

STREAM TURBIDITY MODELING:  
A FOUNDATION FOR WATER QUALITY FORECASTING

By

Amanda L. Mather

A DISSERTATION

Presented to the Division of Environmental and Biomolecular Systems  
and the Oregon Health & Science University  
School of Medicine  
in partial fulfillment of  
the requirements for the degree of

Doctor of Philosophy  
in  
Environmental Science and Engineering

May 2015

School of Medicine  
Oregon Health & Science University

---

CERTIFICATE OF APPROVAL

This is to certify that the PhD dissertation of

Amanda L. Mather

has been approved

---

Richard L. Johnson, PhD  
Research Advisor

---

Joseph A. Needoba, PhD  
Examination Committee Chair

---

Karen H. Watanabe, PhD  
Examination Committee Member

---

Paul G. Tratnyek, PhD  
Examination Committee Member

---

Anthony J. (Jim) Tesoriero, PhD  
U.S. Geological Survey  
External Examination Committee Member

# TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
ACKNOWLEDGEMENTS.....	x
ABSTRACT.....	xi
CHAPTER 1: Introduction.....	1
1.1 Stream Turbidity.....	2
1.2 Water Quality during Hydrologic Events.....	4
1.3 Forecasting Stream Water Quality.....	6
1.4 Dissertation Overview.....	7
CHAPTER 2: Quantitative Characterization of Stream Turbidity-Discharge Behavior using Event Loop Shape Modeling and Power Law Parameter Decorrelation.....	9
2.1 Abstract.....	9
2.2 Introduction.....	10
2.2.1 Background.....	12
2.2.2 Study Objectives.....	15
2.3 Methods.....	15
2.3.1 Loop Modeling.....	15
2.3.2 Parameter Decorrelation.....	17
2.3.3 Data and Hydrologic Event Identification.....	18
2.3.4 Study Sites.....	19
2.3.5 Model Fitting and Evaluation.....	21
2.4 Results.....	22
2.4.1 Event Turbidity Magnitude and Timing.....	22
2.4.2 Turbidity-Discharge Loop Modeling.....	25
2.4.3 Loop Model Parameters.....	28
2.5 Discussion and Conclusions.....	31
CHAPTER 3: Characterization and Exploratory Analysis of Stream Turbidity during 5928 Events from 110 U.S. Gages.....	36
3.1 Abstract.....	36

3.2	Introduction .....	37
3.3	Methods .....	38
3.3.1	Study Gages and Datasets .....	38
3.3.2	Turbidity-Discharge Event Modeling .....	41
3.3.3	Cluster Analysis .....	43
3.4	Results and Discussion .....	44
3.4.1	Model Fit .....	44
3.4.2	Power Law Decorrelation Scaling Factor .....	47
3.4.3	Model parameter $a_0$ .....	49
3.4.4	Model parameter $b$ .....	51
3.4.5	Model parameter $c$ .....	52
3.4.6	Model parameter $r$ .....	53
3.4.7	Gage Similarity Based on Median Parameter Values .....	54
3.5	Conclusions .....	57
CHAPTER 4: Event-based Prediction of Stream Turbidity using Regression and Classification Tree Approaches .....		59
4.1	Abstract .....	59
4.2	Introduction .....	60
4.3	Methods .....	63
4.3.1	Study Catchments .....	63
4.3.2	Event Turbidity Modeling .....	65
4.3.3	Event Characteristics .....	67
4.3.4	Turbidity Prediction Approaches .....	70
4.4	Results and Discussion .....	73
4.4.1	Study Events .....	73
4.4.2	Turbidity Prediction using Cluster Analysis and Classification Trees .....	75
4.4.3	Turbidity Prediction using Regression Analysis .....	80
4.4.4	Comparison of Turbidity Prediction Approaches .....	82
4.5	Conclusions .....	88
CHAPTER 5: Forecasting Turbidity during Streamflow Events for Two Mid-Atlantic U.S. Streams .....		90

5.1	Abstract .....	90
5.2	Introduction .....	91
5.2.1	Study Sites .....	92
5.3	Methods .....	93
5.3.1	Turbidity Forecast Input Data .....	93
5.3.2	Turbidity Forecasting .....	96
5.3.3	Evaluation of Forecasts .....	97
5.3.4	Uncertainty Intervals .....	98
5.4	Results and Discussion .....	99
5.4.1	Streamflow Events and Forecasts .....	99
5.4.2	Streamflow Forecast Errors .....	102
5.4.3	Turbidity Forecasts .....	104
5.4.4	Turbidity Forecast Errors .....	107
5.4.5	Effect of Streamflow Forecast Uncertainty on Turbidity Forecasts .....	110
5.5	Summary and Conclusions .....	113
	CHAPTER 6: Summary and Conclusions .....	115
	REFERENCES .....	121
	APPENDIX A: Supporting Information for Chapter 2 .....	132
	APPENDIX B: Supporting Information for Chapter 3 .....	139
	APPENDIX C: Supporting Information for Chapter 4 .....	150
	APPENDIX D: Supporting Information for Chapter 5 .....	157

