
192 exhibits anti-clockwise behaviour for a large proportion of the storm event the average HI_{new} will
193 produce a negative number. It should be noted that in the unusual case that an exactly symmetrical
194 figure-of-eight storm is presented the index would produce a value of 0, suggesting no hysteresis.
195 Using the HI value in conjunction with loop area will however provide clarification as a storm which has
196 an HI of 0 but a positive loop area has to be a complex loop shape. The advantage with our new
197 technique is that the user can choose to interrogate other output metrics within these results, such
198 as the quantified loop area and the distribution of HI values calculated for each section of the loop in
199 addition to the averaged HI value. By looking at the distribution of values it is simple to identify
200 complex loop shapes such as figure-of-eight (due to both positive and negative values calculated for
201 the various loop sections) and ensures correct interpretation of the HI values. Although we do not
202 explore the advantage of these further analyses here, we suggest they potentially provide a richer
203 analyses of hysteresis dynamics that we aim to explore in future papers.

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

213 *4. Conclusions and recommendations*

214 The concept of using an index to aid the quantification of storm hysteresis has been established for
215 over two decades. However few papers have chosen to use them, perhaps due to the limitations
216 associated with the most common methods. This technical note was designed to test systematically,
217 for the first time, the way that the HI is calculated and to quantify the impact of the chosen method
218 on the results. This technique is useful when the user's interest is in the relative characteristics of the
219 loop geometries. The analysis has led to a number of recommendations concerning how the HI should

220 be calculated in order to produce results which are both statistically robust and comparable between
221 storms and field sites. ~~This technique is useful when the user's interest is in the relative characteristics~~
222 ~~of the loop geometries. These recommendations are:~~

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

236 Undertaking the analysis of hysteresis loops using these guidelines improves the clarity of the
237 hysteresis index as a diagnostic tool for the analysis of storms and how discharge-concentration
238 patterns vary. The new index (HI_{new}) is able to describe robustly the shape and direction of a hysteretic
239 pattern in storms of any size, and can be used to compare storms from multiple catchments. This
240 means that the index becomes more useful as it has the potential to become a standardised analytical
241 technique that can be utilised by the water quality research community. Lloyd et al. (2016a) illustrates
242 the use of the new hysteresis index to investigate storm behaviours across different water
243 quality nutrient parameters and between contrasting catchments. ~~The~~ is cited study exemplifies the
244 power of having such a summary statistic, as different parameters and field sites can be rapidly and
245 robustly compared. The information provided by the HI_{new} can be used in conjunction with other
246 common metrics such as storm maximum concentration to produce a useful and robust quantitative
247 representation of storm hydrochemical behaviour. ~~Standardising approaches for the calculation of HI~~
248 ~~would provide a useful tool for assessing storm behaviour.~~ This is timely given the marked increase in
249 the number of catchment scale water quality monitoring initiatives, which are now employing high
250 temporal resolution monitoring to improve understanding of pollution sources and delivery pathways.
251 Our ongoing research is exploring the use of this new index in understanding changing differences in
252 catchment dynamics associated with storm behaviours.

253 *Acknowledgements*

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

258

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

Percentile Increments	Sections measured	Storms (HI_L /HI _{LA})	Storms (HI _{new})
50% (=HI _L)	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%)

263

264

265 5. References

266

267 Aich, V., Zimmermann, A., and Elsenbeer, H.: Quantification and interpretation of suspended-
268 sediment discharge hysteresis patterns: How much data do we need?, *Catena*, 122, 120-129,
269 10.1016/j.catena.2014.06.020, 2014.

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

272 Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., Gozzard, E.,
273 Newman, J. R., and Palmer-Felgate, E. J.: Characterising phosphorus and nitrate inputs to a rural river
274 using high-frequency concentration-flow relationships, *The Science of the total environment*, 511,
275 608-620, 10.1016/j.scitotenv.2014.12.086, 2015.

