6-1-2018

A New Machine-Learning Approach for Classifying Hysteresis in Suspended-Sediment Discharge Relationships Using High-Frequency Monitoring Data

Scott D. Hamshaw
University of Vermont

Mandar M. Dewoolkar
University of Vermont

Andrew W. Schroth
University of Vermont

Beverley C. Wemple
University of Vermont

Donna M. Rizzo
University of Vermont

Recommended Citation

Hamshaw SD, Dewoolkar MM, Schroth AW, Wemple BC, Rizzo DM. A new machine-learning approach for classifying hysteresis in suspended-sediment discharge relationships using high-frequency monitoring data. Water Resources Research. 2018 Jun;54(6):4040-58.

Water Resources Research

RESEARCH ARTICLE
10.1029/2017WR022238

Key Points:
- Storm-event classification is automated using hysteresis images from high-resolution turbidity sensors and a restricted Boltzmann machine
- New hysteresis patterns in suspended-sediment discharge relationships are identified
- Distributions of hysteresis types show linkages in space and season, and provide insight into stream sediment source connectivity

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Hamshaw,
shamshaw@uvm.edu

Citation:
Hamshaw, S. D., Dewoolkar, M. M., Schroth, A. W., Wemple, B. C., & Rizzo, D. M. (2018). A new machine-learning approach for classifying hysteresis in suspended-sediment discharge relationships using high-frequency monitoring data. *Water Resources Research*, *54*, 4040–4058. https://doi.org/10.1029/2017WR022238

Received 13 NOV 2017
Accepted 7 MAY 2018
Accepted article online 14 MAY 2018
Published online 29 JUN 2018

A New Machine-Learning Approach for Classifying Hysteresis in Suspended-Sediment Discharge Relationships Using High-Frequency Monitoring Data

Scott D. Hamshaw[1,2], Mandar M. Dewoolkar[1], Andrew W. Schroth[3], Beverley C. Wemple[4], and Donna M. Rizzo[1]

[1]Department of Civil & Environmental Engineering, College of Engineering and Mathematical Sciences, University of Vermont, Burlington, VT, USA, [2]Now at Vermont EPSCoR, University of Vermont, Burlington, VT, USA, [3]Department of Geology, College of Arts & Sciences, University of Vermont, Burlington, VT, USA, [4]Department of Geography, College of Arts & Sciences, University of Vermont, Burlington, VT, USA

Abstract Studying the hysteretic relationships embedded in high-frequency suspended-sediment concentration and river discharge data over 600^{+} storm events provides insight into the drivers and sources of riverine sediment during storm events. However, the literature to date remains limited to a simple visual classification system (linear, clockwise, counter-clockwise, and figure-eight patterns) or the collapse of hysteresis patterns to an index. This study leverages 3 years of suspended-sediment and discharge data to show proof-of-concept for automating the classification and assessment of event sediment dynamics using machine learning. Across all catchment sites, 600^{+} storm events were captured and classified into 14 hysteresis patterns. Event classification was automated using a restricted Boltzmann machine (RBM), a type of artificial neural network, trained on 2-D images of the suspended-sediment discharge (hysteresis) plots. Expansion of the hysteresis patterns to 14 classes allowed for new insight into drivers of the sediment-discharge event dynamics including spatial scale, antecedent conditions, hydrology, and rainfall. The probabilistic RBM correctly classified hysteresis patterns (to the exact class or next most similar class) 70% of the time. With increased availability of high-frequency sensor data, this approach can be used to inform watershed management efforts to identify sediment sources and reduce fine sediment export.

Plain Language Summary In this study, the river stage (water level) and amount of suspended sediment (soil particles) within a river and five of its tributaries were monitored for 3 years; more than 600 storm events were captured across all six sites. For each storm event, traces of the sediment concentration and river stage were plotted against each other; and the emerging patterns such as clockwise, counter-clockwise, and figure-eight (hysteresis) loops were grouped into 14 reoccurring patterns. We also developed a machine-learning (artificial intelligence) tool to recognize the 14 patterns using only the visual sediment-stage image, in the same way that handwritten characters are recognized by computers. This allowed classification of the individual storm events to be automated. To better understand what these patterns tell us about the physics associated with the storm events and where on the landscape sediments may originate, we analyzed the 14 storm categories using measured rainfall, soil moisture, sediment, and river level data. The machine-learning tool helped capture the linkages between the visual images and the types and origin of erosion using only data monitored at the river outlet during storm events.

1. Introduction

Quantifying the relationship between riverine sediment export and discharge provides important information for understanding the state of hydrologic systems and ecosystem disturbances/stressors, with implications for downstream water quality. In particular, export of suspended sediment plays a critical role in sediment pollution, water-quality degradation, and ecosystem impairment (Gao, 2008). The association of suspended sediment and sediment-bound nutrients such as particulate phosphorus motivates better characterization and understanding of watershed sediment dynamics, nutrient loading, and potential risks (e.g., eutrophication) to aquatic ecosystems (Walling et al., 2008).

Processes associated with suspended-sediment transport during hydrological events can be referred to as event sediment dynamics. Information on these dynamics as well as sediment sources may be inferred from the storm-runoff response and corresponding sediment concentration response (Asselman, 1999; Gao & Josefson, 2012; Lefrançois et al., 2007; Vercruysse et al., 2017). Given the underlying complexity and nonlinear processes controlling sediment transport, the linkage between suspended sediment and discharge over single storm events often cannot be described by simple linear or univariate relationships (Onderka et al., 2012). This has given rise to numerous studies on suspended-sediment concentration-discharge (SSC-Q) relationships, frequently observed as hysteretic in nature (Duvert et al., 2010; Gao & Josefson, 2012; Lefrançois et al., 2007; Seeger et al., 2004; Sherriff et al., 2016; Williams, 1989).

Williams (1989) appears to be the first to systematically describe single event SSC-Q hysteretic behavior (e.g., linear, clockwise, counter-clockwise, figure-eight loops, and a few variations) using shapes and timing of the hydrograph and corresponding sedigraph and offers examples of physical watershed processes that cause such patterns. Subsequent studies (1) identified these types of hysteresis in the SSC-Q relationships (Megnounif et al., 2013; Soler et al., 2008), (2) validated the physical processes that give rise to these patterns (Cheraghi et al., 2016; Duvert et al., 2010; Pietroń et al., 2015; Seeger et al., 2004), and then (3) inferred sediment dynamics occurring in the study watersheds (Gellis, 2013; Martin et al., 2014; Sherriff et al., 2016; Smith & Dragovich, 2009). Studies also demonstrated broad applicability of the hysteresis patterns beyond the SSC-Q relationship; including, but not limited to, the relationship between discharge and soil moisture (Zuecco et al., 2016), discharge to other solute concentrations (Ramos et al., 2015; Vaughan et al., 2017), and suspended sediment and turbidity (Landers & Sturm, 2013).

SSC-Q hysteresis has also been used to identify (1) the relative contribution of in-stream sediment sources and more distant hillslope sources to overall sediment yields (Aich et al., 2014; Buendia et al., 2016; Martin et al., 2014; Sherriff et al., 2016; Smith & Dragovich, 2009; Yeshaneh et al., 2014), (2) whether individual watersheds are supply limited or transport-limited (Asselman, 1999; Duvert et al., 2010; Gao & Josefson, 2012; Ramos et al., 2015), and (3) the predominant sediment source, such as bank erosion (Lefrançois et al., 2007; Smith & Dragovich, 2009). Studies on the temporal effects on the predominant watershed hysteresis types have inferred seasonal dynamics of sediment supply and transport (Buendia et al., 2016; Gao & Josefson, 2012; Lefrançois et al., 2007; Martin et al., 2014; Sherriff et al., 2016; Smith & Dragovich, 2009; Yeshaneh et al., 2014), and although less extensively, the effect of watershed size (Aich et al., 2014; Asselman, 1999; Smith & Dragovich, 2009), to understand how tributary sediment delivery differs from that of the downstream, main channel outlet.

Classification of hysteretic SSC-Q patterns is important when making inferences about sediment dynamics. To date, researchers have utilized either a qualitative visual classification or hysteresis indices (HI) (Lloyd et al., 2016a) to quantitatively measure differences in the rising and falling limbs of the SSC-Q relationship. Various metrics have been proposed to automate and objectively classify hysteretic behavior (Aich et al., 2014; Landers & Sturm, 2013; Langlois et al., 2005; Lawler et al., 2006; Lloyd et al., 2016a; Smith & Dragovich, 2009; Zuecco et al., 2016). Some indices facilitate compression of information on the shape and pattern of the SSC-Q relationship into a single metric that helps infer event sediment dynamics without the need for classification. However, HI values are not unique (i.e., individual storm events with different hysteresis patterns can have the same index value) and therefore, often require additional metrics such as loop area or direction to preserve information lost during data compression (Lloyd et al., 2016a; Zuecco et al., 2016).

Machine-learning methods can help identify patterns in hydrological data. For example, feed-forward backpropagation algorithms have long been used in rainfall-runoff modeling and streamflow prediction (Abrahart et al., 2012; Maier et al., 2010). More recently, a new variety of pattern recognition networks called deep belief neural networks (DBNNs) excel at classification applications such as handwritten character recognition (O'Connor et al., 2013), sparking extensive research into deep learning. One building block of the DBNN is the restricted Boltzmann machine (RBM), which acts as a feature extractor for pattern recognition and classification (Tieleman, 2008). The suite of RBM algorithms, now readily available in a variety of computer languages that run on a desktop PC, make them attractive in balancing state-of-the art performance with ease of implementation (Testolin et al., 2013).

This study leverages 3 years of high-resolution riverine suspended-sediment time series from multiple sites to show proof-of-concept of (1) expanding the existing visual classification system of storm-event

suspended-sediment hysteresis patterns, and (2) automating the classification of event hysteresis using a novel machine-learning technology designed for high-frequency environmental monitoring data. We then illustrate the utility of the technology for understanding the environmental drivers of suspended-sediment dynamics during storms and suspended-sediment provenance. The discussion includes implications and opportunities for watershed management communities, future applications, and modifications of this approach.

2. Methods

2.1. Study Area

The Mad River watershed, located in the Green Mountains of Vermont within the Lake Champlain Basin (Figure 1), was selected as a study site based on available long-term stream gauge records and ongoing geomorphic and sediment dynamics studies (Stryker et al., 2017; Wemple et al., 2017). Elevation ranges from 132 m to 1,245 m above sea level, with forests (83% of watershed area) dominating all but the valley floors, which are occupied by agriculture (8%) and village centers and other developed lands (8%) (Table 1). Soils range from fine sandy loams derived from glacial till deposits in the uplands to silty loams derived from glacial lacustrine deposits in the lowlands. Erosional watershed processes include bank erosion, agricultural runoff, unpaved road erosion, urban storm water, and hillslope erosion. The Mad River main stem has been subjected to channel management activities (e.g., straightening, dredging, and armoring) as recent as the mid-20th century contributing to present-day decreased access to flood plains and increased erosion hazards (Fitzgerald & Godfrey, 2008). Mean annual precipitation in the watershed ranges from 1,100 mm along the valley floor to ~1,500 mm along the upper watershed slopes (PRISM Climate Group, 2015).

