Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

AUTOMATED CHARACTERIZATION OF STREAMWATER SPECIFIC CONDUCTIVITY RESPONSE TO STORMS

Daniel Demers[1], Mark Green[2], and Scott Bailey[3]

[1]Plymouth State University
[2]Case Western Reserve University
[3]U.S. Forest Service

September 17, 2020

Abstract

Specific electrical conductivity (SC) is a basic, effective indicator of water quality. The recent increase in SC data collected with high-frequency sensors has created a strong need for algorithms that can aid interpretation of these data. This study presents an algorithm that finds and quantifies SC temporal patterns and applies that algorithm to a data set from a forested catchment. During and after rain events, we show three patterns that emerge in SC time series: a solute flush, resulting in an initial increase in SC, followed by a dilution, followed by the SC's recovery toward pre-rain conditions. We compared these SC patterns to precipitation amount and intensity, antecedent wetness, and seasonality. Our results indicate that the magnitude of the flush was driven primarily by precipitation intensity and total rainfall during a storm, and secondarily by antecedent moisture conditions. The magnitude of the dilution was driven mainly by precipitation amount. The rate of SC recovery was driven by precipitation amount and was correlated with the dilution. Overall, the algorithm successfully extracted event-driven characteristics in the SC time series, allowing the development of functional relationships with hydrologic drivers. Applying similar methodologies to more catchments in the future will help identify functional relationships at more sites and use these relationships to identify catchments most sensitive to future precipitation changes.

INTRODUCTION

Data documenting the temporal variation of solute concentrations in streams have advanced the knowledge about how catchments produce stream flow and impact streamwater quality. Lab analysis of discrete water samples has underlain this knowledge such as how piston displacement of stored water is a dominant stream flow generation mechanism in many catchments (Sklash and Farvolden 1979), how mineral weathering and soil development are impacted by human activities (Bailey 2020), and how hydrology and soil geochemistry interact to regulate streamwater solutes (e.g., Bennetin et al., 2015), just to name a few examples. However, weekly or less frequent sampling can inadequately capture temporal patterns in the variation of streamwater chemistry (Kirchner 2004). Advancements in technology have allowed us to supplement the information gained from water sampling with high-frequency data from electronic sensors (e.g., Pellerin et al. 2010; von Freyberg et al. 2017).

Stream solute data allow estimation of the fractional contribution of new precipitation versus older, stored water to streamflow (Hooper and Shoemaker 1986; Pellerin et al. 2008). This chemical hydrograph separation can provide insights into how catchments transport rainfall and how catchment characteristics or human modifications alter this hydrologic transport (Pellerin et al. 2008; Inserillo et al. 2107). Specific electrical conductivity (SC) is a common solute indicator used for "new water" estimates (Sklash and Farvolden 1979; Pellerin et al. 2008). Sensors measuring SC are robust, relatively affordable, and have shown that stream chemistry is highly variable on short time scales (Walling 1975; Vogt et al. 2010). Analysis of temporal SC

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

patterns remains an opportunity for advancing catchment hydrology and biogeochemistry. Improved methods for analysis of these high-frequency time series can help expand our understanding of water movement and solute transport through catchments.

During and immediately following storm events, the SC of streams exhibits patterns that provide insights into catchment hydrology and biogechemistry. During a storm, streamwater SC is diluted by rainwater with lower SC than the groundwater that is sourcing streamwater prior to the event (Sklash and Farvolden 1979). Following this, barring further precipitation events, the rainwater continues to drain from the catchment, is evapotranspired, or is integrated into stores, causing SC to recover towards its pre-storm conditions (Walling 1975). Within certain catchments, a solute flush occurs before dilution (Inamdar et al. 2009; Inserillo et al., 2017). Specific conductivity drastically increases on the rising limb of some streams' hydrographs, indicating a flush of ions entering the stream. This flush may represent a distinct source of water entering the stream beyond that of mean stored groundwater and rainwater (Walling 1975, Creed and Band 1998, Inserillo et al. 2017). Pairing meteorologic conditions with sensed, high-frequency time series of stream solutes will aid understanding how precipitation, antecedent hydrologic conditions, and seasonal changes influence the variability of those solutes (e.g., Robson 1992, Biron et al. 1999, Kirchner 2004, Fovet et al. 2018).

The goal of this study was to analyze the temporal behavior of streamwater SC in response to a variety of different precipitation events. To this end, high frequency data from an *in situ* sensor deployment were used to analyze the magnitude of solute flushing (F_{SC}), the magnitude of SC dilution (D_{SC}), and the rate at which SC recovers after a dilution occurs (R_{SC}). We developed an algorithm to objectively analyze these chemograph patterns and produce functional relationships between storm characteristics and stream SC behavior.

METHODS

This study presents an approach to extract, quantify, and explain three temporal patterns within SC time series, which are related to streamwater responses to rain events. These patterns are found in chemographs produced from *in situ* sensors across numerous New Hampshire lotic systems (Inserillo et al. 2017). The first of these patterns is a rapid increase in SC near the beginning of a precipitation event, on the rising limb of a hydrograph, known as a solute flush (Figure 1). The second is the subsequent dilution of SC, as low SC water (i.e., precipitation) mixes with higher SC water (i.e., groundwater); and the third is the recovery of SC as that low conductivity water leaves the catchment and the streamwater returns to a SC resembling groundwater. Although flushing responses do not occur in all streams, dilutions and their subsequent recoveries after a precipitation event are common (e.g., Inserillo et al. 2017).

We developed an algorithm to consistently and objectively locate these temporal patterns in streamwater SC, quantify the $F_{SC}$and D_{SC} behavior, and calculate the R_{SC}that follows. This process was sequential: a flush following the onset of rain, a dilution following a flush, if one was present, and a recovery following a dilution, if one was present. We then compared those streamwater responses to various environmental conditions surrounding the periods in which they occurred, to understand how such independent variables related to each chemograph pattern. The environmental conditions examined were based upon precipitation, antecedent moisture, and the season in which the precipitation event occurred.

Precipitation variables included the total amount of precipitation (P_T), the intensity at which it fell (I_P), and maximum precipitation intensity per 20-minute interval ($I_{P,max}$). Precipitation that occurred after a flush reached its peak would not influence the F_{SC}; likewise, with the maximum dilution and D_{SC}. So, for both flushing and dilution responses, each precipitation variable was analyzed only from the beginning of the storm to the point in time at which the response occurred. Analysis of the R_{SC}included the precipitation of the entire event, as the tail end of a storm presumably impacts the rate at which SC recovers, and as the recovery is an expression low SC water leaving the system or becoming chemically similar to stored water.

Antecedent moisture variables included the length of the inter-storm period (ISP) and the cumulative vapor pressure deficit (ΣVPD) during that period of time. The ISP was determined by the amount of time between the first measurement of rain for a precipitation event and the last measurement of rain for the previous event.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

The ΣVPD for that period of time was indicative of the intensity of drying pressure and was determined by summing the vapor pressure deficit (VPD) over this ISP. VPD was calculated by subtracting vapor pressure (VP) from saturated vapor pressure (VP_S). VP was calculated by multiplying relative humidity (RH) by VP_S. VP_S was calculated using

where T was air temperature in °C, and the resulting $VP_S$was in kPa (Bolton 1980). These temperature and RH data were measured at 20-minute intervals, and all subsequent calculations maintained this time step.

Variables of seasonal impact were the day of year (DoY) on which the precipitation event began and the median air temperature (T_A) of the day on which the event began. In both instances the day on which the event began was the calendar day that the first rain measurement of that precipitation event occurred.