# LIST OF TABLES

---

Table 2.1. Loop models. ....	16
Table 2.2. Study gages. ....	21
Table 3.1. Summary of the catchment characteristics for the 110 study gages. ....	40
Table 3.2. Spearman’s rank correlation coefficients. ....	47
Table 3.3. Median of the gage median parameter values for each cluster. ....	56
Table 4.1. Study catchments and dataset details. ....	65
Table 4.2. Turbidity event model parameter values used for Power Law and Median Value prediction approaches. ....	75
Table 4.3. Turbidity event model parameter values used for the Cluster/Classification prediction approach. ....	78
Table 4.4. Regression equations for turbidity event model parameters. ....	81
Table 5.1. Streamflow events used in this study. ....	99
Table 5.2. Regression equations to estimate event model parameters. ....	105
Table A 1. Study gages and catchment characteristics. ....	133
Table B 1. Gage data. ....	140

# LIST OF FIGURES

---

Figure 1.1. Three example hydrologic events from the Raritan River with similar hydrographs (blue) and different event turbidity (black). The resulting turbidity-discharge loops are shown below the time series plots. Data points are only shown on the loop plots and the loop direction is as shown. (figure duplicated here from Chapter 4, Figure 4.1). ..4	..4
Figure 2.1. Map of gage locations (triangles) and catchment outlines (shaded areas), showing (a) catchment locations within the U.S., (b) gages and catchments relative to each other, and (c-d) gage and catchment detail..... 20	20
Figure 2.2. Scatterplots of event peak turbidity and peak discharge (scaled by drainage area) with individual gages shown on separate panels. Each data point corresponds to an individual event with the total number of events per gage ( $n$ ) indicated. Event markers indicate season in which the event occurred; Winter (Jan/Feb/Mar), Spring (Apr/May/June), Summer (Jul/Aug/Sep) and Fall (Oct/Nov/Dec)..... 23	23
Figure 2.3. (a) Boxplots of event centroid-to-centroid discharge-turbidity lag ( $t_{cc}$ ) by gage; negative values indicate the centroid of turbidity leads the centroid of discharge. Boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside of the upper and lower whiskers. (b) Gage mean $t_{cc}$ as related to catchment drainage area. Each point represents a gage..... 24	24
Figure 2.4. Example event data and model fits. (a) Raritan R. (G16) event #20, (b) French C. (G9) event #100, (c) Pamunkey R. (G17) event #57, (d) WB Brandywine C. (G6) event #34, and (e) Smith C. (G13) event #20. Bold outline indicates the best model(s) for each event shown. .... 26	26
Figure 2.5. Boxplot of $NSE$ for each model (M1-M5) by gage. Each cluster of five boxes corresponds to models M1-M5, in order from left to right, for the gage shown below the cluster. Boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside of the upper and lower whiskers. .... 27	27
Figure 2.6. Scatterplots by gage of (a) $b_4$ (–) vs. $a_4 \left( FNU(km^2(s/m^3))^{b_4} \right)$ and (b) $b_4$ (–) vs. $a_{4,0}$ ( $FNU$ ). Each marker represents an individual hydrologic event..... 29	29
Figure 2.7. Calculated values of $Q_{n,0}$ for each model and gage. Each cluster of 5 bars corresponds to models M1-M5, in order from left to right..... 29	29
Figure 2.8. M4 model parameters $a_{4,0}$ , $a_4$ , and $b_4$ by month for gages G16 and G9. Data points are individual events and the trendline is a LOWESS smoother. .... 30	30