2.2. Data Collection

We selected six study sites for monitoring, one along the main stem and one on each of five tributaries (Figure 1). The Mad River is a fifth-order stream and the monitored tributaries are all fourth-order except for

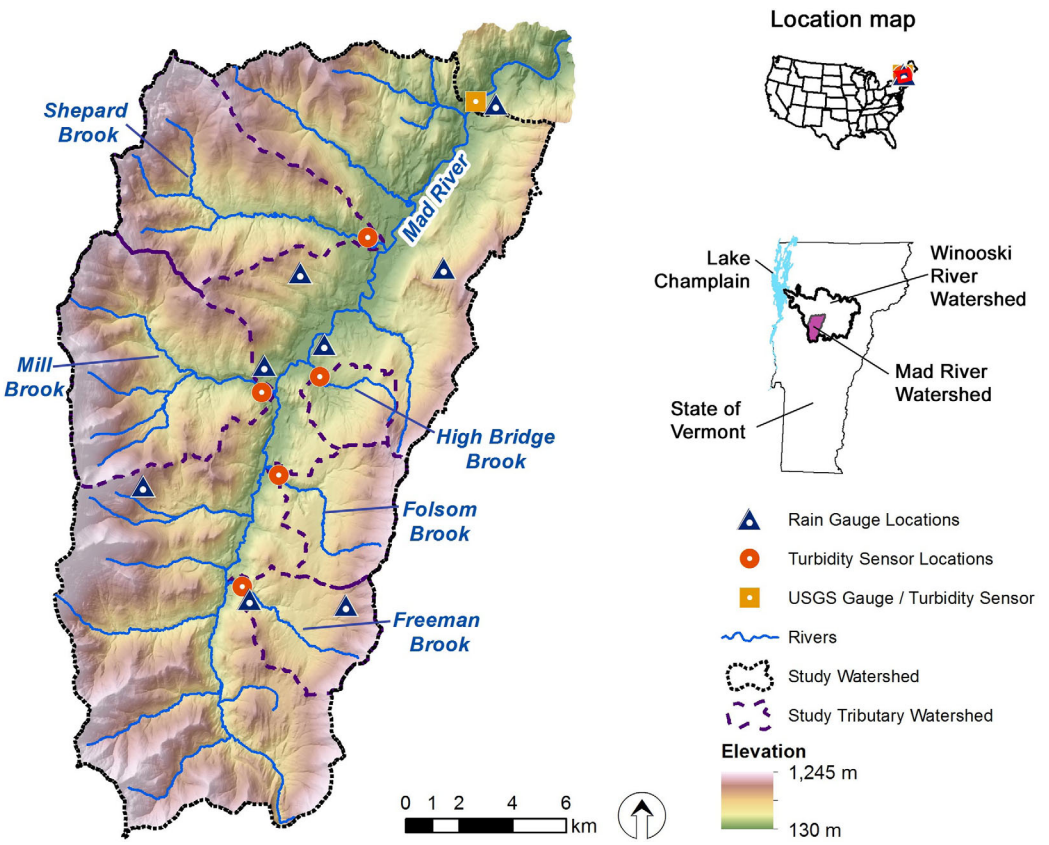

Figure 1. The Mad River watershed and monitoring site locations.

Table 1
Key Characteristics of Study Watersheds

	Shepard Brook	High Bridge Brook	Mill Brook	Folsom Brook	Freeman Brook	Mad River
Area (km^2)	44.6	8.6	49.8	18.2	17.0	344.0
Minimum elevation (m)	195	225	216	229	266	140
Maximum elevation (m)	1,117	796	1,114	886	860	1,245
Elevation range (m)	923	571	898	657	594	1,105
Stream order	Fourth	Third	Fourth	Fourth	Fourth	Fifth
Drainage density (km/km^2)	2.38	2.45	2.16	1.77	1.95	0.97
% Forested land	92.2	66.7	89.2	77.6	76.2	85.5
% Developed land	1.0	16.6	1.5	12.7	8.3	4.7
% Agricultural land	5.6	15.5	7.0	8.8	14.6	8.0
% Other land	1.1	2.1	0.8	0.7	1.7	1.1
Road density (km/km^2)	0.83	2.30	1.19	1.07	1.80	1.32

High Bridge Brook, which is a third-order stream. In situ digital turbidity sensors (Forest Technology Systems model DTS-12) and automated samplers (Teledyne Isco 6712) with stage sensors collected river level and water-quality data for analyzing event sediment dynamics. Both turbidity and discharge data were collected at 15 min intervals. Suspended-sediment concentration (SSC) samples were collected over storm events for laboratory analysis. Samples were processed using the standard gravimetric method (EPA Method 160.2) for measuring total suspended solids; but we discuss using the term SSC for generalizability. To estimate a continuous (15 min interval) SSC time series, relationships between turbidity and SSC were developed for each monitoring station using a rating curve (supporting information Figure S2). At all sites, turbidity was highly correlated with SSC (supporting information Table S1). Tipping bucket rain gauges (Onset HOBO) collected precipitation data at seven locations within the watershed (Figure 1). A meteorological station, located in the central portion of the watershed, provided soil moisture measured as volumetric water content of the soil at multiple depths.

Stage-discharge relations were obtained from the existing USGS (Mad River—Station ID 04288000) rating curve or developed from discharge measured on three of the tributaries (Mill, Folsom, and Shepard Brook). At the High Bridge Brook and Freeman Brook monitoring sites, discharge was estimated by measuring stage and using an approximated stage-discharge rating curve based on the Mad River gauge discharge, scaled on watershed area. Study sites were instrumented in the spring after ice breakup and maintained until the start of ice formation in December from 2013 to 2015. Freeman Brook and High Bridge Brook had unstable channels making stage-discharge relationships impractical to develop; thus, they were monitored only during the 2013 data collection period.

The identification of onset and end of individual storm events was semiautomated. The onset was set as the first positive rate of change between consecutive 15 min Q measurements; while termination was based on manual identification of an end point using the graphical sloped line approximation (Nathan & McMahon, 1990). When multiple, proximate, discharge peaks could be attributed to rainfall periods separated by 12 h, they were divided into separate events. Rainfall for each event was assumed to be equivalent to the nearest rain gauge for each of the five subwatersheds and calculated using a Thiessen polygon weighted average of all rain gauges for the main stem.

A comprehensive range of hydrological conditions characterized the monitoring period. May–June 2013 was the wettest consecutive 2 month period on record and culminated in a large flooding event on 3 July 2013; whereas, late summer 2015 featured very low flows and drought conditions. For both the Mad River and all monitored tributaries, bankfull flow events occurred during the monitoring period. Comparing flows during our 3 year monitoring to a flow duration curve generated from the Mad River USGS stream record shows the monitoring period adequately encompassed a variety of flow conditions (Figure 2a). Across all monitoring sites, 145 storm events were captured resulting in a total of 600 unique events (Figure 2b). The highest number of events (35) were recorded in July and the fewest (7) in December. It should be noted that stations were not always online in May or December of a given year because of sensor deployment limitations (supporting information Figure S1). The number of events monitored at each site (Figure 2c) varied

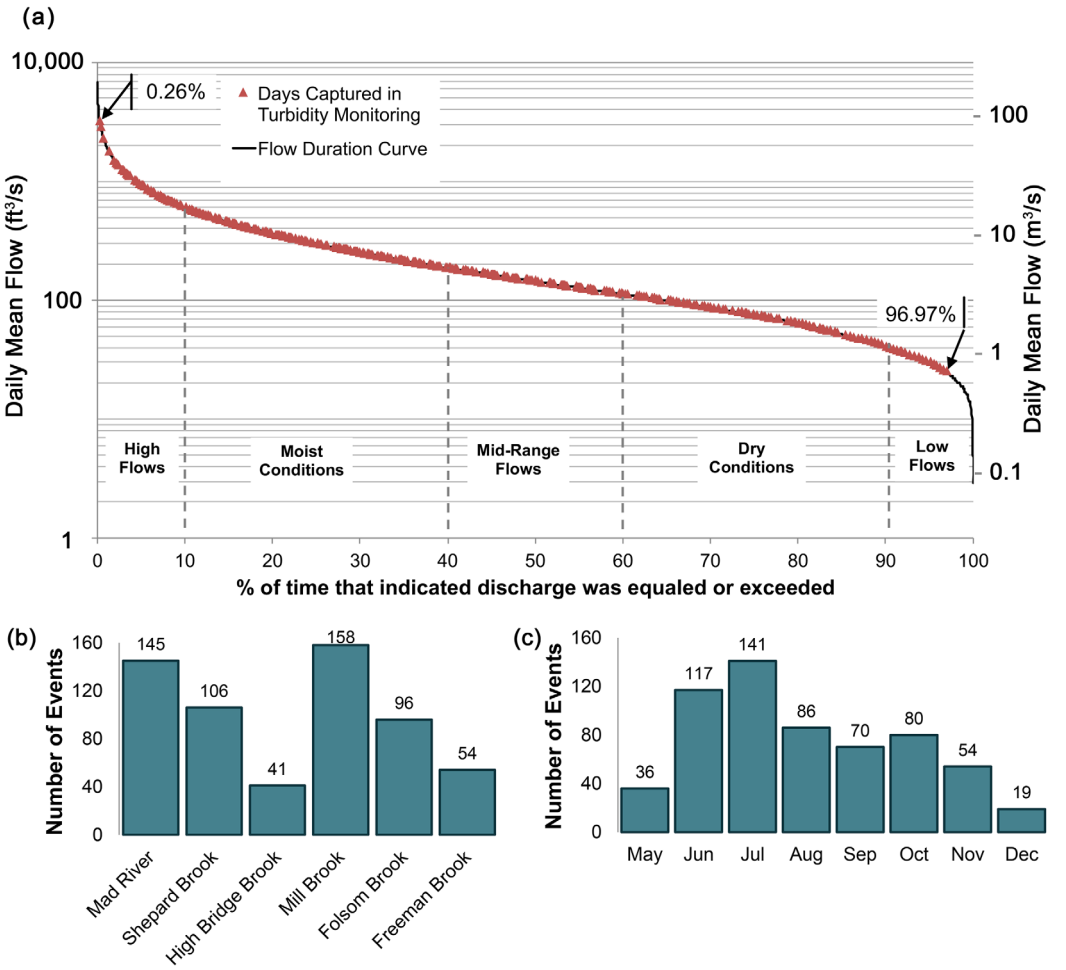

Figure 2. (a) Flow duration curve for the Mad River monitoring site showing days when suspended sediment monitoring occurred and number of storm events captured by (b) site and (c) month across all sites combined.

depending on whether sensors were online (supporting information Figure S1) and because rainfall events do not occur uniformly across all subwatersheds.