Sites and Data Used

The algorithm we developed was designed around a single streamwater SC sensor that was deployed in Clay Brook, a lotic system in New Hampshire, U.S.A. (Figure 2), which repeatedly exhibited all three SC patterns. We used data from May to November for the 2013 and 2014 calendar years. The SC sensor recorded both electrical conductivity (EC) and water temperature. Using these the two measures, the EC was corrected for temperature and reported as if recorded at 25 °C, giving the SC. SC was calculated using

where T was water temperature in °C and both EC and the resulting SC were in μS cm^{-1}.

Clay Brook is located in the hills at the southern end of the White Mountains. Its headwaters begin on Plymouth Mountain and, flowing generally northward, ends with its confluence with the Baker River, a tributary to the Pemigewasset and Merrimack Rivers, which drain to the Atlantic Ocean. The location of the sensor and pour point of the studied catchment area is roughly halfway through the stream's full watercourse, encompassing a catchment area of 7.49 km^2 (Figure 2). The median elevation of the basin is 359 m, with an average slope of 9 degrees. The catchment is 89.3% forested with a mix of hardwood and coniferous trees. Common species include American beech, red maple, sugar maple, red oak, yellow birch, eastern hemlock, and red spruce. The remaining 10% of the catchment's land cover is split between agriculture and wetlands, including 1.3% development. The soils of the region are generally coarse textured Spodosols of a range of drainage classes.

We used meteorologic data from the Plymouth Regional Airport, Plymouth, NH (NOAA ID K1P1) located 4.6 km from the sensor site at 43.78° latitude, -71.75° longitude, at an elevation of 154 m. We used data from the airport's weather station, which measured air temperature, relative humidity, and precipitation in 20-minute intervals.

The SC sensor data from Clay Brook were measured with an Onset Hobo Data Logger U24, which recorded water temperature and EC at a 4-minute interval. The sensor has an EC accuracy of 3% of the reading value, or 5 μS cm^{-1}, whichever is greater, and a temperature accuracy of 0.1°C. The sensor's operating range is -2°C to 36°C in non-freezing conditions and has a resolution of 1μS cm^{-1} and 0.01°C. The data were screened to remove any records that were consistent with the sensor being out of the water at low flows, and random sensor malfunctions that resulted in one data point being excessively high compared to the preceding and following data points. Independent lab analysis of field grab samples at this site suggest that the sensor accuracy is within 5%. We further screened data to only include periods of time from May 1 to October 31 when regional rainfall-runoff responses are generally not complicated by snow and ice (Figure 3).

Extraction of SC Patterns

The algorithm we developed extracted information from *in situ*stream sensors and weather data and sequentially assessed the three aforementioned patterns that occur in response to precipitation events: the flush, dilution, and recovery. This algorithm fully processed the available data 10,000 times, with each iteration using our defined parameters sampled randomly from a uniform distribution (Table 1). The range for each distribution was chosen through trial and error and visual assessment of the SC time series of the catchment.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

The algorithm first identified unique precipitation events, defined by a parameter, B , which was the amount of time between non-zero precipitation measurements (Figure 4). The time between each precipitation measurement was calculated, and if the gap between one measurement and the previous measurement exceeded B , the latter measurement was determined to be the start of a new, separate event. The total amount of precipitation for each of these unique precipitation events was then calculated, and if that value was lower than a fixed amount, defined by a parameter, A , the streamwater SC characteristics associated with that event were not analyzed.

If a storm was detected, the algorithm then analyzed the changes in streamwater SC. It first determined the pre-storm trajectory (PST): the trend of the stream's SC prior to precipitation, projected forward in time as if no rainwater had been added to the catchment. To find the PST, a Sen (1968) slope estimate of the SC data was fit to the time series. The time period used for the estimate ranged between the first precipitation measurement of the event and B . The PST was then projected forward in time to the periods during and after the occurrence of rain. The algorithm determined the residuals surrounding the PST line and using those residuals, found the SC values that deviated far enough from the PST to be considered a flush or dilution caused by the rain event. The magnitude of the deviation was defined by the number of standard deviations of the residuals, and represented by the parameter G .

To find the peak of the flush, the algorithm created a temporal window ranging between the event's first precipitation measurement and a period of time defined by parameter D , and then found the maximum value of SC during the period. If that maximum value was greater than Gstandard deviations from the PST, it was accepted as a legitimate response to the precipitation event.

The algorithm found the maximum dilution differently whether a flush was detected or not. If a flush was found, the range to look for a maximum dilution was set between the time at which the peak flush occurred and a time period after the storm's last precipitation defined by parameter E . If a flush was not detected, the range was created to be between the 6 hours prior to a storm's last precipitation and E . In either instance, the algorithm then found the minimum value of SC within that window. If that minimum value was less than Gstandard deviations from the PST, it was accepted as a legitimate response to the precipitation event.

The values of both the peak flush and maximum dilution were added or subtracted, respectively, from what the SC would have been if no precipitation had occurred. This value was determined by calculating the SC value on the PST at the time in which both responses reached their apex. The difference resulting from our observed stream SC behavior and what the PST suggested the SC would otherwise be was defined as the F_{SC} and D_{SC}, both having units of μS cm^{-1}.

To calculate the R_{SC}, a Sen's slope estimation was again used. The time range for this estimate began at the point of maximum dilution, if one was determined to be present, and ended an amount of time later defined by parameter F . The slope of the resulting line was classified as the R_{SC}, having a unit of μS cm^{-1} day^{-1}.

If a flushing response was either not detected or was not deemed different enough from normal variations in the SC, the algorithm still searched for a dilution response. However, if a dilution was not detected or deemed variable enough, the R_{SC} was not calculated, because there would be no beginning point from which to conduct the Sen's slope estimation.

Creating Robust Results

We did not know the best parameter values, nor their distribution, so we assigned a uniform distribution for each parameter and used a Monte Carlo approach so that we could produce more robust results (Paxton et al 2001). We randomly sampled 10,000 values of the six parameters that were used in the algorithm (Table 1). These varying parameters allowed us to simulate many combinations of parameter values. Therefore, each iteration of the algorithm produced a different estimate of our focus variables (e.g. F_{SC}, P_T). Thus, we created a rule set to decide how to summarize these Monte Carlo iterations. We narrowed the available pool of iterations for each single precipitation event based on first, how frequently a distinct precipitation event was detected, and second, how often that unique event exhibited a chemograph response.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

Subsequently, we considered precipitation events highly likely to represent the reality of conditions and catchment characteristics only if the algorithm detected that unique event within at least 70% of its iterations (7,000 or more instances out of the 10,000). Of the storms that passed the 70% threshold, only those that resulted in a SC response 60% of the time were considered highly likely to have a SC response present (a minimum of between 4,200 responses out of 7,000 existing storm events and 6,000 responses of 10,000 existing storm events). The precipitation events that survived this culling were summarized by the median values of each dependent variable: F_{SC}, D_{SC}, R_{SC}. These were then compared to the independent, environmental factors surrounding the event.

Data Analysis

Kendall's correlation τ was used to quantify the monotonic relationships between independent and dependent variables. The Kendall correlation utilized a two-sided p-value, which we accepted as statistically significant below an α value of 0.05. We also conducted a multivariate analysis utilizing multiple linear regression in order to quantify multivariate relationships with the dependent variables. All dependent and independent variables, and their equivalent base 10 log, were tested for normality using a Shapiro-Wilk test. If the non-transformed variable was deemed normal by having a p-value of greater than 0.05, it was used. If it was non-normal, the log transformed variable was used only of it was deemed to be more normal than the non-transformed variable. We used Bayesian Information Criterion (BIC) to choose the best combination of variables driving the relationships (Burnham and Anderson 2004) and accepted a p-value of 0.05 to be statistically significant.