Figure 3.1. Location of 110 gages (triangles) and corresponding catchment outlines (shaded areas).....	38
Figure 3.2. Boxplots of event $NSE$ , $a_0$ , $b$ , $c$ and $r$ values by gage. Gages are arranged in order of increasing drainage area from left to right. Boxplot boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to the nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside the upper and lower whiskers (outliers). Some outlier events are not shown for scaling reasons; however, the number of events not shown is identified. ....	45
Figure 3.3. Median event $NSE$ by gage. Key to the right of map shows a boxplot of gage median event $NSE$ and color key for mapped markers. Gradient on color key extends from lower (blue) to upper (red) whisker, centered on the median (green). Low outliers are blue and high outliers are red. Mapped marker shape indicates dominant catchment land cover: forest (triangle), urban (square), agriculture (circle) or mixed (diamond). A particular land cover category is considered dominant if it represents at least 60% of the catchment area. ....	46
Figure 3.4. Decorrelation scaling factor, $Q_0$ , for each gage. Key to the right of map shows a boxplot of $Q_0$ and color key for mapped markers. For full description of marker colors and shapes, see Figure 3.3 caption. ....	48
Figure 3.5. Median model parameter values for all events from each gage. Key to the right of each map shows a boxplot of gage median parameter values and color key for mapped markers. For full description of marker colors and shapes, see Figure 3.3 caption. ....	50
Figure 3.6. Study gages clustered by gage median model parameter values.....	55
Figure 4.1. Three example hydrologic events from the Raritan River with similar hydrographs (blue) and different event turbidity (black). The resulting turbidity-discharge loops are shown below the time series plots. Data points are only shown on the loop plots and the loop direction is as shown. ....	61
Figure 4.2. The three study catchments are located in the Mid-Atlantic hydrologic region of the U.S.; red triangles indicate location of gaging stations. Catchments are all shown at the same scale. ....	64
Figure 4.3. Boxplots of best-fit event model parameter values for the events in the training datasets: $a_0$ ( $FNU$ ), $b$ (-), $c$ ( $FNU(km^2s)(s/m^3)$ ), $r$ ( $km^2/m^3$ ) and $NSE$ (-). Boxplot boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to the nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside the upper and lower whiskers (outliers); boxplot colors indicate stream. ....	74

Figure 4.4. Dendrogram showing cluster analysis results and the events associated with each cluster (shape group) for Raritan River. The events are shown as turbidity vs. discharge loops and are colored by shape group. The small plots of individual events all have the same axis scaling. ....	77
Figure 4.5. Classification Tree for Raritan River. Terminal node (circles) class identity colored by shape group, as shown in Figure 4.4. Membership for training events shown with rectangles representing individual events under each terminal node. Only 69 of the total 70 events are shown due to one event at the <i>PRCP</i> branch having missing precipitation data. ....	79
Figure 4.6. Success Rate for training and validation events using each prediction approach. ....	83
Figure 4.7. Success rate by prediction approach for the full dataset (training and validation combined). (a) Relationship with event peak specific discharge ( $Q_p$ ); events are binned into 0.3-intervals on the x-axis, centered on the data markers shown. (b) Relationship with season event occurred; Winter is Jan/Feb/Mar, Spring is Apr/May/June, Summer is Jul/Aug/Sep and Fall is Oct/Nov/Dec. Numbers in gray bars indicate the number of individual events represented by each point below. ....	85
Figure 4.8. Taylor diagrams of the validation events using all four prediction approaches. Each data point represents an event predicted using the approach indicated by the marker color. Left diagrams for each stream show all validation events, middle diagrams are two example events from each stream and time series plots on the right correspond to the two example events shown in the middle. Square and diamond markers on left Taylor diagrams indicate the events used as examples in middle and right. RMSD is the root-mean-square error and OBS is the observed event turbidity reference point. ....	86
Figure 5.1. NWS streamflow forecasts (red) for events A-F and observed specific discharge (black) from 3/19/2014 – 5/21/2014 for Raritan River (top) and Brandywine Creek (bottom). ....	101
Figure 5.2. Mean error ( $ME$ ) and mean absolute error ( $MAE$ ) in NWS streamflow forecasts (events A-F) for various lead times. Forecast lead times are binned at nearest 6-hour interval (i.e., a point at 12 hours is all forecasts within: 9 hrs < lead time $\leq$ 15 hrs). Persistence reference represents the observed specific discharge at the forecast issue time repeated for all lead times in the forecast period. ....	103
Figure 5.3. Turbidity forecasts (orange) for events A-F and observed turbidity (black) from 3/19/2014 – 5/21/2014 for Raritan River (top) and Brandywine Creek (bottom). ....	106
Figure 5.4. Mean error ( $ME$ ) and mean absolute error ( $MAE$ ) in Raritan River (left) and Brandywine Creek (right) turbidity forecasts for various lead times. Forecast lead times are binned at nearest 6-hour interval (i.e., a point at 12 hours is all forecasts with: 9 hrs <	

lead time  $\leq 15$  hrs). Persistence reference represents the observed turbidity at the forecast issue time repeated for all lead times in the forecast period. .... 107

Figure 5.5. Turbidity forecast mean absolute error (*MAE*) for several input data scenarios. “Perfect Forecast” represents a streamflow forecast that perfectly matches the observations at each original forecast time point. “Best Parameters” represents optimal values of turbidity model parameters  $a_0$ ,  $b$ ,  $c$  and  $r$  for each event. .... 109