2.3. SSC-Q Plot Image Processing

For each storm event, SSC-Q plots were generated from the time series. Because the visual hysteresis patterns may be affected by noise in the SSC or Q data, the event time series for SSC and Q were processed using a Savitsky-Golay smoothing filter (Savitzky & Golay, 1964) prior to generating the hysteresis plots (Figure 3). A third-order, 21 step filter for the Mad River site and a third-order, 13 step filter for the tributary sites was selected to reduce noise, yet still preserve the peaks and shape; filter selection was based on visual inspection of event time series. Because smoothing was performed solely to simplify visual classification, it was applied after calculating the storm metrics. The SSC-Q hysteretic loop was shaded to preserve time (i.e., loop direction) (Figure 3b). Images used an 8 bit grayscale color palette, where white indicates the onset of the storm, and dark gray the end of the event, plotted on a black background (Figure 3e). To reduce computational time during classification, the SSC-Q plots were converted to 28 by 28 pixel resolution (Figure 3f); both SSC and Q were normalized on a per event basis. SSC-Q plot images were used to train and test an automated classifier using a supervised machine-learning approach.

2.4. Restricted Boltzmann Machine

The restricted Boltzmann machine (RBM) (Smolensky, 1986) is a type of artificial neural network used for unsupervised learning that has experienced an increase in popularity for its extended use as a classification model (Larochelle et al., 2012) and more recent use in deep learning applications (Hinton et al., 2012). The

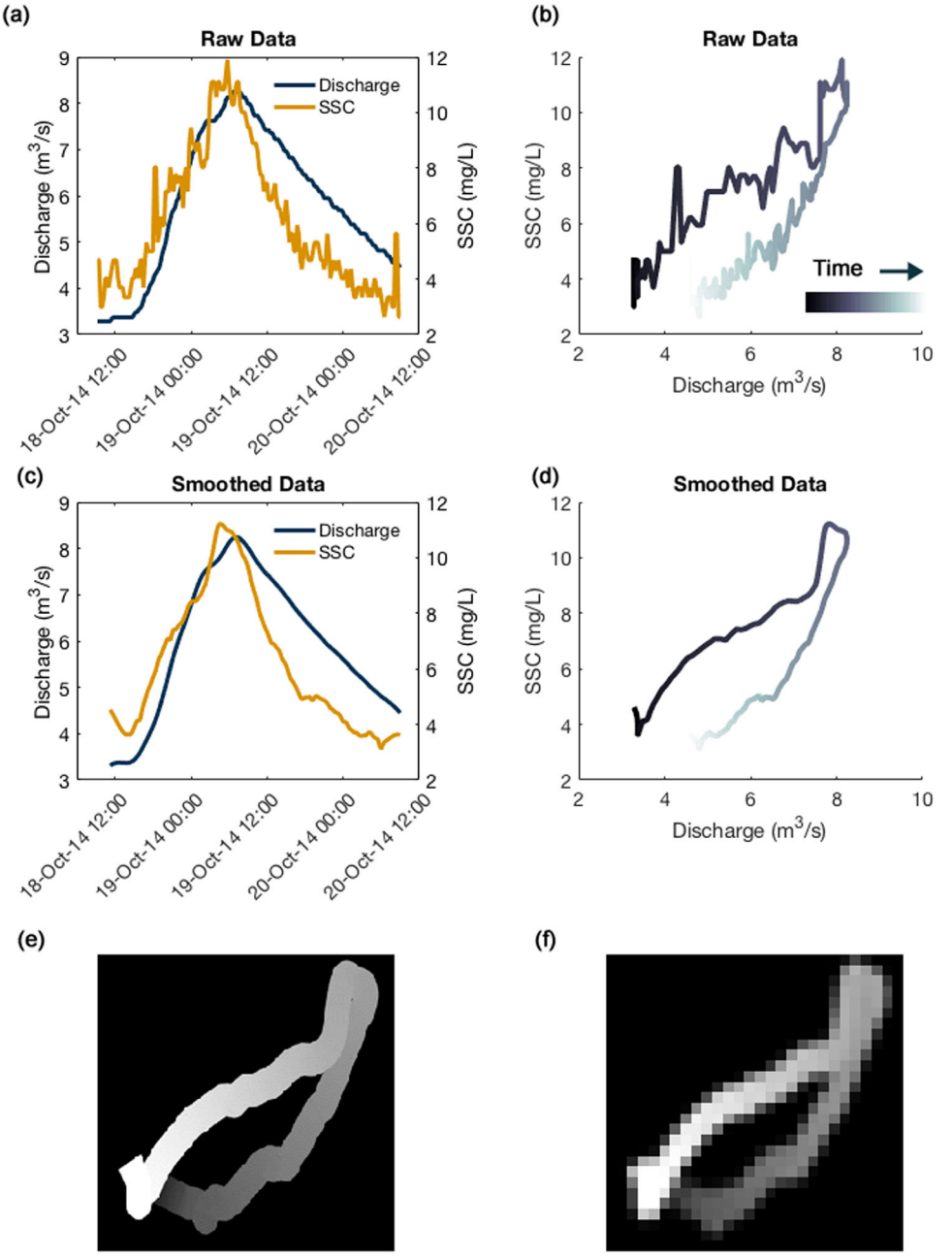

Figure 3. Example of smoothing and image processing applied to discharge and suspended-sediment data from 18 October 2014 storm event at the Mad River. (a and b) The raw 15 min data; (c and d) data after both time series are smoothed. Image of SSC-Q plot used as input to classification tool at (e) high and (f) reduced resolution of 28 × 28 pixels. Shading represents time starting at white and later time in gray.

RBM is a stochastic, generative model consisting of two layers (Figure 4a): a layer of visible nodes (input data) and a layer of hidden nodes (feature detectors).

In a basic RBM configuration, an energy function determines the probability of all configurations of the visible and hidden nodes and is defined as:

$$E(\boldsymbol{v},\boldsymbol{h}) = -\sum_i \sum_j w_{ij} v_i h_j - \sum_i v_i b_i^{(v)} - \sum_j h_j b_j^{(h)},$$

where v_i, h_j are the binary states of visible unit i and hidden unit j, respectively, $b^{(v)}$ and $b^{(h)}$ are their biases, and w_{ij} are the weights connecting these unit layers. The joint probability encoded by the RBM can then be expressed as:

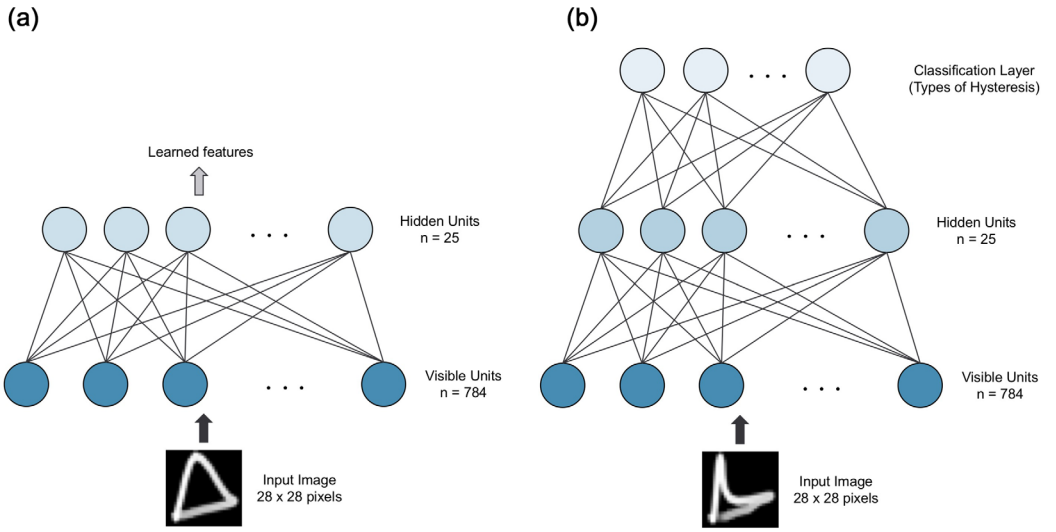

Figure 4. (a) Network architecture of restricted Boltzmann machine (RBM) and (b) RBM classifier neural network.

$$p(\boldsymbol{v},\boldsymbol{h}) = \frac{e^{-E(\boldsymbol{v},\boldsymbol{h})}}{\sum_{\boldsymbol{v},\boldsymbol{h}} e^{-E(\boldsymbol{v},\boldsymbol{h})}}.$$

Given the equations above, the states of the hidden and visible units may be stochastically updated using:

$$p(v_i=1|\boldsymbol{h}) = \sigma\left(b_i^{(v)} + \sum_j w_{ij}h_j\right)$$

$$p(h_j=1|\boldsymbol{v}) = \sigma\left(b_j^{(h)} + \sum_i w_{ij}v_i\right),$$

where $\sigma(x)$ is the sigmoid function $1/(1+e^{-x})$; othersize the units are set to zero. The basic configuration of the RBM uses binary stochastic units for both the visible and hidden layers; however, the visible layer units can alternatively use continuous real values (i.e., normalized grayscale pixel values) with no changes to the weight update rules.

The goal of training a RBM is to adjust the parameters (weights) of the network such that the probability distribution evolved by the RBM becomes similar to the training data (i.e., maximizing the log-likelihood of generating the observed data). To perform a stochastic gradient ascent on the log probability of the data, the following simple learning rule is used:

$$\Delta w_{ij} = \epsilon\left(\langle v_i h_j\rangle_{data} - \langle v_i h_j\rangle_{model}\right),$$

where ϵ is the learning rate, $\langle\cdot\rangle_{data}$ indicates an average over samples with visible units clamped to match actual input data, and $\langle\cdot\rangle_{model}$ indicates the average over samples when the network is allowed to sample freely from the model's distribution. As a result, there are two phases in the RBM learning algorithm: (1) the *positive phase* where a training pattern is clamped to the visible units and hidden unit activations are adjusted and (2) the *negative phase* where the algorithm computes or models the reconstruction of the training data given the hidden unit states. Computation of the positive phase is fairly straightforward; but the negative phase requires inference methods using Markov Chain Monte Carlo sampling methods. In order to avoid the requirement of performing Gibbs sampling over long periods, Hinton et al. (2006) demonstrated a much faster method for training an RBM with contrastive divergence. Different sampling methods have been proposed for the contrastive divergence approach, with block Gibbs sampling over a fixed number of iterations (e.g., Tieleman, 2008) being one of the most common.

To improve the learning speed and convergence, the data are typically split into mini-batches; and the gradient in each learning step is averaged over the patterns of the mini-batch. Readers are referred to Fischer and Igel (2014) and Hinton (2012) for in-depth discussion of RBM training methods.

While RBMs are useful in themselves for feature extraction or image generation after initial training; in many applications, they are extended to a classifier model. When extended to a classifier network algorithm, the trained hidden features are used as input to an additional classifier layer (Figure 4b). In this approach, the trained RBM is then connected to a subsequent classifier layer; and the entire network is "fine-tuned" by the backpropagation learning method (e.g., Sarikaya et al., 2014). RBMs can be used for classification by other methods including training a separate RBM for each class (Hinton, 2012) or by incorporating a "class label" node in the visual input layer (Larochelle et al., 2012). We selected the former approach as it has been demonstrated to be both practical and is similar to the approach implemented in DBNN applications (Testolin et al., 2013).