RESULTS

Assessing the Algorithm

We visually assessed the accuracy of the algorithm by plotting the algorithm-generated chemograph characteristics with the actual precipitation and SC data. Each analyzed storm was evaluated graphically, including the distribution of estimated $F_{SC}$and D_{SC} values (Figures 5 and 6). The median values of the slopes of both the PST and R_{SC} were matched with their median intercepts so that each could be plotted. This resulted in a total of 34 days on which a storm began, 29 of which showed a flushing response and 28 showed a dilution and recovery. For one storm event, the algorithm did not detect the chemograph responses which we were assessing. Each of the remaining 33 storms had either a flush or dilution and subsequent recovery, or both.

Univariate Correlations

The correlations between the F_{SC} and the independent variables revealed statistically significant relationships with I_P, $I_{P,max}$, and P_T(Table 2, Figure 7). These 3 relationships were drawn from 27 algorithm-identified flushes in response to precipitation events. This number differed from the 29 available flushes due to a gap in the VPD data record. The Kendall's τ value for I_P was 0.502 ($p<0.001$), for $I_{P,max}$ was 0.486 ($p<0.001$), and for P_T was 0.422 ($p=0.001$).

The correlations between the D_{SC} and the independent variables also showed statistically significant relationships with I_P, $I_{P,max}$, and P_T(Figure 8). These 3 relationships were drawn from 28 algorithm-identified dilutions in response to precipitation events. The Kendall's τ value for I_P was 0.407 ($p=0.003$), for $I_{P,max}$ was 0.38 ($p=0.006$), and for $P_T$was 0.628 ($p<0.001$).

The correlations between the R_{SC} and the independent variables resulted in a statistically significant relationship with P_T (Figure 9). This relationship was drawn from 28 algorithm-identified recoveries in response to precipitation events. The Kendall's τ value for P_T was 0.564 ($p<0.001$). There was also a significant relationship with another dependent variable: D_{SC} ($\tau=0.45$; $p<0.001$).

Multiple Linear Regressions

The multiple linear regression between the log transformed F_{SC} and the independent variables showed three variables driving the relationship (Table 3). The correlation with log transformed I_P, log transformed $I_{P,Max}$,

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

and log transformed ΣVPD resulted in $p<0.001$ and an adjusted R^2 of 0.54. The equation of the regression was

The multiple linear regression between the D_{SC} and the independent variables showed two variables driving the relationship. The correlation with log transformed P_T and the slope of the PST resulted in $p<0.001$ and an adjusted $R^2$of 0.78. The equation of the regression was

The multiple linear regression between R_{SC} and the independent variables showed a single variable driving the relationship. The correlation with log transformed P_T resulted in a $p<0.001$ and an adjusted R^2 of 0.39. The equation of the regression was .

DISCUSSION

The algorithm was effective at extracting patterns of solute flushing, dilution, and recovery. Quantifying these chemograph characteristics and comparing them to the environmental variables revealed that the three precipitation variables exerted the greatest influence on these patterns. The methodology we used here can be expanded to other catchments in order to characterize and compare them using the easily obtained, high-frequency data sources that have been produced and collected in recent years. Given the ability of the algorithm to extract both solute flush and dilution characteristics, we expect that other temporal behaviors that occur in other streams could be extracted with the general approach that we used.

Efficacy of the Algorithm

There has been a major increase in *in situ* sensor-obtained streamwater data (e.g., Pellerin et al. 2010, Rode et al. 2016, Duncan et al. 2018, Fovet et al. 2018), creating a need for automated analyses that explore the data, reveal patterns, and aid in interpretation. There are numerous methods to analyze water quality time series (Hirch et al. 1982, Cun and Vilagines 1997, Faruk 2009), however these were not well-suited to extracting the characteristic hydrologic patterns which were our focus, due to the thresholds and conditional decisions needed to identify each pattern. Our methodology presents a novel way to extract useful information from the deluge of data provided by modern high-yield sensors. While we demonstrate the approach for SC, the general approach may be useful for data from other types of water quality sensors. Our algorithm's extraction of SC patterns was assessed to be accurate and similar to what would be obtained from human estimation (Figures 5 and 6). The rare instances in which the algorithm was not accurate are acceptable, given the beneficial nature of the automation; essentially that the process is highly robust due to the Monte Carlo approach, which has previously been shown to accurately extract complicated patterns from time series (e.g., Contosta et al. 2016). Similar approaches using machine learning would possibly outperform our algorithm, and we see this as a likely next step in analyzing streamwater sensing data.

Environmental Controls on Chemograph Characteristics

Both flushing and dilution are strongly correlated with rainfall-based variables. This leads to the assessment that both are primarily driven by the mixing of different sources of water, rather than more nuanced controls like seasonality of solute generation or variation in storage (as represented by antecedent moisture indicators). The importance of source water mixing, and the relative contributions of new and old water to streamflow, has been previously discussed at length (Pinder and Jones 1969, Sklash and Farvolden 1979, Kirchner 2003) and are consistent with our results involving SC dilution. However, the functional relationships we show (Figure 8), obtained solely from SC and precipitation data, may be useful in characterizing hydrologic systems for comparison.

In the case of flushing events, our results suggest that a third source of solute-laden water (aside from the typically assumed two members: groundwater and rainwater) is entering the stream system before precipitation is transported to the stream, resulting in SC dilution (Robson et al. 1992, Creed and Band 1998, Inamdar et al. 2009). This third source, however, may quickly exhaust its supply, resulting in water following the same pathway but exhibiting a lower concentration as a storm progresses. For example, at the nearby Hubbard Brook Experimental Forest, the upper intermittent stream reaches are characterized by high dissolved organic carbon and aluminum, and low pH, due to eluvial processes in the bedrock controlled

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

soils that predominate the upper catchment (Bailey et al. 2019). As the catchment wets up, the active portion of the stream network may quickly expand into the upper catchment, causing a flush of solutes from near-stream soils , that may then be quickly diluted by precipitation or from lower SC water as deeper soils lower in the catchment increase their contribution to runoff. Analyses of flushing would benefit from three-end member mixing models being used in future studies, requiring sensing of additional chemical tracers (Burns et al. 2001, Hooper 2003). SC has been an effective catchment hydrology tracer in the past (Davis et al. 1980, Pellerin et al. 2008, Cox et al. 2007), however, the role of the solute flushing in mixing analysis needs to be addressed more thoroughly.

The univariate results indicate that this third source of water is driven into the system not simply by the amount of rainwater added, but maybe more importantly by the intensity at which it is added. Rainwater can cause flushing events in two ways. First, it can displace stored high SC water into the stream during precipitation events, or second, it can reach the stream as high SC water after gaining solutes via mixing during transport (Robson et al. 1992). Because F_{SC} is more correlated with I_P and $I_{P,max}$ than P_T, we hypothesize that the flush of high SC water is from either near-stream soils, ephemeral streambeds, or development of saturated conditions in shallow to bedrock soils: pathways which allow more rapid movement of water to the watershed outlet, whether that water is precipitation that picked up solutes or is the displacement of ionically enriched stored water.