276 Burt, T. P., Worrall, F., Howden, N. J. K., and Anderson, M. G.: Shifts in discharge-concentration
277 relationships as a small catchment recover from severe drought, *Hydrological Processes*, 29, 498-
278 507, 10.1002/hyp.10169, 2015.

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

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

285 Darwiche-Criado, N., Comin, F. A., Sorando, R., and Sanchez-Perez, J. M.: Seasonal variability of NO₃-
286 mobilization during flood events in a Mediterranean catchment: The influence of intensive
287 agricultural irrigation, *Agriculture Ecosystems & Environment*, 200, 208-218,
288 10.1016/j.agee.2014.11.002, 2015.

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

292 Evans, C., and Davies, T. D.: Causes of concentration/discharge hysteresis and its potential as a tool
293 for analysis of episode hydrochemistry, *Water Resources Research*, 34, 129-137,
294 10.1029/97wr01881, 1998.

295 Evans, D. J., and Johnes, P.: Physico-chemical controls on phosphorus cycling in two lowland streams.
296 Part 1 - the water column, *Science of the Total Environment*, 329, 145-163,
297 10.1016/j.scitotenv.2004.02.016, 2004.

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

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

303 Jordan, P., Arnscheidt, A., McGrogan, H., and McCormick, S.: Characterising phosphorus transfers in
304 rural catchments using a continuous bank-side analyser, *Hydrology and Earth System Sciences*, 11,
305 372-381, 2007.

306 Klein, M.: ANTI CLOCKWISE HYSTERESIS IN SUSPENDED SEDIMENT CONCENTRATION DURING
307 INDIVIDUAL STORMS - HOLBECK CATCHMENT - YORKSHIRE, ENGLAND, *Catena*, 11, 251-257,
308 10.1016/s0341-8162(84)80024-7, 1984.

309 Langlois, J. L., Johnson, D. W., and Mehuys, G. R.: Suspended sediment dynamics associated with
310 snowmelt runoff in a small mountain stream of Lake Tahoe (Nevada), *Hydrological Processes*, 19,
311 3569-3580, 10.1002/hyp.5844, 2005.

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

315 Lloyd, C. E. M., Freer, J. E., Johnes, P. J., and Collins, A. L.: Using hysteresis analysis of high-resolution
316 water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment
317 transfer in catchments, *Science of The Total Environment*, 543, Part A, 388-404,
318 <http://dx.doi.org/10.1016/j.scitotenv.2015.11.028>, 2016a.

319 Lloyd, C. E. M., Freer, J. E., Johnes, P. J., Coxon, G., and Collins, A. L.: Discharge and nutrient
320 uncertainty: implications for nutrient flux estimation in small streams, *Hydrological Processes*, 30,
321 135-152, 10.1002/hyp.10574, 2016b.

322 McDonald, D. M., and Lamoureux, S. F.: Hydroclimatic and channel snowpack controls over
323 suspended sediment and grain size transport in a High Arctic catchment, *Earth Surface Processes and
324 Landforms*, 34, 424-436, 10.1002/esp.1751, 2009.

325 McGonigle, D. F., Burke, S. P., Collins, A. L., Gartner, R., Haft, M. R., Harris, R. C., Haygarth, P. M.,
326 Hedges, M. C., Hiscock, K. M., and Lovett, A. A.: Developing Demonstration Test Catchments as a
327 platform for transdisciplinary land management research in England and Wales, *Environmental
328 Science: Processes & Impacts*, 10.1039/C3EM00658A, 2014.

329 Oeurng, C., Sauvage, S., and Sanchez-Perez, J.-M.: Temporal variability of nitrate transport through
330 hydrological response during flood events within a large agricultural catchment in south-west
331 France, *Science of the Total Environment*, 409, 140-149, 10.1016/j.scitotenv.2010.09.006, 2010.