2.5. Automated SSC Hysteresis Classification of Storm Events

Given the 600^{+} SSC-Q images, a set of distinct, yet repeating, hysteresis pattern types were first selected by a single domain expert. This manual clustering of the SSC-Q images was initially guided by a nonparametric clustering method (Hamshaw et al., 2016). We subsequently recruited five domain experts and repeated the manual classification on a random subset of 100 hysteresis patterns to help evaluate precision across the small subset of human experts and have a baseline measure of classification accuracy for human experts. SSC-Q plots were then manually labeled with the best-matching type of hysteresis pattern. To automate the storm-event classification, the RBM classifier (Figure 4b) was trained on a portion of the data and tested on the remaining data. Three examples of each hysteretic type were selected for the training data to ensure each type was represented; the training data were then supplemented with a random selection of events until one-half of the events (294) were included for training. The remaining 306 events were withheld for testing. Events that could not be categorized into one of the hysteresis pattern types were labeled "complex" and excluded from the training data because prediction of the complex (i.e., no pattern) type was not desired.

The RBM classifier performance was compared to the manually assigned hysteresis type of each storm event. Because some hysteresis patterns may be a transition between two types, we also evaluated the network classification accuracy by defining "correct" classification as an exact match with the domain expert or with the most similar state in the transition. This provided a second measure of the RBM classifier performance where "near misses" are assessed as valid classifications. For each hysteresis type the two similar patterns (transition states) were defined by domain experts (supporting information Figure S4). The number of hidden nodes, mini-batches, and amount of training data all influence the prediction results; therefore, we varied the number of each as part of evaluating the RBM performance. The results were averaged over 25 model runs, which comprised five network simulations each using five different realizations of training data.

2.6. Analysis of Event Variables

We analyzed a suite of additional hydrological and meteorological variables (Figure 8) on a per event basis in conjunction with the raw sedigraph and hydrograph data to determine whether particular conditions are more conducive to producing a given hysteresis type. To determine if differences in the typical storm event were more likely associated with a particular hysteresis type, we compared the variable mean between each hysteresis type and all other types, repeating for each variable. Means were compared using between group *t* tests on normally distributed variables, and Wilcoxon rank tests when normality could not be assumed. In addition, a HI was calculated following procedures in Lloyd et al. (2016a).

Two additional event variables were calculated for each event at the Mad River site. The coefficient of variation (CV) of the total event rainfall from all rain gauges within the watershed was computed to assess the spatial variability of rainfall. Second, the total stormflows from all events were fit to a lognormal distribution, repeated for each site. This distribution was then used to estimate the quantile of stormflow for each event providing a measure of the size of hydrological response associated with each storm event.

3. Results

3.1. Hysteresis in SSC-Q Event Relationships

Events fell into 14 SSC-Q hysteresis types that could be grouped into five main categories (Classes I–V) corresponding to those originally proposed by Williams (1989); however, Classes II, III, and V are now further subdivided into newly proposed types based on patterns observed repeatedly at the study sites (Figure 5).

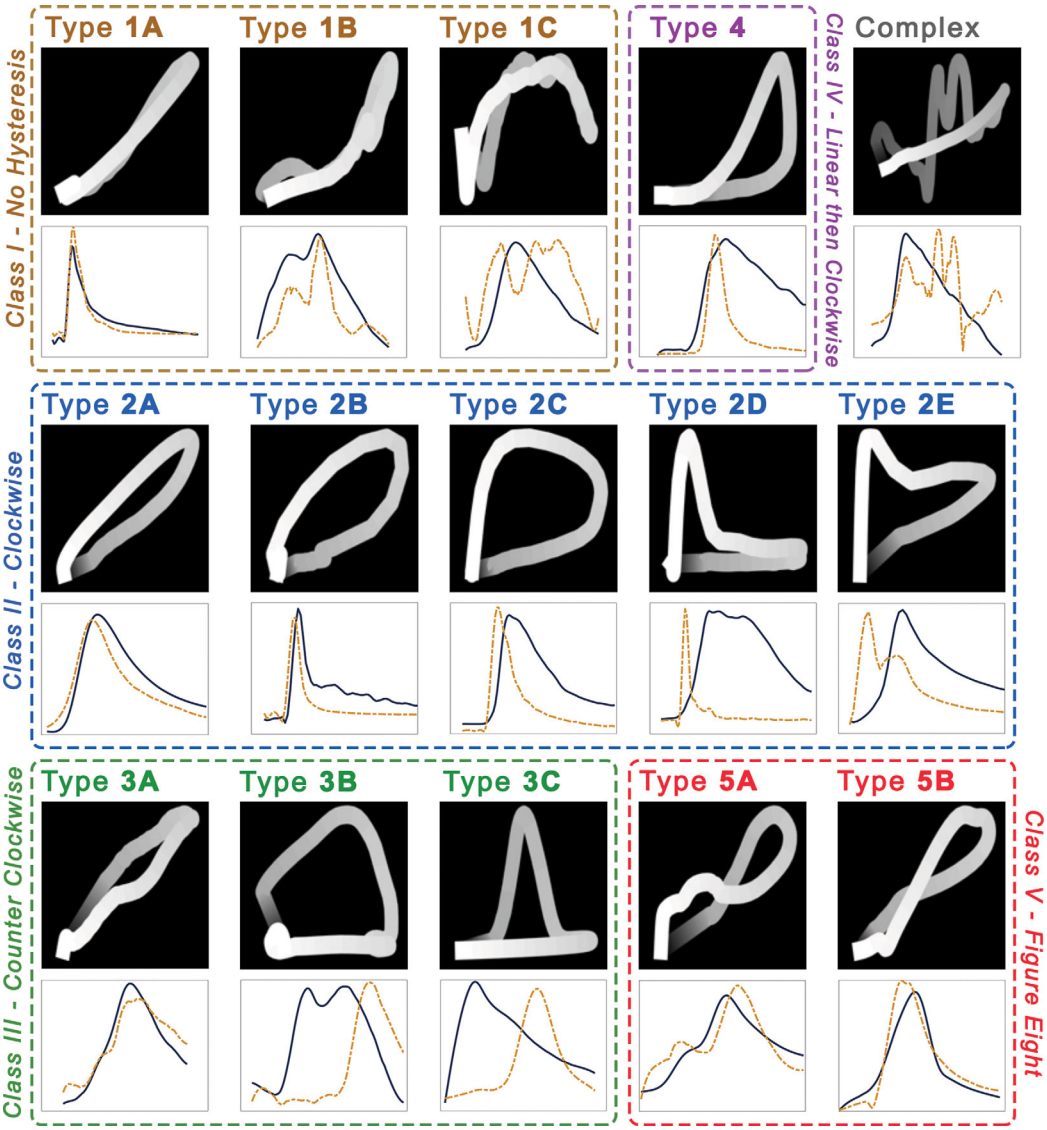

Figure 5. Classes of hysteresis in SSC-Q plots from events observed in the Mad River watershed. Solid line indicates hydrograph and dashed line indicates sedigraph. (*Note*. scales are not consistent between storms.)

Class I, consists of Types 1A, 1B, and 1C and represents variations on the SSC-Q relationships that do not exhibit any hysteretic behavior. Class II behavior (clockwise loops) consists of types 2A, 2B, 2C, 2D, and 2E with differences in the timing of the peak SSC and peak Q influencing the shape of the hysteresis. Type 2A is indicative of the peak SSC occurring just prior to the peak Q; therefore, resulting in a SSC-Q plot with minor amount of hysteresis. In contrast, when the SSC peak occurs well before peak Q (i.e., Type 2D), it results in an "L" shaped loop. Type 2E is a variation where the peak SSC occurs well before peak Q but has a secondary peak SSC occurring near the peak Q. The Class III SSC-Q relationships (counter-clockwise loops) were similarly subdivided into Types 3A, 3B, and 3C reflecting separation differences in the timing between the SSC and Q peaks. An SSC-Q plot with a linear relationship followed by clockwise loop is indicative of Class IV behavior. We note that although Type 4 hysteresis patterns could reasonably be considered a subset of Class II (clockwise) hysteresis types, we labeled them as Class IV in this work to be consistent with the five classes originally proposed by Williams (1989). The figure-eight shaped SSC-Q loops are represented as Class V with subcategories Type 5A and Type 5B discriminating between the loop direction.

The majority (90.0%) of the monitored SSC-Q event relationships displayed hysteretic behavior. Across all study sites, the most commonly observed hysteresis types (63.8%, or 388 of 600 events) were Class II

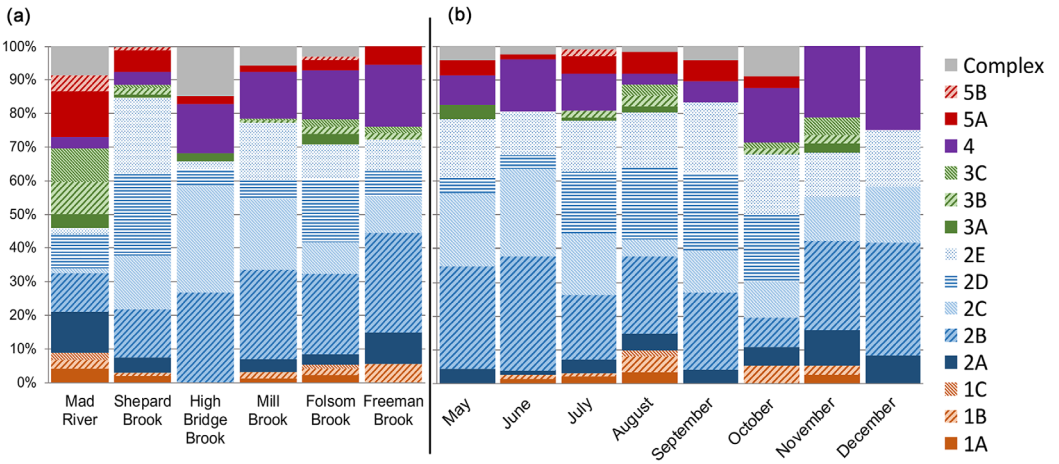

Figure 6. (a) Distribution of hysteresis types by site. (b) Distribution of hysteresis types by month for only fourth-order tributaries (Shepard, Mill, Folsom, and Freeman Brook) combined.

(clockwise) patterns, indicative of a peak SSC occurring prior to the peak discharge. Class III (counter-clockwise) types occurred less frequently (8.5%). We were unable to identify a clear hysteresis type in 5.1% of the events and labeled those "complex" type hysteresis.

The frequency of hysteresis types varied among sites (Figure 6a and supporting information Table S2). The Mad River (main stem) site featured the most varied hysteresis types including more frequent occurrence of counter-clockwise and figure-eight patterns compared to the tributaries. The tributaries predominantly featured Class II (clockwise) patterns (73%); however, the distribution of Class II subcategories varied among sites. The most common type observed at Freeman, Folsom, and Mill Brook was Type 2B; whereas Types 2C and 2D were the most common pattern at High Bridge Brook and Shepard Brook, respectively. There was variability in median HI by hysteresis type (Figure 7). Types 1A, 1B, and 1C had median values of HI near zero as well as the figure-eight patterns (Types 5A and 5B). As expected, Class II (clockwise) types had positive median HI and Class II (counter-clockwise) had negative. However, with the exception of two types (2C and 3B) median values of HI between types were not statistically different from other types (Figure 7).