If this is the case, we predict that the solutes responsible for this flush are those which are shallow (e.g. dissolved organic C and nitrate) rather than those which are weathering-derived (e.g. Ca, Na). If piston flow later takes over as the major contributor of high SC water, as has been shown to be the case in systems without flushing events (Sklash and Farvolden 1979), then we hypothesize that the solutes responsible would be from deeper, weathering-derived sources. Sample collection during precipitation events would indicate which specific solutes are entering the stream system and allow an evaluation of which of these processes drives SC dynamics.

The R_{SC} relationship with P_T is expected given it is driven by the dissipation of a pulse of low SC rainwater through higher SC stored water, however we see potential to further explore this chemograph characteristic to understand catchment hydrologic function. As more rainwater enters the system and increases the deviation from the catchment's normal, high-SC state (shown by a higher D_{SC}), the system responds by recovering more quickly. As the new water leaves the system, the SC will move towards pre-event concentrations. The rate at which this rebound occurs may be used to characterize the catchment's export of new water. It may be possible to derive hydrogeologic characteristics using R_{SC}, similar to previous hydrograph recession analysis (Brutsaert and Nieber 1977).

The best multiple linear regression models we produced for D_{SC} and R_{SC} ended up with P_T as a primary driver, accompanied in the case of D_{SC} with PST, which is due to the manner in which the algorithm quantifies the D_{SC}. The current analysis for these two SC patterns suggests that the system was so influenced by precipitation that the importance of other controls, while physically relevant, were either very difficult to detect, were relatively unimportant, or were restricted by the constraints of available data. Similarly, no seasonal variables significantly influenced flushing, dilution, or recovery behaviors, suggesting that seasonal variation of runoff generation was similarly masked by precipitation characteristics or due to us limiting our analysis to the snow-free season.

The lack of correlation between antecedent conditions (ΣVPD and ISP) and D_{SC} does not agree with previous studies, and accordingly is unexpected, as the drying pressure exerted on a catchment influences the hydrologic storage capacity of its soils, and thus runoff generation (Biron et al. 1999, Detty and McGuire 2010, Grand-Clement et al. 2014). It is possible that there is a threshold system response to drying, rather than a gradient response, thus revealing no correlations in our analyses pertaining to ΣVPD or ISP. Or, drying pressure's effect may be masked by the importance of precipitation amount and intensity. We expect that this result may change if we had more storms to analyze and could then apply data-intensive, non-parametric multivariate models (e.g., regression trees) due to the complex nature of hydrologic systems.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

Cumulative vapor pressure deficit did, however, emerge as a driving factor in our multiple linear regression model for F_{SC}, indicating that the drying pressure exerted on the catchment is a factor in flushing responses. The antecedent conditions here likely influence the connectivity of soil water to the stream, becoming more disconnected with greater ΣVPD. As rainfall connects soils to the stream, a stronger pathway is formed, allowing ions to more freely mobilize. Again, this relationship may become stronger and more clearly important to F_{SC} with the study of additional storms.

Climate change is causing more intense precipitation events with longer inter-storm periods (Allan and Soden 2008, Yu et al. 2016). Our results indicate a likelihood of higher variability in streamwater SC as these climate-driven effects take hold. The greater precipitation amounts and intensities will cause higher spikes in the concentration of solutes at the beginning of a precipitation event and higher dilution of stream solutes later in events. This may have a negative effect on lotic organisms that rely on certain ranges of SC, or concentrations of the ions that SC represents, to carry out biologic functions (Daley et al. 2009). And, depending on the solutes making up the ionic composition of the streamwater, these may impact the waters used for human consumption. Thus monitoring of SC and automated detection of responses provides an effective tool for surface water assessments.

CONCLUSION

Our algorithm extracted three chemograph patterns, which routinely appear in response to rain events. Our comparison of these patterns to independent environmental conditions suggested that the major factors driving the variability in flushing, dilution, and recovery behavior were the precipitation amount and intensity. Seasonal variables were largely unimportant at this catchment or overshadowed by the heavy influence of rainfall. When analyzed using multiple linear regression, antecedent moisture was shown to be a factor in driving flushing responses, but not in driving dilution or recovery responses. Our methodology can be useful for analysis of other catchments, revealing functional relationships between independent, environmental variables and streamwater SC responses. The results of such investigations could help characterize catchments in a robust, objective, and repeatable manner, with implications for guiding water resource management decisions.

ACKNOWLEDGMENTS

We thank Errin Volitis, Ashley Inserillo, Jen Taylor, and Donovan King for help maintaining sensors, and Shannon Rogers for useful comments on earrlier drafts of this work.

DATA AVAILABILITY STATEMENT

The water quality sensing data that support the findings of this study are openly available from the University of New Hampshire Data Discovery Center at http://ddc-lovotecs.sr.unh.edu/about/project_description/. The meteorological data are openly available from the United States National Oceanic and Atmopheric Administration at https://mesonet.agron.iastate.edu/request/download.phtml. Code used for this analysis is available at https://github.com/danieljdemers/streamresponse.

Works Cited

Allan, R. P., & Soden, B. J. (2008). Atmospheric warming and the amplification of precipitation extremes. *Science* , *321* (5895), 1481-1484.

Bailey, S. W. (2020) Tracking the Fate of Plagioclase Weathering Products. *Biogeochemical Cycles: Ecological Drivers and Environmental Impact* , 248:151-162.

Benettin, P., Bailey, S. W., Campbell, J. L., Green, M. B., Rinaldo, A., Likens, G. E., ... & Botter, G. (2015). Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH, USA. Water Resources Research, 51(11), 9256-9272.

Biron, P. M., Roy, A. G., Courschesne, F., Hendershot, W. H., Côté, B., & Fyles, J. (1999). The effects of antecedent moisture conditions on the relationship of hydrology to hydrochemistry in a small forested

watershed. *Hydrological Processes* , *13* (11), 1541-1555.

Bolton, D. (1980). The computation of equivalent potential temperature.*Monthly weather review* , *108* (7), 1046-1053.

Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resour. Res* , *13* (3), 637-643.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological methods & research* , *33* (2), 261-304.

Burns, D. A., McDonnell, J. J., Hooper, R. P., Peters, N. E., Freer, J. E., Kendall, C., & Beven, K. (2001). Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research Watershed (Georgia, USA). *Hydrological Processes* , *15* (10), 1903-1924.

Contosta, A. R., Adolph, A., Burchsted, D., Burakowski, E., Green, M., Guerra, D., ... & Routhier, M. (2016). A longer vernal window: the role of winter coldness and snowpack in driving spring transitions and lags. *Global change biology* .

Cox, M. H., Su, G. W., & Constantz, J. (2007). Heat, chloride, and specific conductance as ground water tracers near streams. *Ground Water* , *45* (2), 187-195.

Creed, I. F., & Band, L. E. (1998). Export of nitrogen from catchments within a temperate forest: evidence for a unifying mechanism regulated by variable source area dynamics. *Water Resources Research* , *34* (11), 3105-3120.

Cun, C., & Vilagines, R. (1997). Time series analysis on chlorides, nitrates, ammonium and dissolved oxygen concentrations in the Seine river near Paris. *Science of the total environment* , *208* (1), 59-69.

Daley, M. L., Potter, J. D., & McDowell, W. H. (2009). Salinization of urbanizing New Hampshire streams and groundwater: effects of road salt and hydrologic variability. *Journal of the North American Benthological Society* ,*28* (4), 929-940.

Davis, S. N., Thompson, G. M., Bentley, H. W., & Stiles, G. (1980). Ground-Water Tracers—A Short Review. *Ground water* , *18* (1), 14-23.