332 Outram, F. N., Lloyd, C. E. M., Jonczyk, J., Benskin, C. M. H., Grant, F., Perks, M. T., Deasy, C., Burke,
333 S. P., Collins, A. L., Freer, J., Haygarth, P. M., Hiscock, K. M., Johnes, P. J., and Lovett, A. L.: High-
334 frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of
335 the 2011–2012 drought in England, *Hydrol. Earth Syst. Sci.*, 18, 3429-3448, 10.5194/hess-18-3429-
336 2014, 2014.

337 Rodriguez-Blanco, M. L., Taboada-Castro, M. M., and Taboada-Castro, M. T.: Phosphorus transport
338 into a stream draining from a mixed land use catchment in Galicia (NW Spain): Significance of runoff
339 events, *Journal of Hydrology*, 481, 12-21, 10.1016/j.jhydrol.2012.11.046, 2013.

340 Tena, A., Vericat, D., and Batalla, R. J.: Suspended sediment dynamics during flushing flows in a large
341 impounded river (the lower River Ebro), *Journal of Soils and Sediments*, 14, 2057-2069,
342 10.1007/s11368-014-0987-0, 2014.

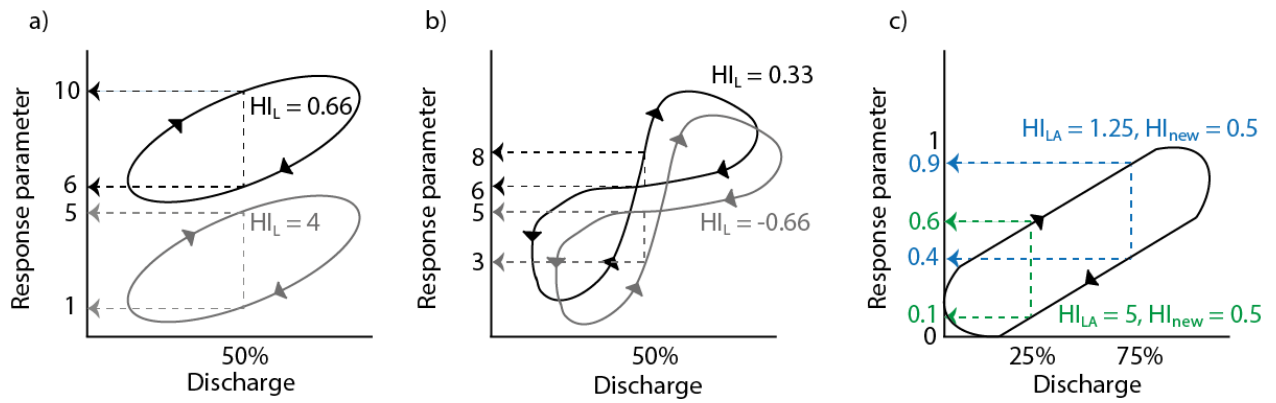
343 Whiting, P. J., Samm, J. F., Moog, D. B., and Orndorff, R. L.: Sediment-transporting flows in
344 headwater streams, *Geological Society of America Bulletin*, 111, 450-466, 10.1130/0016-
345 7606(1999)111<0450:stfihs>2.3.co;2, 1999.

346 Williams, G. P.: Sediment concentration versus water discharge during single hydrologic events in
347 rivers, *Journal of Hydrology*, 111, 89-106, 10.1016/0022-1694(89)90254-0, 1989.

348 Ziegler, A. D., Benner, S. G., Tantasirin, C., Wood, S. H., Sutherland, R. A., Sidle, R. C., Jachowski, N.,
349 Nullet, M. A., Xi, L. X., Snidvongs, A., Giambelluca, T. W., and Fox, J. M.: Turbidity-based sediment
350 monitoring in northern Thailand: Hysteresis, variability, and uncertainty, *Journal of Hydrology*, 519,
351 2020-2039, 10.1016/j.jhydrol.2014.09.010, 2014.