3.2. Automated Classification of Hysteresis

An RBM configuration with 25 hidden nodes and a mini-batch size of 14 provided sufficiently good prediction. Tests using smaller numbers of hidden nodes and mini-batch size degraded performance; while more hidden nodes resulted in negligible improvement and overfitting. This was consistent with the RBM training guidelines of Hinton (2012), which recommend a mini-batch size equal to the number of classes.

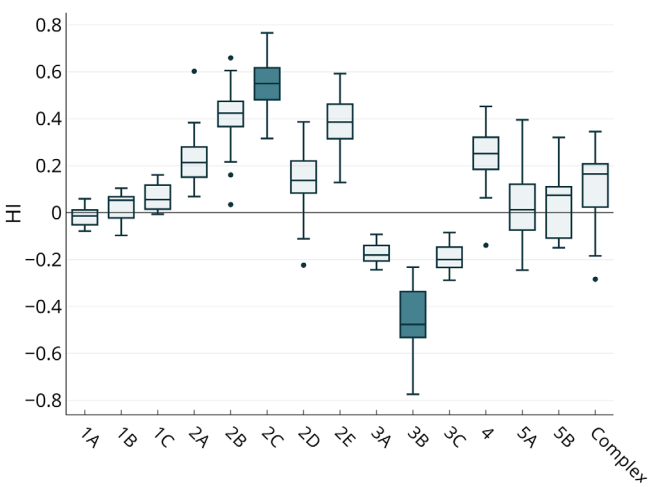

Figure 7. Box plots of Hysteresis Index plotted by hysteresis type. Shaded box plots indicate statistical difference from other types.

The accuracy (total percentage of exact classification match with the domain expert) was 77% for the training data and 53% for the testing data. Alternatively, the classification accuracy using the exact match or one of the transition states is 85% for the training data and 70% for the testing data (supporting information Figures S3 and S4). For some individual event classifications, discriminating the hysteresis direction was difficult (e.g., Type 3B classified as clockwise instead of counter-clockwise as seen in supporting information Figure S5f). Generally, only 8% of events had misclassified direction (i.e., Type 2A versus Type 3A as seen in supporting information Figure S4). The manual classification test by domain experts resulted in an average of 67% of events classified into the correct class, or 79% into the correct or similar class.

3.3. Storm-Event Characteristics of Hysteresis Types

Using only a single watershed scale (the fourth-order tributaries) allowed the analysis between typical event characteristics and

Figure 8. Comparison of hysteresis types for all fourth-order tributaries (Shepard, Mill, Freeman, and Folsom brook) against the mean of all storm-event variables (variable names are shown in center of top plot and described in lower portion of the figure). Significant differences above or below the mean storm event were tested using a *t* test for normally distributed variables and the Wilcoxon test for nonnormally distributed variables. We provide an example interpretation of this figure as follows: storm events exhibiting Type 5A hysteresis (in red) have, on average, slightly lower (lower portion of top plot) 14 day antecedent rainfall (A14P), flood intensity (FI), and total sediment load (SSL_{NORM}); these patterns also have moderately lower 3 day antecedent rainfall (A3P), shallow soil moisture levels ($SM_{SHALLOW}$), prestorm base flow (BF_{NORM}), and total rainfall (P); and much lower event stormflow (Q_{NORM}) and stormflow event quantile ($log[Q_{NORM}]$) compared to all other events.

hysteresis types to be performed at a consistent spatial scale (Figure 8). Patterns exhibiting no hysteresis (Types 1A and 1B) or minor hysteresis (Type 2A) occurred relatively infrequently (36 out of 414 events) across the four sites. These events were characterized as having low, 14 day antecedent rainfall, low deep soil moisture, and a lower and shorter event rainfall duration. In addition, despite having little separation between the peaks of the sedigraph and hydrograph, the time between the sedigraph peak and rainfall pulse was longer on average than other events.

Type 2B hysteresis (moderate clockwise loop) was the most commonly observed (96 events). Compared to all events, larger storms on average were classified as Type 2B, featuring greater 3 and 14 day antecedent rainfall, higher prestorm base flows, greater rainfall total, and maximum rainfall rate. Hydrological metrics also trended higher with large flood intensity, stormflow, sediment load, and peak sediment concentration. Similarities in Types 2C and 2B events were observed in terms of larger flow metrics and antecedent rainfall. However, Type 2C (pronounced clockwise loops) events differed in that they have higher antecedent soil moisture conditions, shorter time span between the rain and sedigraph peaks, and less intense and shorter rainfall than Type 2B. Type 2C events would therefore be characterized by very wet antecedent conditions in which a nonintense rainfall results in a rapid sediment response.

The Type 2D events had very different characteristics than Types 2B or 2C events. Compared to Types 2B and 2C events, Type 2D events were characterized by drier antecedent conditions with on average lower base flow, less 3 and 14 day antecedent rainfall, and lower soil moisture. In addition, Type 2D events tended to feature lower and shorter rainfall and correspondingly less stormflow and less sediment yield. These events were also characterized as having the shortest time span between the rainfall peak and sedigraph peak. Type 2E clockwise events are characterized by both a rapid sediment response and a delayed, second sediment pulse closer to the hydrograph peak. It therefore may be considered a combination of Type 2D and Type 2A or 2B patterns. Perhaps owing to its "mixing" of two event types, Type 2E events are not

distinguished statistically from the overall data for the majority of variables. Of note, Type 2E events feature little separation in the timing of the rainfall peak and sedigraph peak, which is consistent with the very early sediment pulse, and the tendency for low sediment concentrations and low flow normalized sediment flux.

Although Classes III, IV, and V patterns occurred less frequently (83 out of 414 events) in the four tributaries, some patterns emerged between the event metrics and hysteresis types classes. The Types 3A and 3B counter-clockwise patterns were observed in events that featured on average less antecedent rainfall and lower rainfall amount and duration. The flows were also characterized as being smaller in magnitude with longer durations and time to peak discharge. Interestingly, they also tended to have higher flow normalized sediment fluxes, which is likely a result of the sediment peak occurring during a portion of the hydrograph with higher flows compared to Type 2D. Type 4 hysteresis, which features a more linear SSC-Q relationship followed by a clockwise loop, occurred most frequently during events with a few unique characteristics. These events tended to have low rainfall rates that were long in duration with corresponding low flood intensity and longer times to peak discharge. These storms also occurred most frequently in late spring and late autumn when more typical climate patterns comprise slower moving, less intense frontal storms in spring and autumn, as opposed to shorter, more intense convective storms typical of the summer. Events displaying Type 5A hysteresis were on average small events that occurred with dry antecedent conditions.

4. Discussion

4.1. Expanded Classification of Hysteresis Types and Comparison to HI

Expansion of the hysteresis classes helped identify differences among our sites that would not be possible using an analysis based solely on the five original classes proposed by Williams (1989). For example, three tributaries (Shepard, Folsom, and Freeman Brook) were essentially dominated by the same percentage (65–67%) clockwise (Class II) hysteresis patterns across the three sites. However, classification using this one category obscures the fact that in Shepard Brook the most common Type 2D hysteresis differs from that observed in Folsom and Freeman Brook where Type 2B is the most frequent. In addition, at the High Bridge Brook tributary site a different type of clockwise pattern was most frequent (Type 2C). Thus, our expanded classification helps further distinguish in-storm sediment export regimes that could provide more insight into dominant drivers of suspended-sediment export during storm events across different catchments.

While other studies have suggested the need for an additional new class, we believe this study is the first to systematically expand the types of hysteresis observed in the SSC-Q relationship. Asselman (1999) differentiated between two types of clockwise loops (*typical* and *pronounced*), which would be most similar to our Types 2B and 2C, respectively. Similarly, a weak clockwise loop was identified by Gao and Josefson (2012), which corresponds to Type 2A in our classification. Previous studies have also identified a double clockwise loop (Megnounif et al., 2013; Smith & Dragovich, 2009). As noted by Zuecco et al. (2016) hydrographs that feature multiple peaks can create complex patterns such as the double clockwise loop; but generally, they can be avoided by separating the peaks into individual events. The limitation posed by the singular clockwise loop classification was also identified by Duvert et al. (2010), where flash flood type events produced hysteresis patterns that could not be easily classified. Our expanded classifications comprise patterns commonly observed in our study; they also comprise types considered or proposed in the above studies.

Hysteresis indices have been adopted widely as a way to expand the analysis of hysteresis loop in terms of direction and prominence. The nonuniqueness of HI across types as identified by Lloyd et al. (2016a) was evident in our data set. For example, Types 2A and 2D have similar HI values (Figures 9b and 9d) as do Types 3A and 3C. However, both pairs of classes differ significantly in the distance between peak SSC and Q. Various enhancements to hysteresis indices have been proposed to overcome the issue of nonuniqueness. Lloyd et al. (2016a) proposed also computing a loop area; and Zuecco et al. (2016) proposed using the minimum and maximum differences between the rising and falling limbs of the hysteresis plot. However, all the hysteresis indices, by design, lose information about the shape of the loop during the compression to an index value. Using the entire SSC-Q hysteretic image as proposed here avoids such loss of potentially valuable information, as illustrated by the following discussion.