Detty, J. M., & McGuire, K. J. (2010). Threshold changes in storm runoff generation at a till-mantled headwater catchment. *Water Resources Research* , *46* (7).

Duncan, J. M., Band, L. E., & Groffman, P. M. (2017). Variable nitrate concentration–discharge relationships in a forested watershed.*Hydrological Processes* , *31* (9), 1817-1824.

Faruk, D. O. (2009). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence* , *23* (4), 586-594.

Fovet, O., Humbert, G., Dupas, R., Gascuel-Odoux, C., Gruau, G., Jaffrezic, A., ... & Grimaldi, C. (2018). Seasonal variability of stream water quality response to storm events captured using high-frequency and multi-parameter data. *Journal of Hydrology* ,*559* , 282-293.

Grand-Clement, E., Luscombe, D. J., Anderson, K., Gatis, N., Benaud, P., & Brazier, R. E. (2014). Antecedent conditions control carbon loss and downstream water quality from shallow, damaged peatlands. *Science of the Total Environment* , *493* , 961-973.

Hirsch, R. M., Slack, J. R., & Smith, R. A. (1982). Techniques of trend analysis for monthly water quality data. *Water resources research* , *18* (1), 107-121.

Hooper, Richard P., and Christine A. Shoemaker. "A comparison of chemical and isotopic hydrograph separation." Water Resources Research 22.10 (1986): 1444-1454.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

Hooper, R. P. (2003). Diagnostic tools for mixing models of stream water chemistry. *Water Resources Research* , *39* (3).

Inamdar, S., Rupp, J., & Mitchell, M. (2009). Groundwater flushing of solutes at wetland and hillslope positions during storm events in a small glaciated catchment in western New York, USA. *Hydrological processes* , *23* (13), 1912-1926.

Inserillo, A., Green, M., Shanley, J. B., & Boyer, J. (2017). Comparing Catchment Hydrologic Response to a Regional Storm Using Specific Conductivity Sensors. Hydrological Processes.

Kirchner, J. W. (2003). A double paradox in catchment hydrology and geochemistry. *Hydrological Processes* , *17* (4), 871-874.

Kirchner, J. W., Feng, X., Neal, C., & Robson, A. J. (2004). The fine structure of water-quality dynamics: the (high-frequency) wave of the future. *Hydrological Processes* , *18* (7), 1353-1359.

Likens, G. E. (2013). *Biogeochemistry of a forested ecosystem* . Springer Science & Business Media.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling* , *8* (2), 287-312.

Pellerin, B. A., Wollheim, W. M., Feng, X., & Vorosmarty, C. J. (2008). The application of electrical conductivity as a tracer for hydrograph separation in urban catchments. *Hydrological Processes* , *22* (12), 1810-1818.

Pellerin, B. A., Saraceno, J. F., Shanley, J. B., Sebestyen, S. D., Aiken, G. R., Wollheim, W. M., & Bergamaschi, B. A. (2010). Taking the pulse of snowmelt: in situ sensors reveal seasonal, event and diurnal patterns of nitrate and dissolved organic matter variability in an upland forest stream. *Biogeochemistry* , *108* (1-3), 183-198.

Pinder, G. F., & Jones, J. F. (1969). Determination of the ground-water component of peak discharge from the chemistry of total runoff. *Water Resources Research* , *5* (2), 438-445.

Robson, A., Neal, C., Smith, C. J., & Hill, S. (1992). Short-term variations in rain and stream water conductivity at a forested site in mid-Wales—implications for water movement. *Science of the total environment* , *119* , 1-18.

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., ... & Skeffington, R. (2016). Sensors in the stream: the high-frequency wave of the present. *Environmental Science & Technology* .

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* , *63* (324), 1379-1389.

Sklash, M. G., & Farvolden, R. N. (1979). The role of groundwater in storm runoff. *Developments in Water Science* , *12* , 45-65.

Vogt, T., Hoehn, E., Schneider, P., Freund, A., Schirmer, M., & Cirpka, O. A. (2010). Fluctuations of electrical conductivity as a natural tracer for bank filtration in a losing stream. *Advances in Water Resources* , *33* (11), 1296-1308.

von Freyberg, J., Studer, B., & Kirchner, J. W. (2017). A lab in the field: high-frequency analysis of water quality and stable isotopes in stream water and precipitation. Hydrology and Earth System Sciences, 21, 1721-1739.

Walling, D. E. (1975). Solute variations in small catchment streams: some comments. *Transactions of the Institute of British Geographers* , 141-147.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

Yu, L., Zhong, S., Pei, L., Bian, X., & Heilman, W. E. (2016). Contribution of large-scale circulation anomalies to changes in extreme precipitation frequency in the United States. *Environmental Research Letters* , *11* (4), 044003.

Figure Captions and Tables

Figure 1: A conceptual model of the specific conductivity (SC) patterns that streamwater takes during and after a precipitation event. Highlighted are the aspects that we use to characterize the watershed response to individual precipitation events: the flush, dilution, and rate of SC recovery. Also shown is the pre-storm trajectory, or the course of streamwater SC prior to the influence of a precipitation event, which is used to quantify both the magnitude of the flush and the magnitude of the dilution.

Figure 2: A map of the Clay Brook catchment.

Figure 3: The full record of specific conductivity (SC) and precipitation data for the months we analyzed in both 2013 and 2014.

Figure 4: A visual representation of the parameter ranges for the Monte Carlo calculations. Parameter A defined the minimum amount of precipitation for an event to be considered for analysis. Parameter B defined both the time between individual precipitation measurements that separated unique events, and the range of time to define the pre-storm trajectory (PST). Parameter D defined the amount of time for a flushing response to be found after the storm began. Parameter E defined the amount of time for a dilution response to be found after the flush occurred, or after 6 hours prior to the end of the storm's precipitation if no substantial flush was present. Parameter F defined the amount of time in which the slope of the rate of specific conductivity (SC) recovery was calculated after the dilution. Parameter G defined the number of standard deviations above or below the PST line that a peak flush or maximum dilution must have exceeded in order to be declared substantial enough for analysis.

Figure 5: The algorithm-derived points of peak flush and maximum dilution for a precipitation event beginning on May 8, 2013. and their respective distributions. Of note are the three points in which a peak flush was found across iterations, due to the variable range of Parameter D, and the increasing specific conductivity (SC) values of those points occurring in time. In this way, the algorithm seeks the truest representation of a flush over time. A) The distributions of algorithm derived peak flush values (number of times a peak flush was detected / number of times the precipitation event was detected). B) The median of the extracted peak flush and maximum dilution values from each of the algorithm's iterations. These points overlay the SC data surrounding the precipitation event. The pre-storm trajectory (PST) and rate of SC recovery (R_{SC}) are also displayed with estimated intercepts for appropriate visual representation. C) The distribution of algorithm derived maximum dilution values (number of times a maximum dilution was detected / number of times the precipitation event was detected).

Figure 6: The algorithm-derived points of peak flush and maximum dilution for a precipitation event beginning on October 4, 2014. and their respective distributions. Unlike Figure 5, the earlier point for peak flush is chosen more frequently by the algorithm because the first has a greater specific conductivity (SC). In this way, the algorithm retains the truest representation of a flush despite later options being available. A) The distributions of algorithm derived peak flush values (number of times a peak flush was detected / number of times the precipitation event was detected). B) The median of the extracted peak flush and maximum dilution values from each of the algorithm's iterations. These points overlay the SC data surrounding the precipitation event. The pre-storm trajectory (PST) and rate of SC recovery (R_{SC}) are also displayed with estimated intercepts for appropriate visual representation. C) The distribution of algorithm derived maximum dilution values (number of times a maximum dilution was detected / number of times the precipitation event was detected).