Figure 5.6. Relative forecast discharge error for Raritan River (left) and Brandywine Creek (right) as a function of forecast specific discharge for individual forecast data points over events A-F. Empirical uncertainty intervals (70% dark gray, 90% light grey) shaded for  $\sim 1.6X$  (2 dB) x-axis bins. Uncertainty intervals for forecast specific discharge greater than  $\sim 0.2 (m^3/s)/km^2$  were not calculated because data points beyond this threshold were sparse and nearly exclusively from a single event (event E). .... 111

Figure 5.7. Three example forecasts from Raritan River event D. Left panels show NWS streamflow forecast and empirical uncertainty intervals. Right panels show turbidity forecasts and uncertainty intervals produced from streamflow forecast uncertainty. Observed (solid black), observed but not yet known at forecast issue time (dotted black), streamflow forecast (red), turbidity forecast (orange), 90% uncertainty interval (light gray), 70% uncertainty interval (dark grey) and forecast issue time (green)..... 112

Figure B 1. Number of events per gage used in this study. Histogram displays distribution of number of events per gage and shows color key for map of gage event totals..... 148

Figure C 1. Boxplots of best-fit event model parameter values for the events in the validation datasets:  $a_0$  (*FNU*),  $b$  (-),  $c$  (*FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)*),  $r$  (*km<sup>2</sup>/m<sup>3</sup>*) and *NSE* (-). Boxplot boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to the nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside the upper and lower whiskers (outliers); boxplot colors indicate stream. .... 150

Figure C 2. Dendrogram showing cluster analysis results and the events associated with each cluster (shape group) for Brandywine Creek. The events are shown as turbidity vs. discharge loops and are colored by shape group. The small plots of individual events all have the same axis scaling. .... 151

Figure C 3. Dendrogram showing cluster analysis results and the events associated with each cluster (shape group) for S.B. Potomac River. The events are shown as turbidity vs. discharge loops and are colored by shape group. The small plots of individual events all have the same axis scaling. .... 152

Figure C 4. Classification Tree for Brandywine Creek. Terminal node (circles) class identity colored by shape group, as shown in Figure C 2. Membership for training events

shown with rectangles representing individual events under each terminal node. Only 59 of the total 62 events are shown due to three events having missing characteristic data for a branch the event encounters. .... 153

Figure C 5. Classification Tree for S.B. Potomac River. Terminal node (circles) class identity colored by shape group, as shown in Figure C 3. Membership for training events shown with rectangles representing individual events under each terminal node. .... 154

Figure C 6. Comparison of predicted parameter values and fitted parameter values for Raritan River. Units are:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)) and  $r$  (km<sup>2</sup>/m<sup>3</sup>). . 155

Figure C 7. Comparison of predicted parameter values and fitted parameter values for Brandywine Creek. Units are:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)) and  $r$  (km<sup>2</sup>/m<sup>3</sup>). ..... 155

Figure C 8. Comparison of predicted parameter values and fitted parameter values for S.B. Potomac River. Units are:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)) and  $r$  (km<sup>2</sup>/m<sup>3</sup>). ..... 156

Figure D 1. Flow diagram showing turbidity model calibration and operational forecasting. .... 158

Figure D 2. Diagram showing calculation of  $V(t)$ , which is used in the turbidity model (equation (5.1) in Chapter 5). ..... 159

Figure D 3. NWS streamflow forecasts for Raritan River events A-F. .... 163

Figure D 4. NWS streamflow forecasts for Brandywine Creek events A-F. .... 164

Figure D 5. Turbidity forecasts for Raritan River events A-F. .... 165

Figure D 6. Turbidity forecasts for Brandywine Creek events A-F. .... 166

# ACKNOWLEDGEMENTS

---

I would like to express my sincere appreciation to my advisor, Dr. Rick Johnson, for his continuous encouragement, patient guidance and steadfast support, which have made it possible to accomplish this research. I would also like to thank Dr. Paul Tratnyek for providing me with the original opportunity to start down the path towards this degree. There have been many others over the years that have helped me along the way, and I would particularly like to thank Dr. Paul Capel and Dr. Chris Green. Finally, I must express my deepest gratitude to my best friend and beloved husband, Joel, for his unfailing encouragement and support throughout this endeavor.

# ABSTRACT

---

Streams and rivers provide many important functions and services for society, of which many depend not only on the quantity of water available, but also the stream water quality. While quantity (i.e., streamflow) is routinely forecast for many U.S. streams several days into the future, the forecasting of stream water quality remains relatively limited. This dissertation utilizes the extensive existing networks of observed and forecast streamflow, precipitation and air temperature as the basis for practical forecasting of stream turbidity for three days into the future. To accomplish this, an improved power-law based model for turbidity during hydrologic events was developed and applied to ~6000 events from 110 U.S. gages. The resulting event model parameters were examined in the context of catchment characteristics and event characteristics (e.g., hydrologic, meteorologic and antecedent moisture conditions) in order to understand the variability of turbidity response between streams and for different events within a particular stream. The results indicated that gage median parameter values were mainly correlated with catchment land cover and baseflow index, while the individual-event parameter values for a particular stream were largely correlated with the antecedent moisture conditions preceding the event. These analyses were, in part, facilitated by the use of a power-law parameter decorrelation methodology, which clarified the relationships between the power-law coefficient and the catchment and event characteristics. The information gained from these analyses was used to develop regression equations to forecast turbidity model parameters based on the event characteristics. Using archived streamflow and meteorologic forecast inputs and gage-specific regression equations, turbidity forecasts were made for events from two mid-sized streams. The turbidity forecast errors were