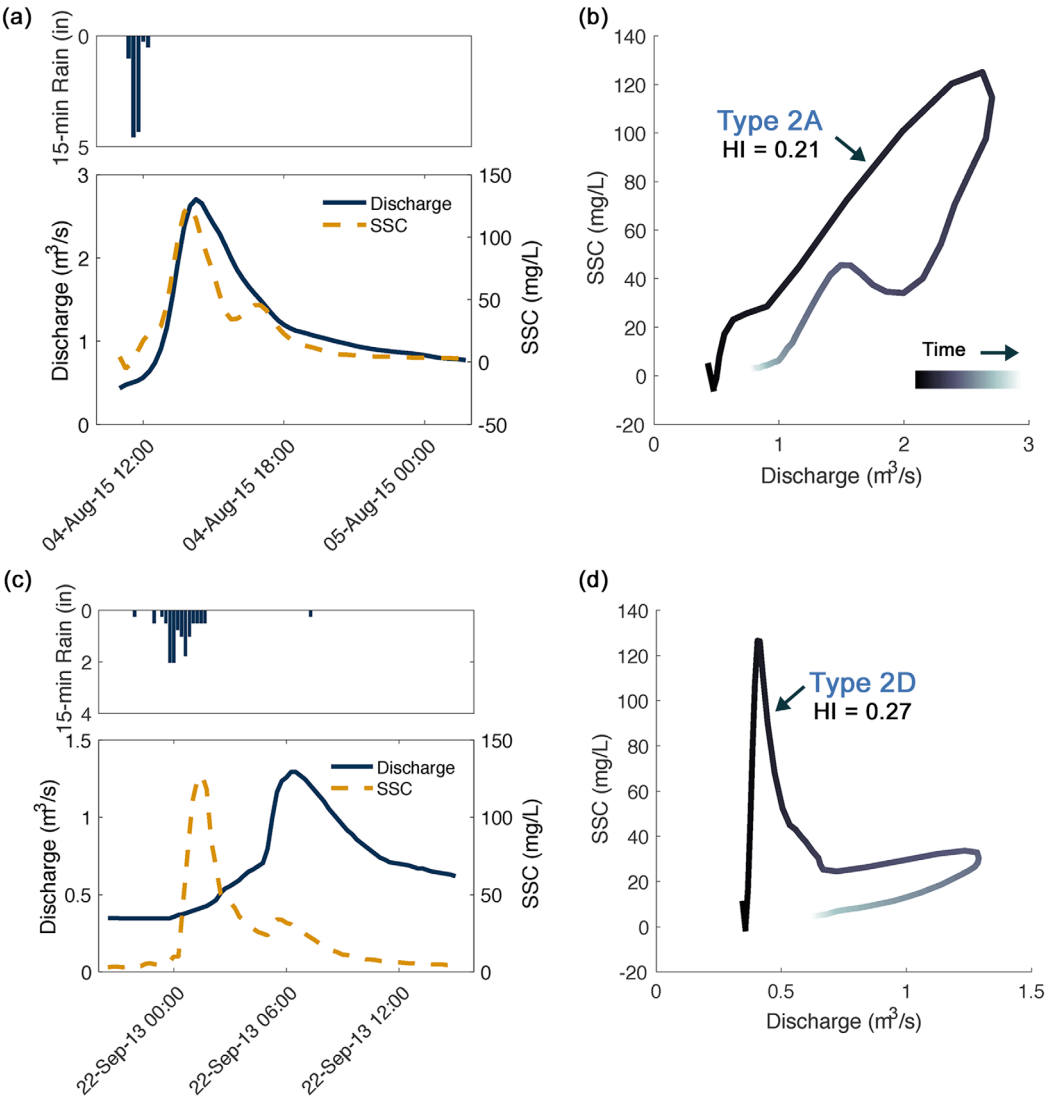

Figure 9. Examples of two storm events from Shepard Brook exhibiting different hysteretic behavior. (a) 4 August 2015 event features a peak SSC aligned closely with peak Q and (c) 22 September 2013 event features a peak SSC aligned closely with peak rainfall. (b and d) Resulting SSC-Q plots for the two events illustrating two different hysteresis types with similar HI measures clockwise loop direction.

4.2. Automated Classification of SSC-Q Hysteresis Patterns

Given the limitations of only having 600^{+} storm events to date and the simple RBM model design, we believe the performance of the automated classification to be satisfactory. We did observe a lower accuracy in testing than training (77% and 53%, respectively). In part, this likely is attributed to having a small, unbalanced data set and relatively large number (14) of classes where small variation in patterns within the testing data are not well represented in the training data. Efforts to augment the training data set with either synthetic data or perturbed observation data may improve the model performance. In comparison to the manual classification by domain experts, the RBM classifier network was ~14% less accurate. However, we attribute a portion of this to the network's ability to reliably recognize the correct hysteresis shape, but sometimes misclassify the direction indicating the learning algorithm places biases the image shape over its shading (hysteresis direction). Approximately 8% of the events had misclassified direction. Therefore, if this bias alone were remedied, the current classification accuracy (53%) would increase to 61%, which is more comparable to the manual classifications by domain experts (67%).

The classification accuracy is also linked to the relatively high number of hysteresis classes (14) selected for our analysis. If we had aggregated events into the traditional five classes of Williams (1989) (Figure 5), the

classification accuracy would have increased from 53% to 75%. Similarly, when using the alternative classification metric (i.e., exact match or one of the transition states), the model accuracy was increased to 70%. This implies we may have too many classes. However, the purpose of this research was to provide a framework for automating the pattern recognition and identify the physical drivers of hysteresis patterns; thus, retaining the additional class types was desired. We note that selecting the number of classes could be different depending on the application (e.g., hysteresis associated with hydrological variables or solutes). Finally, we highlight that SSC-Q hysteresis data are "noisy"; and even the classification accuracy of our domain experts' results in nearly 25% not being easily recognized/classified into one class, which is similar to previous studies (Sherriff et al., 2016; Yeshaneh et al., 2014). Therefore, the automated classification performance in this study can be considered acceptable; and more advanced machine-learning architectures (i.e., deep learning neural networks) are worthy of exploring, particularly when combined with additional training data from other monitoring sites or collected using much larger sensor networks.

4.3. Drivers of Hysteresis Type

We were able to interpret different drivers of hysteresis type (Figure 8). Antecedent watershed conditions related to soil moisture levels at the start of events, and prior rainfall amounts had a clear impact on event SSC-Q hysteresis type classification. Previously, wet antecedent conditions in the watershed have been connected to clockwise patterns (Buendia et al., 2016), and in general, our findings were consistent. However, our results indicate a more nuanced analysis is warranted because different types of clockwise patterns were observed to have different event drivers including antecedent conditions and storm-event characteristics. For example, two similar magnitude storm events in Shepard Brook result in very different hysteresis patterns resulting from the lag between sedigraph and hydrograph peaks (Figures 9a and 9b). These hysteresis patterns have similar directions and HI values (Figures 9c and 9d), and therefore would be classified the same using these metrics in contrast to our expanded classification.

Moderate clockwise hysteresis (i.e., Type 2B) events were on average larger storm events that occurred after significant rainfall and with higher prestorm base flows. Because these events featured higher than average sediment load and peak concentrations, this suggests that large rainfall events occurring on relatively wet antecedent watershed conditions cause widespread connection and mobilization of sediment sources; but with limited supply. This is in line with previous studies that found clockwise patterns were indicative of sediment supply in areas nearby the channel (Buendia et al., 2016), and ample sediment supply from channel and gully erosion (Smith & Dragovich, 2009). Bank erosion and gully erosion are suspected to be sediment sources in all four tributary watersheds. Conversely, smaller rainfall events occurring with wet antecedent conditions in the watershed were likely to exhibit stronger hysteresis (Type 2C) than observed with larger events. Therefore, sediment sources are still quickly mobilized, but are more limited in supply, suggesting less lateral connectivity with sediment sources. Thus, in these systems, small storms with high antecedent landscape moisture can quickly mobilize sediment from available nearby sources, yet have limited supply given the rapid decline in sediment concentration around the hydrograph peak.

Events with dry antecedent conditions likely have limited availability or connection to sediment sources, thereby limiting loading. Interestingly, a varied set of hysteresis patterns (primarily Types 2D, 3A, 3B, and 5A) were associated with various measures of dry antecedent conditions. Type 2D events occurred most frequently; and the very short time from rainfall to peak SSC (Figure 9c) suggests the sediment supply is limited to nearby, highly connected sediment sources. Thus, once rainfall ceases, the sediment supply decreases very rapidly; and typically, dry antecedent conditions suggest that Type 2D and similar patterns would not have significant overland flow or connection to remote sediment sources. This conclusion is consistent with Asselman's (1999) observation that pronounced clockwise hysteresis was attributable to erosion and sediment supply just upstream of the measurement location.

Counter-clockwise (Class III) patterns are typically attributed to sediment sources being more distant from the channel (Gellis, 2013) and may be attributed to storm events where erosion is primarily occurring in the far upstream or upslope portions of the watershed. These events feature long time delay between rainfall and sediment response. That these events had similar characteristics (dry antecedent conditions and smaller storms) to Type 2D events where sediment is proximal highlights the spatial complexity likely in the sediment response. In the larger Mad River watershed, rainfall events, particularly those during summer months, are often spatially isolated. Therefore, events occurring under similar conditions but with very

different sediment pulse timing, may be the result of spatially isolated areas within the watershed mobilizing sediment. Figure-eight Type 5A events also were associated with dry antecedent conditions. This is similar to findings of Seeger et al. (2004) but in contrast to findings reported by Buendia et al. (2016), who observed figure-eight patterns associated with large events and wet antecedent conditions. Figure-eight patterns have also been attributed to a combination of stream bank and streambed sources with a delayed sediment contribution from distinct upstream sources (Eder et al., 2010; Megnounif et al., 2013; Seeger et al., 2004); and therefore, may be considered the result of fairly complex processes with drivers that may be more watershed specific. We also note, that our analysis indicates dry antecedent conditions promote more spatially and temporally heterogeneous hysteresis types.

Seasonal changes in hysteresis types suggest shifting drivers and sources of sediment to streams. We observed Types 2C and 2B events more frequently during the late spring (May/June) and late autumn (November/December) months (Figure 6b). In the Northeastern U.S., stream banks, fields, and hillslopes have less ground cover and are more susceptible to erosion during these months. Other studies observed a similar seasonal trend with clockwise patterns more prevalent in spring or late autumn/winter conditions (Asselman, 1999; Martin et al., 2014; Sherriff et al., 2016). Type 2D events occurred most frequently during summer and early autumn months (July to October). Presentation of hysteresis types was consistent with typical climate patterns observed in the Northeastern U.S. (i.e., more frequent rainfall and wetter antecedent conditions in June/July and drier conditions in August/September followed by increased rainfall and soil moisture again in October/November). During late autumn conditions, Type 4 events were more prevalent. This type of hysteresis has not been studied as extensively as the more common clockwise, counter-clockwise, and Figure-eight classes; but Williams (1989) alluded to a threshold effect at higher discharges. This may suggest sediment sources for Type 4 events are primarily in-channel or bank sources activated by some discharge threshold. Since this event is more common during late autumn and lacks the early sediment pulse present in Type 2D and 2E events, it also suggests possible occurrence of seasonal depletion of proximal sediment sources.

4.4. Effect of Watershed Scale

The timing of sediment delivery from sources within the watershed to the outlet affects the type of hysteresis observed; and therefore, watershed scale needs to be considered. We observed a scale effect in our study that manifest in a significantly different distribution of hysteresis patterns observed at the Mad River outlet (watershed scale) compared to the tributary scale. When comparing the distributions of hysteresis types at Mill, Shepard, and Folsom Brook to the distribution at the Mad River outlet, we observed a notable shift from predominantly clockwise (Class II) hysteretic types to a more uniform distribution (Figure 6a). Asselman (1999) noted a similar shift from clockwise loops in upstream tributaries to counter-clockwise loops in downstream locations. Analysis also suggests that the river flood wave can propagate downstream faster than the sediment pulse and that tributaries of this size often produce clockwise SSC-Q loops, both consistent with previous work (Buendia et al., 2016; Martin et al., 2014; Smith & Dragovich, 2009; Williams, 1989).