Figure 7: Kendall's τ relationships for the magnitude of peak flush (F_{SC}) and A) precipitation intensity (I_P), B) maximum precipitation intensity ($I_{P,max}$), and C) total precipitation (P_T). Whiskers indicate the interquartile range of the values obtained from the 10,000 iterations. An absence of whiskers indicates 100%

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

confidence in the value across all iterations.

Figure 8: Kendall's τ relationships for the magnitude of maximum dilution (D_{SC}) and A) precipitation intensity (I_P), B) maximum precipitation intensity ($I_{P,max}$), and C) total precipitation (P_T). Whiskers indicate the interquartile range of the values obtained from the 10,000 iterations. An absence of whiskers indicates 100% confidence in the value across all iterations.

Figure 9: Kendall's τ relationships for the rate of specific conductivity recovery (R_{SC}) and A) precipitation intensity (P_T) and B) magnitude of the dilution ($D_{SC)}$. Whiskers indicate the interquartile range of the values obtained from the 10,000 iterations. An absence of whiskers indicates 100% confidence in the value across all iterations.

Table 1: Ranges for each parameter's uniform distribution. The P_T is the total precipitation for a storm, the R_{SC} is the recovery rate of SC after peak dilution, and PST is the pre-storm trajectory.

Parameter	**Description**	**Distribution**
A	Minimum P_T to qualify a precipitation event	0 to 5 mm
B	Maximum amount of time between precipitation measurements to qualify as same event	12 to 24 hours
D	Maximum time after beginning of event to find the peak flush	0 to 48 hours
E	Maximum time after end of an event to find the maximum dilution	20 min to 48 hour
F	Amount of time after a dilution to calculate R_{SC}	0 to 72 hours
G	Minimum SC deviation from the PST to qualify a flush or dilution	1 to 3.5 standard

Table 2: Kendall's τ values of every relationship analyzed. P-values are indicated as follows: *<=0.1; **<=0.05; ***<=0.01; ****<=0.001.

Variable	**F_{SC}**	**D_{SC}**	**R_{SC}**
I_P	0.499****	0.407***	0.048
$I_{P,max}$	0.532****	0.38***	-0.124
P_T	0.405***	0.628****	0.564****
ISP	0.064	-0.125	0.077
ΣVPD	0.118	0.016	0.069
T_A	0.171	0.175	0.249*
DoY	0.032	-0.132	-0.175
PST	-0.212	0.138	-0.021
F_{SC}	—	-0.011	-0.032
D_{SC}	—	—	0.45****

F_{SC} = Magnitude of the Flush µS cm^{-1}

D_{SC} = Magnitude of the Dilution µS cm^{-1}

R_{SC} = Rate of Specific Conductivity Recovery µS cm^{-1} day^{-1}

I_P = Precipitation Intensity mm hr^{-1}

$I_{P,max}$ = Maximum Precipitation Intensity mm $20min^{-1}$

P_T = Total Precipitation mm

ISP = Inter-storm Period days

ΣVPD = Cumulative Vapor Pressure Deficit kPa

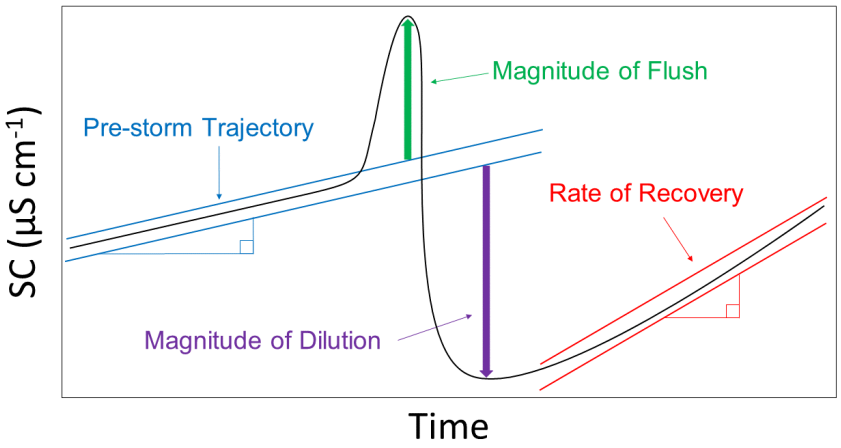
T_A = Median Air Temperature ^oC

DoY = Day of Year —

PST = Pre-Storm Trajectory μS cm^{-1} day^{-1}

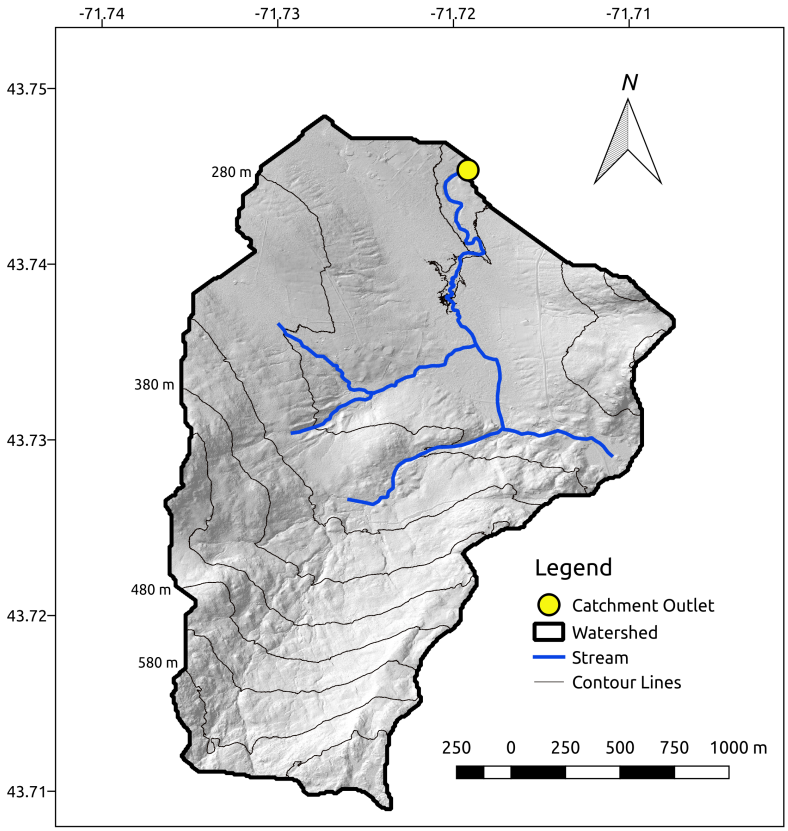
Table 3: Results of the multiple linear regression analysis. The F_{SC} is the SC flush, the D_{SC} is the maximum dilution of SC, and the R_{SC} is the recovery rate of the SC from peak dilution.