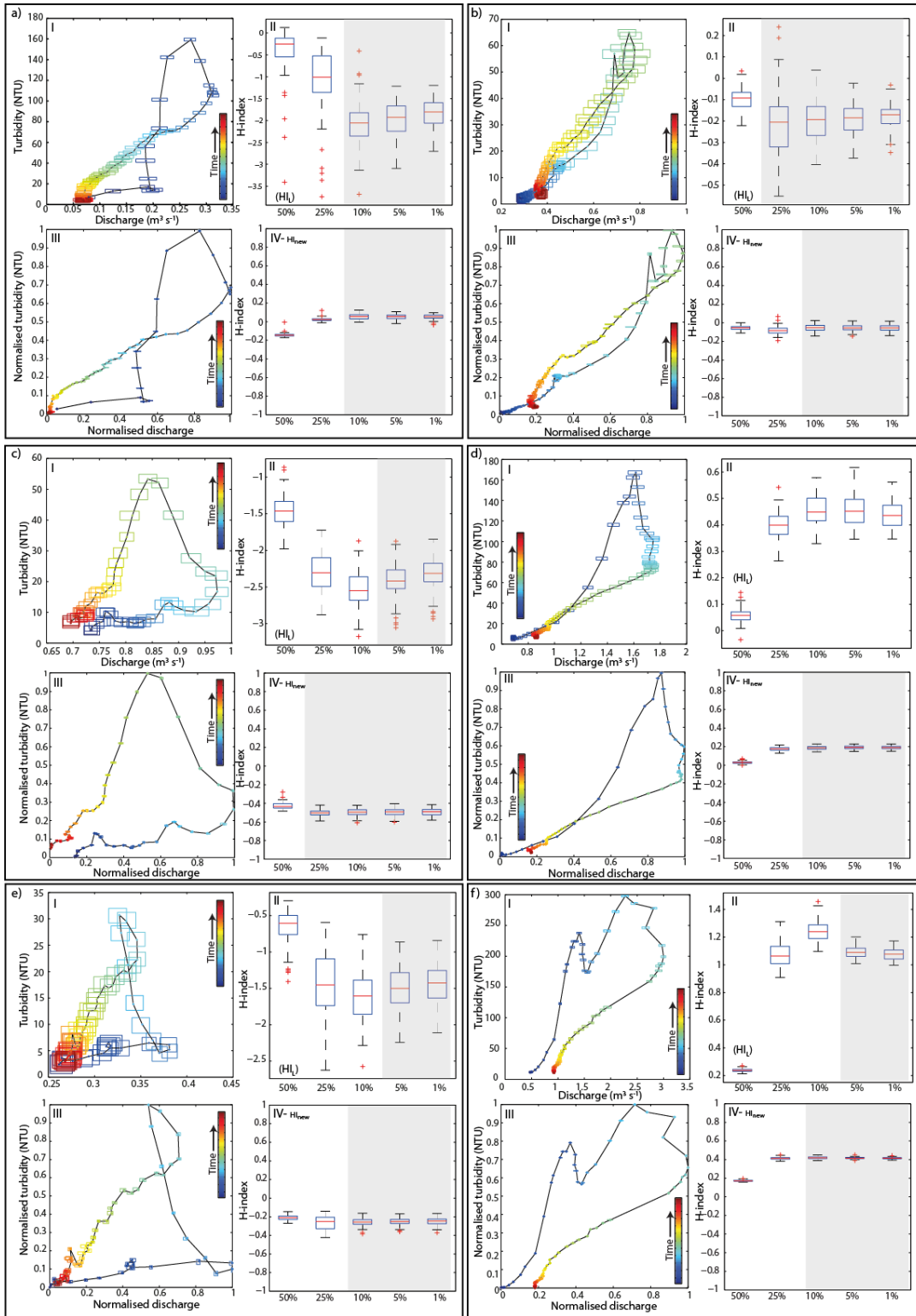
352

353



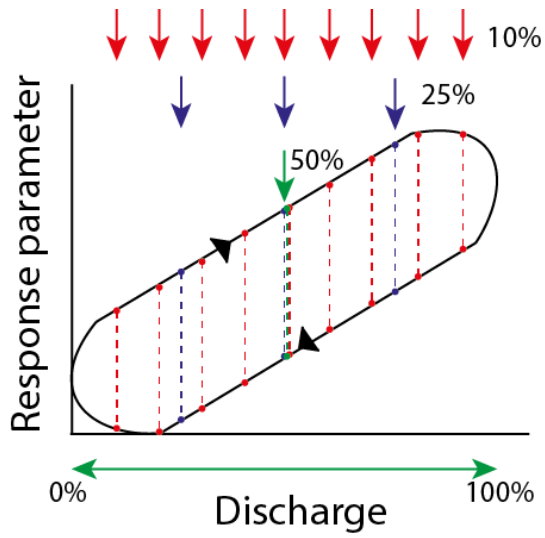
354

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



361

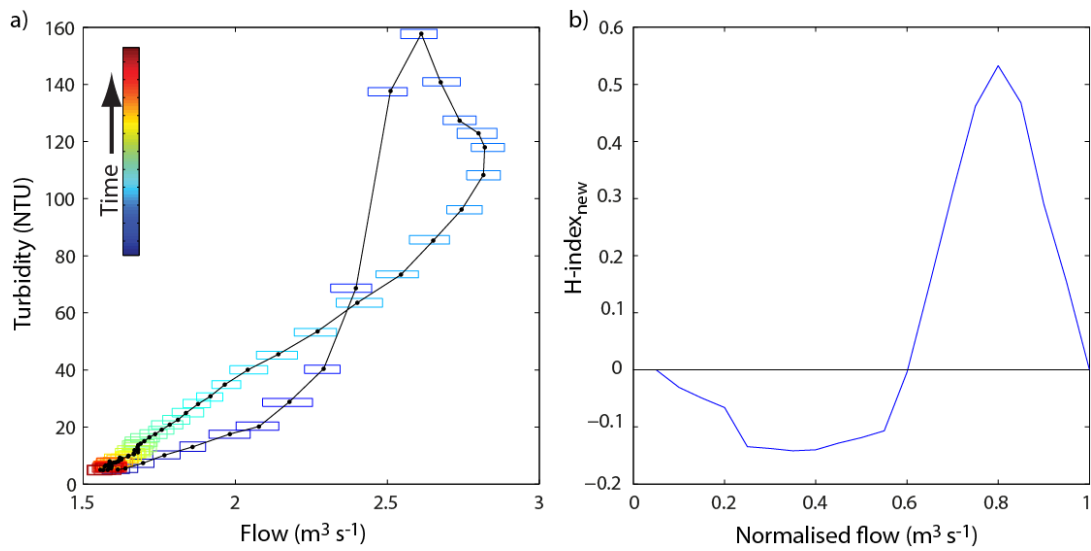
362 Figure 2: Plots showing six storms with varying loop shapes and sizes (a-f), where (I) is the hysteresis
 363 loop using the raw data, (II) is the distribution of HI values using the original and adapted Lawler et
 364 al. (2006) methods (HI_L/HI_{LA}) using varying percentiles of flow, (III) is the hysteresis loop plotted using
 365 normalised data, and (IV) is the distribution of HI values using the new method (HI_{new}) using varying
 366 percentiles of flow. The grey areas show the distributions which are not statistically different from
 367 each other. In panels I and III, the black line represents the median and the boxes represent the 5th-
 368 95th percentiles of the uncertainty range.



369

370 Figure 3: diagram showing examples of how the sampling intervals for the calculation of the HI_{LA} and
 371 HI_{new} are determined. The coloured arrows and dashed lines illustrate the position of sections used
 372 for the calculation of the HI, where 50%, 25% or 10% intervals are used. The coloured dots show the
 373 positions on the rising and falling limbs used to calculate the HI.

374



375

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