The effect of upstream sediment source delivery on downstream hysteresis type was also observed in our rain gauge data. The Mad River subwatersheds are on the spatial scale of 10–50 km^2; as a result, they are more uniformly impacted by rainfall events than the entire Mad River watershed, which frequently experiences rainfall limited to portions of the watershed. Events with large separation between sedigraph and hydrograph peaks (Types 2D, 3B, 3C) tend to have both smaller total stormflow and higher variability among rain gauge totals (Figure 10). This suggests that at the Mad River scale, hysteresis types with the sediment pulse well before or after the hydrograph peak are indicative of rainfall localized in only a portion of the watershed. These findings are consistent with those of Smith and Dragovich (2009) who observed small events were dominated by flow and sediment delivery from a localized portion of the upstream watershed. Gao and Josefson (2012) noted that hysteresis analysis is most often performed for watersheds smaller than 100 km^2; and that application to larger watersheds is largely affected by sediment delivery from upstream tributaries and, therefore, more difficult to isolate the specific source. Our results support this conclusion; however, our simultaneous monitoring of upstream and downstream locations as well as rainfall variability within watersheds allowed more robust analysis at the larger watershed scale.

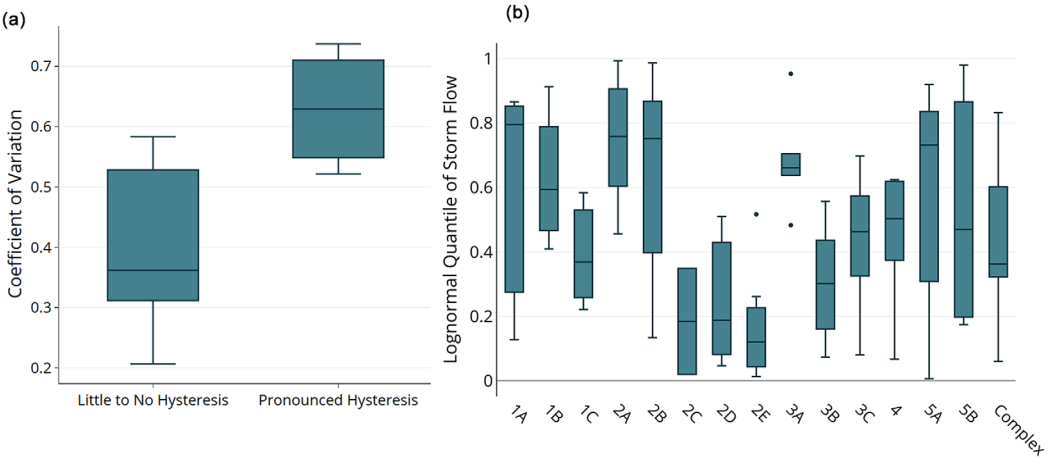

Figure 10. (a) Comparison of coefficient of variation in total event rainfall recorded by rainfall gauges between storms categorized as little to moderate hysteresis (Types 1A, 1B, 1C, 2A, 3A) and those with large separation of rainfall and sediment pulse (Types 2D, 3B, 3C). (b) Lognormal distribution quantile for stormflow of Mad River events by hysteresis type.

4.5. Implications for Watershed Management

Hysteresis analysis has been used to infer sediment dynamics worldwide with the goal of reducing soil loss and sediment export (Lefrançois et al., 2007; Ramos et al., 2015; Sherriff et al., 2016). Our study offers an expanded classification scheme of hysteresis patterns that leverages high-frequency sediment data to better understand sediment dynamics within a watershed. With SSC-Q hysteresis being monitored in watersheds around the world, we believe the pattern types observed in our study are likely transferable to other locations. However, we do not presume this classification as an exhaustive set of patterns.

Hysteresis analysis relies solely on the timing and shape of the sedigraph and hydrograph, relatively low-cost monitoring equipment such as stage and turbidity sensors may be used without necessarily creating rating curves for discharge and SSC. Aich et al. (2014) demonstrated a rapid assessment survey, in which only a small number of storms are monitored, that can provide insight into sediment transport processes within a watershed. However, as this study demonstrates, longer term monitoring may uncover important seasonal trends. Monitoring for shifts in hysteresis patterns also offers potential to assess changes in sediment production and delivery in response to watershed management activities.

We noted differences among the subwatersheds in terms of the most frequently observed hysteresis types, which help provide insights into sediment sources. In Shepard Brook, for example, Type 2D (as seen in Figures 9c and 9d) and 2E hysteresis were observed more frequently; and Type 4 were less frequent than in other subwatersheds. This suggests, that these storm events were characterized by the availability of nearby, rainfall-activated sediment sources and the absence of distant, upslope sediment sources. Sediment export control efforts could, therefore, investigate sediment sources near the outlet such as gravel road ditches or bare agricultural soils. In Mill Brook and Freeman Brook, Type 4 hysteresis was observed frequently, which is indicative of sediment sources activated only above some discharge threshold. Thus, in those subwatersheds, as part of sediment control efforts, watershed managers could investigate areas with soils subject to erosion under higher discharges (e.g., stream banks and slope failures along the river corridor). These inferences match with potential sediment sources identified during field observations; namely, our observation of bare-earth crop land and nonvegetated drainage ditches short distances upstream of the monitoring site in Shepard Brook and the presence of stream bank erosion and active slope erosion short distances upstream of the monitoring sites on Freeman Brook and Mill Brook.

5. Challenges and Opportunities for Expansion

The tools for automating visual pattern recognition from hydrological data presented here represent a novel application of artificial neural networks and an advancement beyond existing time series forecasting (Abrahart et al., 2012; Maier et al., 2010) and clustering applications (Chea et al., 2016; Pearce et al., 2013).

Our proof-of-concept using an RBM is readily scalable to deep learning algorithms that offer the potential for analyzing even larger data sets. The RBM-based classification of SSC-Q relationships also highlights opportunities for expansion of the methodology beyond analysis of event sediment dynamics. In this study, we combined high-frequency monitoring and detailed pattern recognition to identify differences in SSC-Q event relationships that would be obscured if using only the overall hysteretic behavior (e.g., clockwise versus counter-clockwise) of events. Concentration-discharge relationships using other constituents have found hysteresis patterns not observed in SSC-Q relationships (Lloyd et al., 2016b; Zuecco et al., 2016). Therefore, potential for broader application of this approach to the analysis of event dynamics of other solutes or constituents exists provided high-frequency monitoring data are available.

In this study, the preprocessing routines and the RBM classifier network are developed in MATLAB (version 9.2) using the RBM algorithm presented in Hinton and Salakhutdinov (2006) and Testolin et al. (2013) as a starting point. While both require programing, and therefore, present challenges over a simpler method (e.g., hysteresis index (HI) of Lloyd et al., 2016a, 2016b), we believe the availability of machine-learning algorithms are more accessible than ever; pretrained algorithms are available off the shelf or as a starting point (Cireşan et al., 2012). The RBM used as proof-of-concept here is now available in a variety of common programing languages (e.g., R, MATLAB, and Python). We trained and tested the RBM network in less than one hour on a standard desktop PC with a basic graphics card. In addition, once trained, ANNs may be used directly (e.g., embedded directly on sensor chips) without requiring additional network training. Finally, because this approach is analogous to machine-learning applications used for handwritten character recognition, a number of tutorials are now available based on the similar MNIST benchmark data set (e.g., Eclipse Deeplearning4j Development Team, 2017; Pedregosa et al., 2011). However, the challenges of automating the separation of storm events must be tackled and currently requires a priori programing, regardless of type of event hysteresis analysis.

A challenge encountered in our study was the identification of distinct hysteresis types used in the classification scheme. A possible solution lies with the nature of the RBM classifier model. Because the RBM is a probabilistic network, the probability of correctly classifying each input pattern is computed (i.e., the marginal distribution across classes of hysteresis) along with the suggested classification (supporting information Figure S5). If a group of patterns are routinely unable to be classified with a significant level of confidence, it may suggest that additional hysteresis types exist and that the RBM should be retrained on a larger set of hysteresis types. Complex hysteresis patterns that arise from multiple peak hydrographs (supporting information Figure S6a) present an additional challenge for visual classification and highlights the need for careful consideration of event separation techniques. To minimize the effect of multipeak hydrographs, we followed an approach similar to Sherriff et al. (2016) and accordingly, whenever possible, split back-to-back hydrograph rises into individual rainfall events.

As more sediment data become available, particularly from other watershed locations, classification performance should improve through training RBM models on greater numbers of events. We observed differences in the distribution of hysteresis types between relatively similar watersheds. Therefore, having a greater number of events from a variety of watersheds with different land use, climate, geology, topography, and drainage area offer an opportunity for building a greater understanding of drivers of sediment loading during storms across both time and space. Building this knowledge and meaning behind an expanded set of hysteresis types in SSC-Q relationships offers a practical tool for applications focused on sediment connectivity and sources in watersheds. Furthermore, as tools for interpreting big data evolve, environmental researchers should continue their application in interpreting large, high-frequency data sets to better understand the complex environmental processes and their drivers.

Acknowledgments
Support provided by Vermont EPSCoR, with funds from the National Science Foundation (NSF) grants EPS-1101317 and OIA-1556770, is acknowledged. Additional support was provided by NSF under grant DGE-0925179NSF, the Vermont Water Resources and Lake Studies Center, the Gund Institute for Environment, and the Robert & Patricia Switzer Foundation. We thank Kristen Underwood for fieldwork contributions and acknowledge landowners of study sites for their support. The data used in this study are freely available from CUAHSI HydroShare (http://www.hydroshare.org/resource/6b1783d011d94928a2b7d26a4d5ecb96). MATLAB code for the RBM implementation used in this study is available on GitHub (https://github.com/shamshaw/RBM-Hysteresis-Classifier).

References

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., et al. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography*, *36*(4), 480–513. https://doi.org/10.1177/0309133312444943

Aich, V., Zimmermann, A., & Elsenbeer, H. (2014). Quantification and interpretation of suspended-sediment discharge hysteresis patterns: How much data do we need? *CATENA*, *122*, 120–129. https://doi.org/10.1016/j.catena.2014.06.020

Asselman, N. E. M. (1999). Suspended sediment dynamics in a large drainage basin: The River Rhine. *Hydrological Processes*, *13*(10), 1437–1450. https://doi.org/10.1002/(SICI)1099-1085(199907)13:10<1437::AID-HYP821>3.0.CO;2-J

Buendia, C., Vericat, D., Batalla, R. J., & Gibbins, C. N. (2016). Temporal dynamics of sediment transport and transient in-channel storage in a highly erodible catchment. *Land Degradation & Development*, *27*(4), 1045–1063. https://doi.org/10.1002/ldr.2348

Chea, R., Grenouillet, G., & Lek, S. (2016). Evidence of water quality degradation in lower Mekong basin revealed by self-organizing map. *Plos One*, *11*(1), e0145527. https://doi.org/10.1371/journal.pone.0145527

Cheraghi, M., Jomaa, S., Sander, G. C., & Barry, D. A. (2016). Hysteretic sediment fluxes in rainfall-driven soil erosion: Particle size effects. *Water Resources Research*, *52*, 8613–8629. https://doi.org/10.1002/2016WR019314

Cireşan, D. C., Meier, U., & Schmidhuber, J. (2012). Transfer learning for Latin and Chinese characters with Deep Neural Networks. In *The 2012 international joint conference on neural networks (IJCNN)* (pp. 1–6). Brisbane, Australia: IEEE. https://doi.org/10.1109/IJCNN.2012.6252544

Duvert, C., Gratiot, N., Evrard, O., Navratil, O., Némery, J., Prat, C., et al. (2010). Drivers of erosion and suspended sediment transport in three headwater catchments of the Mexican Central Highlands. *Geomorphology*, *123*(3–4), 243–256. https://doi.org/10.1016/j.geomorph.2010.07.016

Eclipse Deeplearning4j Development Team (2017). *Deeplearning4j: Open-source distributed deep learning for the JVM. Apache Software Foundation License 2.0*. San Fancisco, CA: SkyMind. Retrieved from http://deeplearning4j.org

Eder, A., Strauss, P., Krueger, T., & Quinton, J. N. (2010). Comparative calculation of suspended sediment loads with respect to hysteresis effects (in the Petzenkirchen catchment, Austria). *Journal of Hydrology*, *389*(1–2), 168–176. https://doi.org/10.1016/j.jhydrol.2010.05.043

Fischer, A., & Igel, C. (2014). Training restricted Boltzmann machines: An introduction. *Pattern Recognition*, *47*(1), 25–39. https://doi.org/10.1016/j.patcog.2013.05.025

Fitzgerald, E. P., & Godfrey, L. C. (2008). *Upper mad river corridor plan*. Waitsfield, VT: Friends of the Mad River.