examined, and the results indicated that the forecasts were “useful” compared to a persistence reference. The uncertainty in the turbidity forecasts due to uncertainty in the streamflow forecasts was also explored. Overall, the results indicated that practical and useful turbidity forecasts can be produced from currently-available observed and forecast inputs. The widespread availability of these inputs for streams across the U.S., and the value of turbidity as a surrogate for other water quality constituents, suggests that the forecasting demonstrated here could be implemented for many other streams and for additional water quality constituents.

# CHAPTER 1

---

## Introduction

Rivers, streams and creeks (herein referred to simply as “streams”) provide many important functions and services, such as transportation, recreation, power generation and a source of drinking water. The quantity of water flowing in streams (i.e., “streamflow”) can change on a daily basis and is frequently forecast for several days into the future to support these uses, most notably by the National Weather Service (NWS) in the U.S. For many of these functions and services, we also care a great deal about the water quality (i.e., the physical, chemical and biological characteristics). However, the extension of routine forecasting to include stream water quality is not yet commonplace. This dissertation demonstrates practical stream water quality forecasting, using turbidity as an example.

## 1.1 Stream Turbidity

Turbidity is an expression of the optical clarity of water. It is an aggregate measurement in which any suspended or dissolved material that causes light to be scattered or absorbed may be expressed [Gray and Glysson, 2003]. Stream turbidity is influenced by inorganic (e.g., clay and silt) and organic (e.g., algae and plankton) suspended solids, and dissolved substances (e.g., dyes and organic matter) [U.S. EPA, 1999; Anderson, 2005]. Because of the diverse sources of turbidity-causing materials, the dominant material(s) contributing to stream turbidity levels can vary over time and between streams [U.S. EPA, 1999]. The U.S. Geological Survey commonly measures turbidity using the ISO 7027 standard.<sup>1</sup> The near-IR light source specified by this method is minimally absorbed by colored molecules, and the method is therefore quite resistant to influences of dissolved materials [Anderson, 2005]. All of the turbidity data used for the research presented in this dissertation was collected by the U.S. Geological Survey using ISO 7027 and has units of Formazin Nephelometric Units (*FNU*).

High levels of turbidity can have negative consequences for the functions and services a stream provides. Almost half of the U.S. population relies on streams as a source of drinking water [Wickham *et al.*, 2011]. The turbidity level of source water from stream intakes has an impact on subsequent treatment costs. It has been estimated that

---

<sup>1</sup> Two well-known water quality standard methods and device designs used to measure stream turbidity are EPA Method 180.1 and ISO 7027. The primary difference between these two methods is related to the light source. For EPA Method 180.1, the light source must be a tungsten lamp with broad spectrum emissions peaking between 400-680 nm. Whereas for the ISO 7027 method, the light source is monochromatic, either an LED or filtered tungsten lamp, emitting near-IR light with a wavelength of 860 nm and a bandwidth of 60 nm. Measurements made using EPA Method 180.1 are assigned units of Nephelometric Turbidity Units (*NTU*), while ISO 7027 method measurements are given units of Formazin Nephelometric Units (*FNU*). In both cases, devices are calibrated using a formazin standard; however, because of the different light sources, direct comparison of measurements with different turbidity units is not recommended. [Anderson, 2005]

public drinking water treatment systems bear 0.25% increase in treatment cost per 1% increase in turbidity [*Dearmont et al.*, 1998]. Additionally, episodes of extreme turbidity may require switching to alternative water sources [e.g., *Portland Water Bureau*, 2011] or temporary shutdown of treatment systems [e.g., *Duncan and Grant*, 2003]. As a consequence, stream turbidity is important from an economic and water supply perspective.