Dependent Variable	**Coefficient**	**Independent Variable 1 (Value)**	**Independent Variable 2 (Value)**	**Independent Variable 3 (Value)**
Log F_{SC}	-0.41	Log I_P (0.39)	Log $I_{P,Max}$ (0.47)	Log ΣVPD (0.30)
Log D_{SC}	-17.14	Log P_T (21.29)	Pre-Storm Slope (2.27)	—
R_{SC}	-1.34	Log P_T (2.43)	—	—

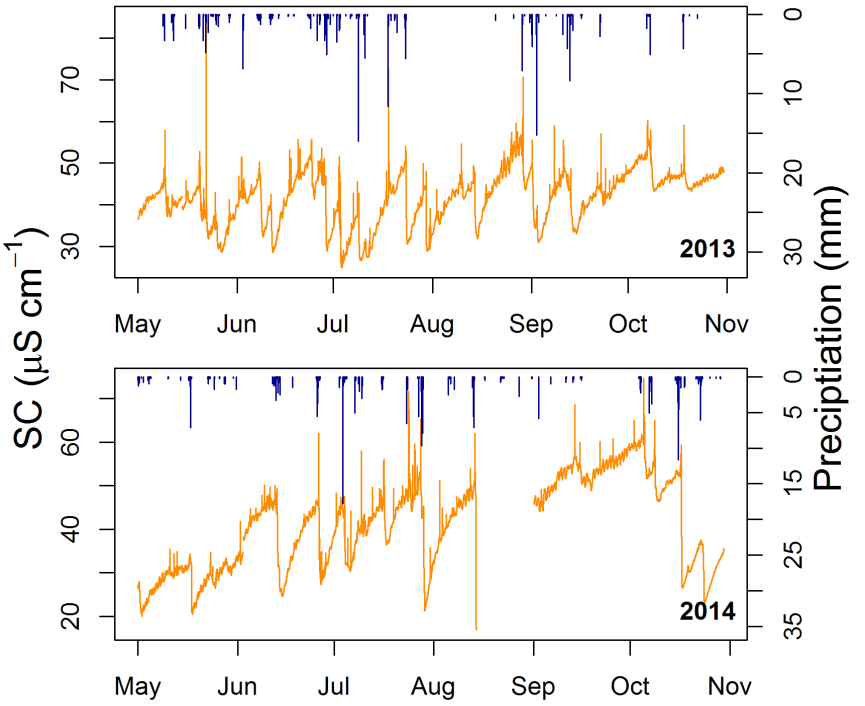


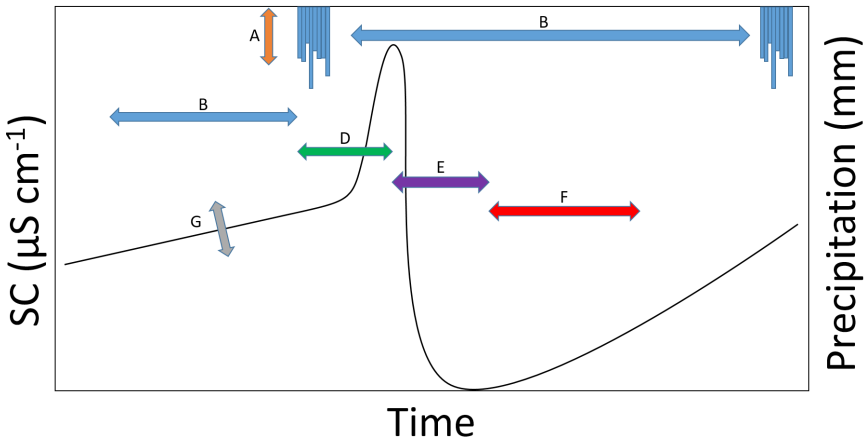
Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

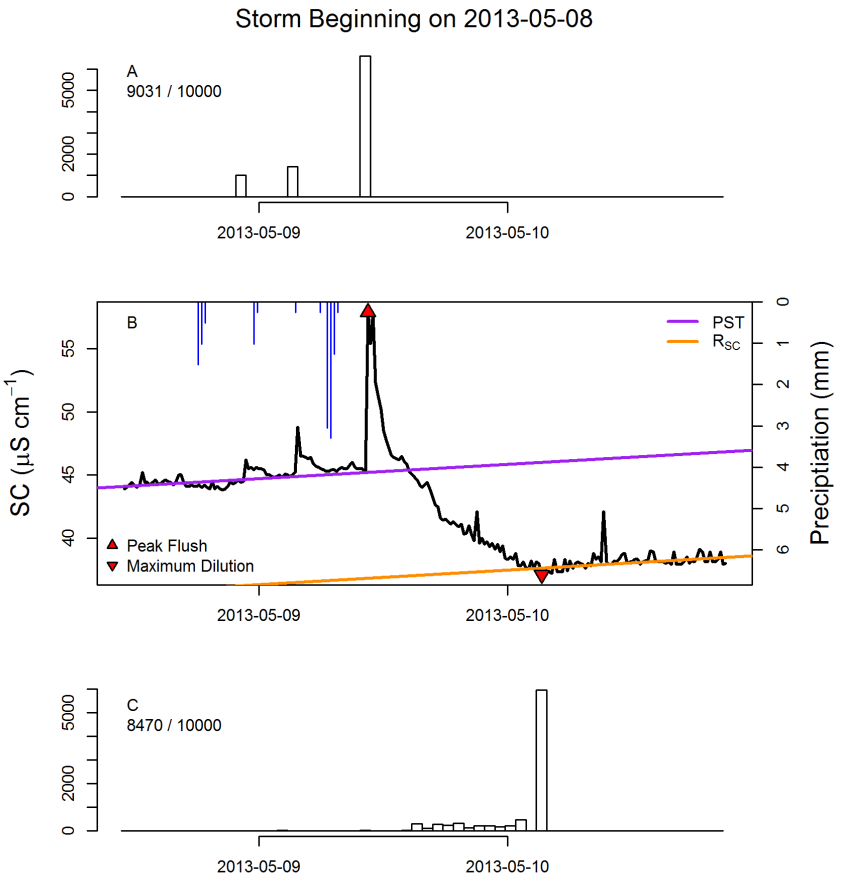


Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.39223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