Gao, P. (2008). Understanding watershed suspended sediment transport. *Progress in Physical Geography*, *32*(3), 243–263. https://doi.org/10.1177/0309133308094849

Gao, P., & Josefson, M. (2012). Event-based suspended sediment dynamics in a central New York watershed. *Geomorphology*, *139–140*, 425–437. https://doi.org/10.1016/j.geomorph.2011.11.007

Gellis, A. C. (2013). Factors influencing storm-generated suspended-sediment concentrations and loads in four basins of contrasting land use, humid-tropical Puerto Rico. *CATENA*, *104*, 39–57. https://doi.org/10.1016/j.catena.2012.10.018

Hamshaw, S. D., Rizzo, D. M., Wemple, B. C., & Underwood, K. L. (2016). Classification and prediction of event-based suspended sediment dynamics using artificial neural networks. Poster presented at the American Geophysical Union 2016 Fall Meeting, San Francisco, CA.

Hinton, G. E. (2012). A practical guide to training restricted Boltzmann machines. In G. Montavon, G. B. Orr, & K.-R. Müller (Eds.), *Neural networks: Tricks of the trade*. (pp. 599–619). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-35289-8_32

Hinton, G. E., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, *29*(6), 82–97. https://doi.org/10.1109/MSP.2012.2205597

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*(7), 1527–1554.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Landers, M. N., & Sturm, T. (2013). Hysteresis in suspended sediment to turbidity relations due to changing particle size distributions. *Water Resources Research*, *49*, 5487–5500. https://doi.org/10.1002/wrcr.20394

Langlois, J. L., Johnson, D. W., & Mehuys, G. R. (2005). Suspended sediment dynamics associated with snowmelt runoff in a small mountain stream of Lake Tahoe (Nevada). *Hydrological Processes*, *19*(18), 3569–3580. https://doi.org/10.1002/hyp.5844

Larochelle, H., Mandel, M., Pascanu, R., & Bengio, Y. (2012). Learning algorithms for the classification restricted Boltzmann machine. *Journal of Machine Learning Research*, *13*, 643–669.

Lawler, D. M., Petts, G. E., Foster, I. D. L., & Harper, S. (2006). Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK. *Science of the Total Environment*, *360*(1–3), 109–126. https://doi.org/10.1016/j.scitotenv.2005.08.032

Lefrançois, J., Grimaldi, C., Gascuel-Odoux, C., & Gilliet, N. (2007). Suspended sediment and discharge relationships to identify bank degradation as a main sediment source on small agricultural catchments. *Hydrological Processes*, *21*(21), 2923–2933. https://doi.org/10.1002/hyp.6509

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., & Collins, A. L. (2016a). Technical Note: Testing an improved index for analysing storm discharge-concentration hysteresis. *Hydrology and Earth System Sciences*, *20*(2), 625–632. https://doi.org/10.5194/hess-20-625-2016

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., & Collins, A. L. (2016b). 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. https://doi.org/10.1016/j.scitotenv.2015.11.028

Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software*, *25*(8), 891–909.

Martin, S., Conklin, M., & Bales, R. (2014). Seasonal accumulation and depletion of local sediment stores of four headwater catchments. *Water*, *6*(7), 2144–2163. https://doi.org/10.3390/w6072144

Megnounif, A., Terfous, A., & Ouillon, S. (2013). A graphical method to study suspended sediment dynamics during flood events in the Wadi Sebdou, NW Algeria (1973–2004). *Journal of Hydrology*, *497*, 24–36. https://doi.org/10.1016/j.jhydrol.2013.05.029

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, *26*(7), 1465–1473. https://doi.org/10.1029/WR026i007p01465

O'Connor, P., Neil, D., Liu, S.-C., Delbruck, T., & Pfeiffer, M. (2013). Real-time classification and sensor fusion with a spiking deep belief network. *Frontiers in Neuroscience*, *7*, 178. https://doi.org/10.3389/fnins.2013.00178

Onderka, M., Krein, A., Wrede, S., Martínez-Carreras, N., & Hoffmann, L. (2012). Dynamics of storm-driven suspended sediments in a headwater catchment described by multivariable modeling. *Journal of Soils and Sediments*, *12*(4), 620–635. https://doi.org/10.1007/s11368-012-0480-6

Pearce, A. R., Rizzo, D. M., Watzin, M. C., & Druschel, G. K. (2013). Unraveling associations between cyanobacteria blooms and in-lake environmental conditions in Missisquoi Bay, Lake Champlain, USA, using a modified self-organizing map. *Environmental Science & Technology*, *47*(24), 14267–14274. https://doi.org/10.1021/es403490g

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pietroń, J., Jarsjö, J., Romanchenko, A. O., & Chalov, S. R. (2015). Model analyses of the contribution of in-channel processes to sediment concentration hysteresis loops. *Journal of Hydrology*, *527*, 576–589. https://doi.org/10.1016/j.jhydrol.2015.05.009

PRISM Climate Group (2015). *30-yr Normal precipitation: Annual, period: 1981–2010*. Corvallis, OR: Oregon State University. Retrieved from http://prism.oregonstate.edu

Ramos, T. B., Gonçalves, M. C., Branco, M. A., Brito, D., Rodrigues, S., Sánchez-Pérez, J.-M., et al. (2015). Sediment and nutrient dynamics during storm events in the Enxoé temporary river, southern Portugal. *CATENA*, *127*, 177–190. https://doi.org/10.1016/j.catena.2015.01.001

Sarikaya, R., Hinton, G. E., & Deoras, A. (2014). Application of deep belief networks for natural language understanding. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *22*(4), 778–784. https://doi.org/10.1109/TASLP.2014.2303296

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639.

Seeger, M., Errea, M.-P., Beguería, S., Arnáez, J., Martí, C., & García-Ruiz, J. (2004). Catchment soil moisture and rainfall characteristics as determinant factors for discharge/suspended sediment hysteretic loops in a small headwater catchment in the Spanish Pyrenees. *Journal of Hydrology*, *288*(3–4), 299–311. https://doi.org/10.1016/j.jhydrol.2003.10.012

Sherriff, S. C., Rowan, J. S., Fenton, O., Jordan, P., Melland, A. R., Mellander, P.-E., et al. (2016). Storm event suspended sediment-discharge hysteresis and controls in agricultural watersheds: Implications for watershed scale sediment management. *Environmental Science & Technology*, *50*(4), 1769–1778. https://doi.org/10.1021/acs.est.5b04573

Smith, H. G., & Dragovich, D. (2009). Interpreting sediment delivery processes using suspended sediment-discharge hysteresis patterns from nested upland catchments, south-eastern Australia. *Hydrological Processes*, *23*(17), 2415–2426. https://doi.org/10.1002/hyp.7357

Smolensky, P. (1986). *Information processing in dynamical systems: Foundations of harmony theory*. Boulder, CO: University of Colorado at Boulder, Department of Computer Science. Retrieved from http://www.dtic.mil/docs/citations/ADA620727

Soler, M., Latron, J., & Gallart, F. (2008). Relationships between suspended sediment concentrations and discharge in two small research basins in a mountainous Mediterranean area (Vallcebre, Eastern Pyrenees). *Geomorphology*, *98*(1–2), 143–152. https://doi.org/10.1016/j.geomorph.2007.02.032

Stryker, J., Wemple, B., & Bomblies, A. (2017). Modeling sediment mobilization using a distributed hydrological model coupled with a bank stability model. *Water Resources Research*, *53*, 2051–2073. https://doi.org/10.1002/2016WR019143

Testolin, A., Stoianov, I., De Filippo De Grazia, M., & Zorzi, M. (2013). Deep unsupervised learning on a desktop PC: A primer for cognitive scientists. *Frontiers in Psychology*, *4*, 251. https://doi.org/10.3389/fpsyg.2013.00251

Tieleman, T. (2008). Training restricted Boltzmann machines using approximations to the likelihood gradient. In *Proceedings of the 25th international conference on Machine learning* (pp. 1064–1071). New York, NY: ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1390290

Vaughan, M. C. H., Bowden, W. B., Shanley, J. B., Vermilyea, A., Sleeper, R., Gold, A. J., et al. (2017). High-frequency dissolved organic carbon and nitrate measurements reveal differences in storm hysteresis and loading in relation to land cover and seasonality: High-resolution doc and nitrate dynamics. *Water Resources Research*, *53*, 5345–5363. https://doi.org/10.1002/2017WR020491

Vercruysse, K., Grabowski, R. C., & Rickson, R. J. (2017). Suspended sediment transport dynamics in rivers: Multi-scale drivers of temporal variation. *Earth-Science Reviews*, *166*, 38–52. https://doi.org/10.1016/j.earscirev.2016.12.016

Walling, D. E., Collins, A. L., & Stroud, R. W. (2008). Tracing suspended sediment and particulate phosphorus sources in catchments. *Journal of Hydrology*, *350*(3–4), 274–289. https://doi.org/10.1016/j.jhydrol.2007.10.047

Wemple, B. C., Clark, G. E., Ross, D. S., & Rizzo, D. M. (2017). Identifying the spatial pattern and importance of hydro-geomorphic drainage impairments on unpaved roads in the northeastern USA. *Earth Surface Processes and Landforms*, *42*(11), 1652–1665. https://doi.org/10.1002/esp.4113

Williams, G. P. (1989). Sediment concentration versus water discharge during single hydrologic events in rivers. *Journal of Hydrology*, *111*(1), 89–106.

Yeshaneh, E., Eder, A., & Blöschl, G. (2014). Temporal variation of suspended sediment transport in the Koga catchment, North Western Ethiopia and environmental implications. *Hydrological Processes*, *28*(24), 5972–5984. https://doi.org/10.1002/hyp.10090

Zuecco, G., Penna, D., Borga, M., & van Meerveld, H. J. (2016). A versatile index to characterize hysteresis between hydrological variables at the runoff event timescale. *Hydrological Processes*, *30*(9), 1449–1466. https://doi.org/10.1002/hyp.10681