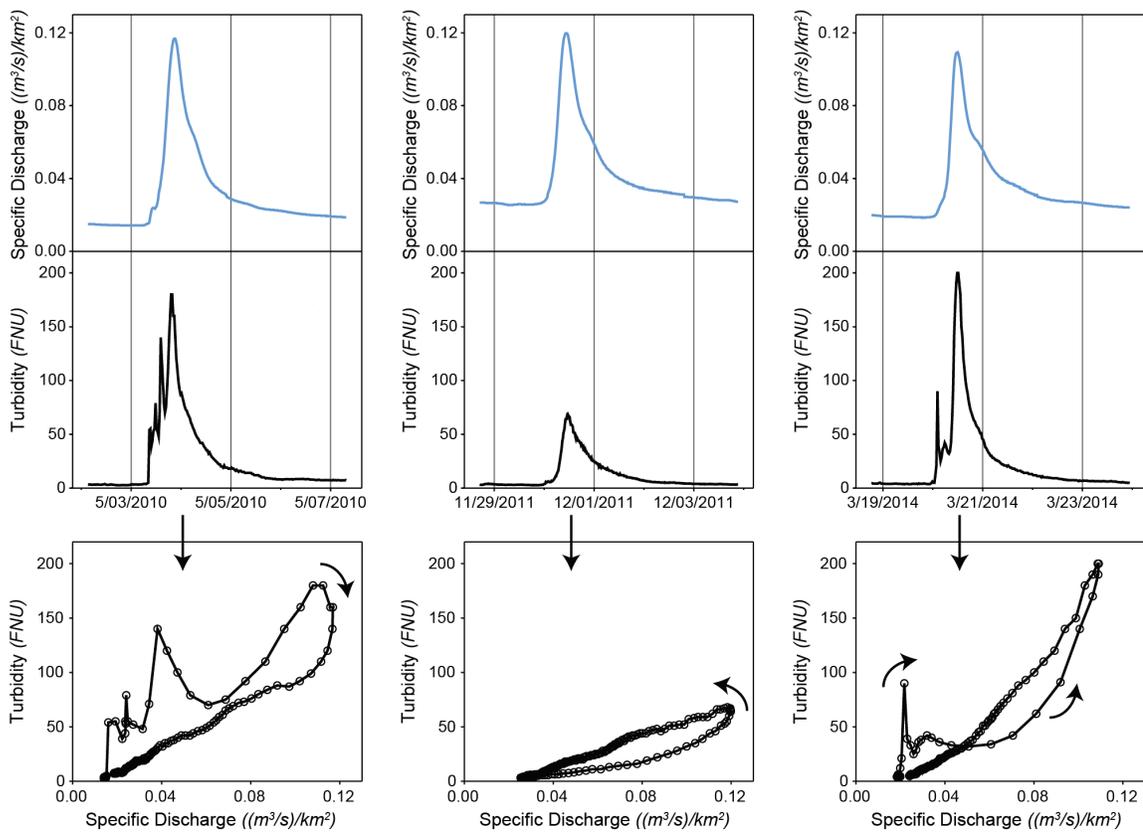
Perhaps more importantly, turbidity is a remarkably versatile stream water quality surrogate—often enabling the estimation of many other water quality constituents that are not easily measured or are expensive to measure, including:

- Suspended sediment [e.g., *Gray and Glysson*, 2003]
- Agricultural chemicals [e.g., *Hickman*, 2004]
- Pathogens [e.g., *Christensen et al.*, 2000]
- Heavy metals [e.g., *Miller*, 1997]

Because turbidity is relatively simple and inexpensive to monitor continuously in-stream [*Davies-Colley and Smith*, 2001], it is an important constituent in water quality monitoring and regulation [*Gray and Glysson*, 2003] and an ideal candidate for water quality forecasting.

## 1.2 Water Quality during Hydrologic Events

Stream water quality varies over time—some changes occurring slowly over the scale of years (e.g., due to changing climate or watershed development) while other variations occur over much shorter time periods and may even be cyclical in character (e.g., seasonal patterns or diurnal variations). The most significant changes in short-term water quality (i.e., at the scale of days), however, are in response to hydrologic events (Figure 1.1) [e.g., *Kirchner et al.*, 2004].



**Figure 1.1. Three example hydrologic events from the Raritan River with similar hydrographs (blue) and different event turbidity (black). The resulting turbidity-discharge loops are shown below the time series plots. Data points are only shown on the loop plots and the loop direction is as shown. (figure duplicated here from Chapter 4, Figure 4.1).**

Hydrologic events, also referred to as discharge or streamflow events, occur when a stream responds to a period of precipitation and are typically characterized by a rise and subsequent retreat of streamflow levels (i.e., a hydrograph). Most stream water quality constituents experience significant changes in “concentration” during hydrologic events—either increasing (i.e., flushing effect) or decreasing (i.e., dilution effect) [e.g., *Chang and Carlson, 2005*]. Hydrologic events (subsequently referred to sometimes simply as “events”) therefore represent an important timescale of interest for stream water quality.