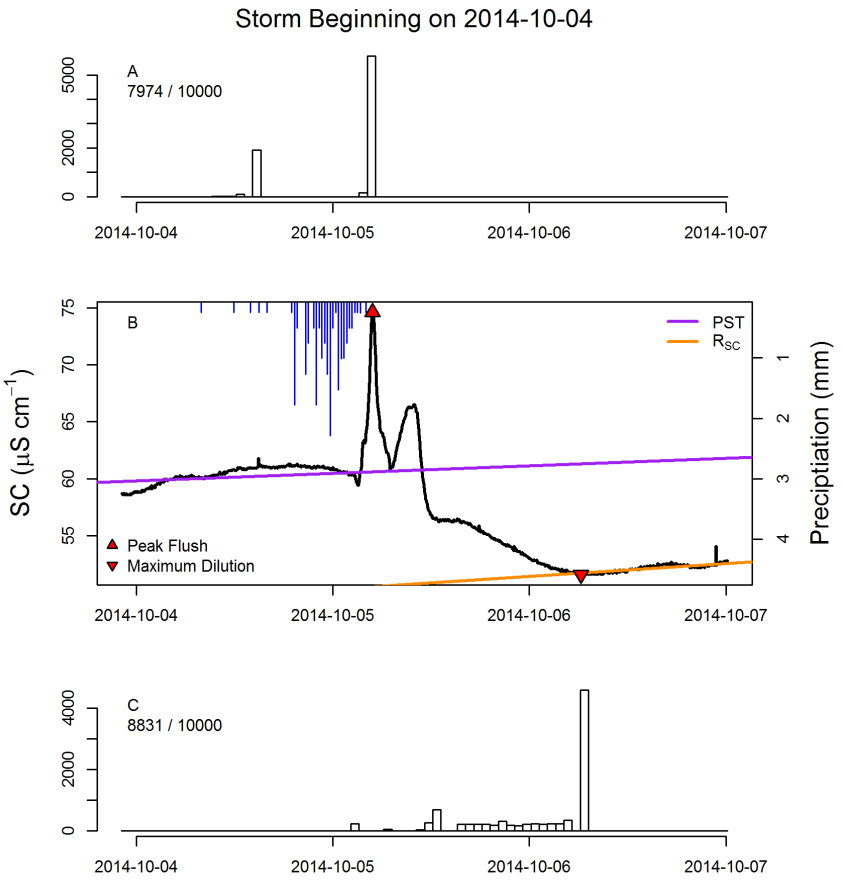




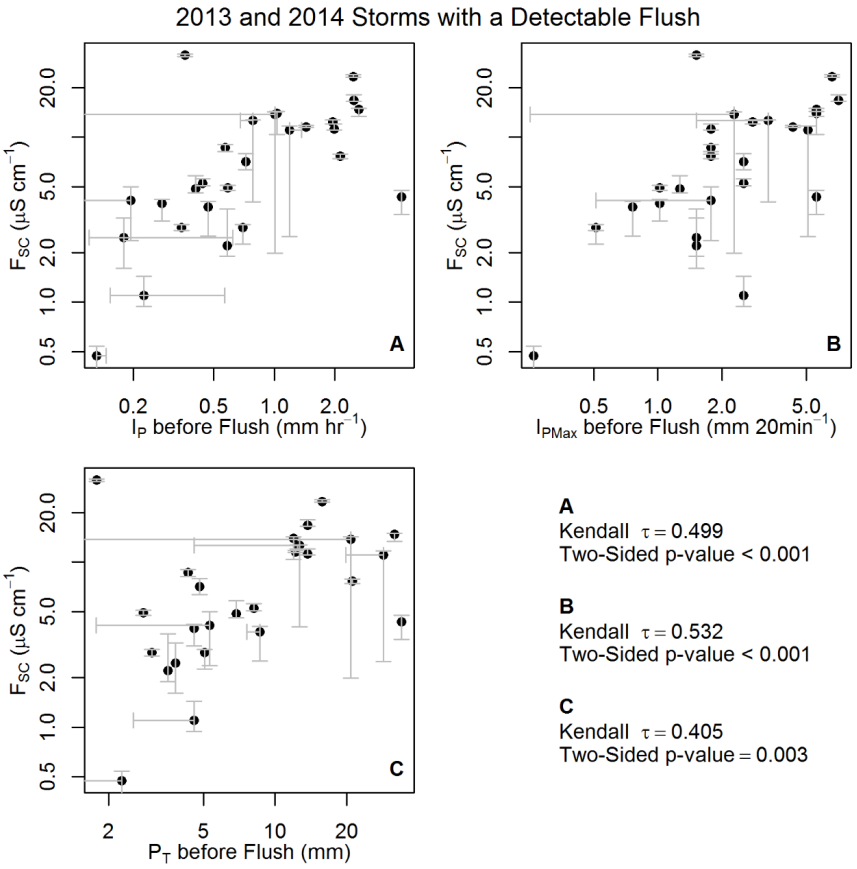
Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.



Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.



Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

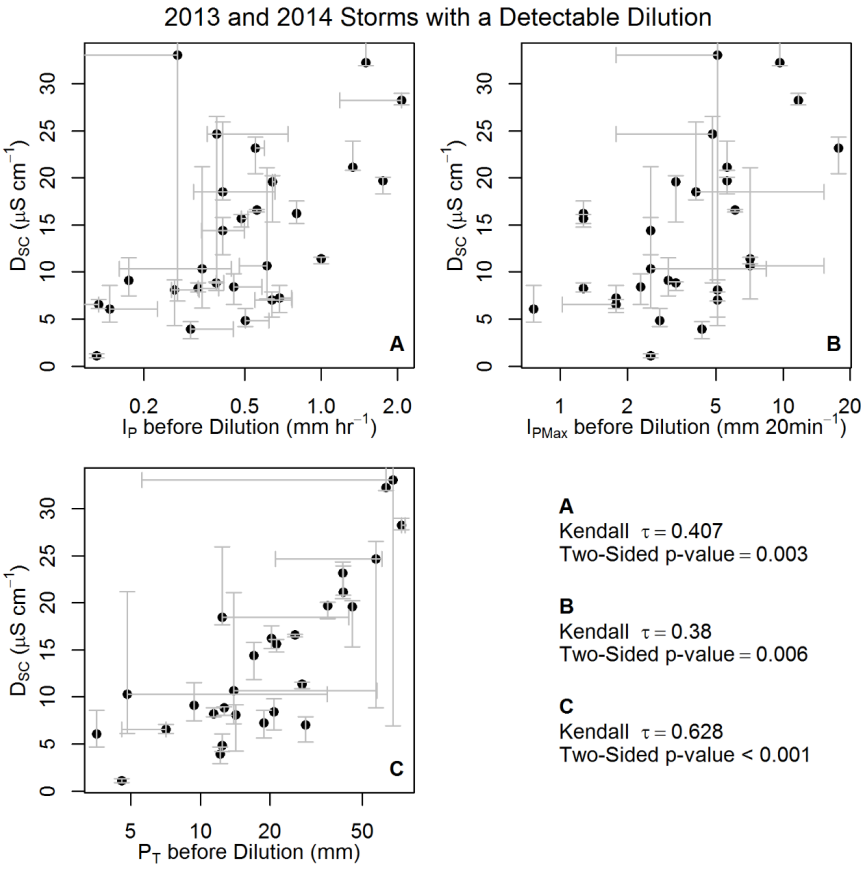


A
Kendall $\tau = 0.499$
Two-Sided p-value < 0.001

B
Kendall $\tau = 0.532$
Two-Sided p-value < 0.001

C
Kendall $\tau = 0.405$
Two-Sided p-value = 0.003

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.

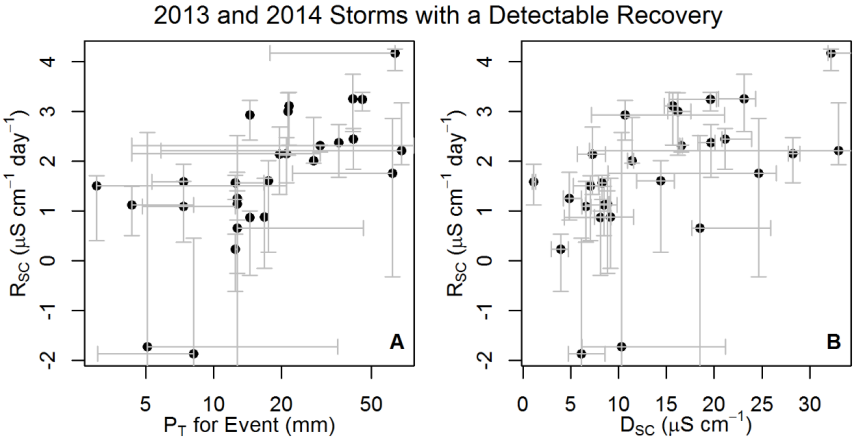


A
Kendall $\tau = 0.407$
Two-Sided p-value = 0.003

B
Kendall $\tau = 0.38$
Two-Sided p-value = 0.006

C
Kendall $\tau = 0.628$
Two-Sided p-value < 0.001

Posted on Authorea 17 Sep 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.160037753.30223767 — This a preprint and has not been peer reviewed. Data may be preliminary.



A
Kendall $\tau = 0.564$
Two-Sided p-value < 0.001

B
Kendall $\tau = 0.45$
Two-Sided p-value < 0.001