For turbidity, as with many other water quality constituents, the relationship between “concentration” and discharge typically varies throughout an event [e.g., *Rose, 2003; Lawler et al., 2006; Megnounif et al., 2013; Mukundan et al., 2013; Siwek et al., 2013*]. This means that when event turbidity is plotted as a function of stream discharge, the resulting shape is a hysteretic loop rather than a straight line (Figure 1.1, lower panels). Additionally, as in the three example events shown in Figure 1.1, turbidity behavior varies between different events, even on the same stream and even when the hydrograph is nearly identical. Almost certainly, the underlying reasons for these differences have their origin in some combination of catchment characteristics (e.g., slope, land cover, baseflow index, etc.), event characteristics (e.g., peak discharge, precipitation, antecedent moisture conditions, etc.) and season. However, to date this has not been demonstrated. Analysis of the catchment, event and seasonal characteristics that influence event turbidity, and the development of predictive relationships from these characteristics, is an important component of stream turbidity forecasting.

### **1.3 Forecasting Stream Water Quality**

Streamflow forecasts are important aspects of modern water management, and the extension of routine forecasting to include water quality is equally important. Widespread sub-hourly monitoring of turbidity and many other water quality constituents now exists for many streams across the U.S., and as a result, we know more than ever about how water quality changes over the scale of hours. But in most cases, without assuming tomorrow will be the same as today, we do not know the answer to the question: What will the water quality be 12 hours in the future? Or two days in the future? Knowing the answer to these questions would allow informed and economical operational decisions to be made today, in advance and in response to anticipated water quality tomorrow.

Understanding short-term variability in stream water quality, and being able to anticipate influent water quality at intakes, may be critical for management of drinking water systems in the future. Climate change is expected to result in alteration of water distribution, availability and quality [*Levin et al.*, 2002; *IPCC*, 2014]. This may further stress already fully-utilized water supplies and prompt the need for greater day-to-day adaptation to changes in the water quality of drinking water sources [*IPCC*, 2014]. Our demand for drinking water is only going to grow, both here in the U.S. and elsewhere [*WWAP (World Water Assessment Programme)*, 2012]. Forecasting stream water quality is an important advancement that would assist drinking water providers in continuing to provide safe and sufficient water in a future of changing stream water quality.

Forecasting stream turbidity will require the use of both recently observed data and forecast input (i.e., streamflow, precipitation and air temperature); however, these inputs are already available for many streams across the U.S. The NWS currently

forecasts streamflow for several days into the future on a routine-basis (i.e., daily) for 1500+ stream locations and makes forecasts for an additional ~1500 locations during times of hydrologic activity [*NOAA National Weather Service*, n.d.]. Couple this with the comprehensive forecasting of precipitation and air temperature across the country, and the result is a complete input data framework on which to base water quality forecasts.

## **1.4 Dissertation Overview**

The most commonly used approach for modeling turbidity based on stream discharge is a simple power law; however, this approach has many shortcomings. As a result, Chapter 2 describes an improved event model to quantitatively characterize stream turbidity during hydrologic events. Chapter 2 also tackles several long-standing issues with power-law based models through the use of a decorrelation and unit simplification scheme.

In Chapter 3 the model developed in Chapter 2 is applied to ~6000 events from 110 stream gages across the U.S. The motivation for the research in Chapter 3 is twofold—it permits a large-sample evaluation of the turbidity event model developed in Chapter 2, and it allows an exploration of how the turbidity model parameters vary between streams and how this relates to the observable characteristics of the catchment. Chapter 3 verifies the applicability of the event model to streams with diverse characteristics and offers a peek into the potential prediction of event turbidity model parameters for streams which do not currently have turbidity monitoring (i.e., “ungaged”).

Chapter 4 addresses the fundamental question of how to predict stream turbidity from the hydrologic, meteorologic and antecedent moisture characteristics of the event.

Stream turbidity for three streams is predicted using several approaches, including multiple linear regression and a classification tree approach. The success rates of the prediction approaches are compared and their utility within an operational forecasting context is considered. The event characteristics which are the most useful for predicting stream turbidity are also discussed.

Chapter 5 demonstrates stream turbidity forecasting by producing forecasts for six events from two streams (118 individual forecasts). The turbidity forecasts have 3-day duration with 6-hr resolution (to match discharge forecasts) and are based on a combination of observed and forecasted streamflow, precipitation and temperature. The turbidity forecast errors, uncertainty due to streamflow forecast uncertainty and “usefulness” compared to a persistence reference (i.e., assuming the future will be the same as the present) are examined. Empirical uncertainty intervals are presented along with the forecasts, providing the opportunity to visually observe the uncertainty in turbidity forecasts resulting from streamflow forecast uncertainty.

Finally, Chapter 6 provides a summary and some overall conclusions for the body of the research presented in Chapters 2-5. Several chapters have associated supporting information which can be found in the Appendices at the end of the dissertation.

# CHAPTER 2

---

## **Quantitative Characterization of Stream Turbidity-Discharge Behavior using Event Loop Shape Modeling and Power Law Parameter Decorrelation<sup>2</sup>**

*Amanda L. Mather and Richard L. Johnson*

### **2.1 Abstract**

Turbidity behavior in streams is a complex and dynamic function of both source material supply and event-driven transport. While the primary controls on turbidity behavior across time and space are still not fully understood, recent increases in the availability of high temporal resolution, co-located stream turbidity and discharge data provide an opportunity for more-detailed analysis. Here we examine methods to quantitatively characterize event responses by modeling the shape of turbidity-discharge hysteresis loops. A total of 1559 events from 20 gages in the Mid-Atlantic region of the U.S. were modeled using both previously-reported and new models combining elements

---

<sup>2</sup> Reproduced (with edits) with permission from Mather, A., and R. Johnson (2014), Quantitative characterization of stream turbidity-discharge behavior using event loop shape modeling and power law parameter decorrelation, *Water Resour. Res.*, 50(10), 7766–7779, doi:10.1002/2014WR015417. Copyright 2014. John Wiley & Sons, Inc.

of existing models. The results suggest that a more general power law based model, utilizing both a discharge rate of change term and a “supply” term, allows characterization of a wide range of simple and complex events. Additionally, this study explores a decorrelation approach to address the strong correlation frequently observed between the power law model coefficient ( $a$ ) and exponent ( $b$ ), with the goal of exposing the underlying behavior of each parameter individually. An examination of seasonal parameter behavior suggests that this approach may facilitate greater physically based interpretation of the power law coefficient. The power law parameter decorrelation strategy and the loop models examined here provide a step toward the larger goal of understanding the physical controls on turbidity-discharge hysteretic behavior.

## 2.2 Introduction

Despite the recognized importance of stream turbidity on aquatic life [e.g., *Henley et al.*, 2000], contaminant transport [e.g., *Juracic et al.*, 1986] and drinking water quality [e.g., *U.S. EPA*, 1999], the primary controls on its behavior during a hydrologic event are still not fully understood. During a single hydrologic event, the point at which the maximum event turbidity is reached may or may not coincide with the timing of the peak discharge. As a result, when event turbidity is plotted against event discharge it frequently forms a hysteretic loop (Figure 1.1) [*Seeger et al.*, 2004; *Lawler et al.*, 2006; *Eder et al.*, 2010; *Mukundan et al.*, 2013b]. The loop may be clockwise or counter-clockwise, depending on whether discharge or turbidity reaches a maximum value first, and on differences in the shapes of the event time series of discharge and turbidity. More complex variations in which the loop direction reverses during the event (e.g., figure-eight) have also been observed. There is considerable variability in the dominant loop

behavior observed between catchments as well as the event-to-event variability within individual catchments [Seeger *et al.*, 2004; Lawler *et al.*, 2006; Eder *et al.*, 2010; Mukundan *et al.*, 2013b]. The specific features of each individual event loop are believed to be a direct reflection of the catchment and hydrologic processes and their relative importance [Seeger *et al.*, 2004; Mukundan *et al.*, 2013b]. However with turbidity, as well as other water quality constituents [e.g., Rose, 2003; Megnounif *et al.*, 2013; Siwek *et al.*, 2013], we have yet to decipher the full extent of the information contained within the shape of hysteresis loops.

Turbidity behavior in streams is a complex and dynamic function of source material supply and event-driven transport [VanSickle and Beschta, 1983; Brasington and Richards, 2000; Zhang and Summer, 2002]. The supply of turbidity causing material within a catchment varies in amount [Asselman, 1999] and location [Duvert *et al.*, 2010; Wilson *et al.*, 2012], and it is believed to have a strong influence on the turbidity-discharge loop shape. The shapes of turbidity and/or suspended sediment hysteresis loops have been widely scrutinized in attempts to understand the spatial source material distribution and the catchment geomorphic processes involved in its transport [e.g., Jansson, 2002; Lawler *et al.*, 2006; López-Tarazón *et al.*, 2009; Smith and Dragovich, 2009; Vongvixay *et al.*, 2010]. From a watershed management standpoint, the identification of sources is a critical initial step towards subsequent design and implementation of management and control strategies [Walling, 2005].

In addition to supply, event turbidity response is affected by energy and flow characteristics. The amount of energy available for transport varies in time throughout a hydrologic event and is a factor contributing to looped behavior [Zhang and Summer,

2002]. Furthermore, in much the same way that river flow shows different characteristics at various scales (headwater to large rivers) [*Rinaldo et al.*, 1991; *Saco and Kumar*, 2002; *Merz et al.*, 2009], catchment turbidity response is a combined result of hydrologic and geomorphologic processes that dominate at one scale, but may not be visible in the response at a different catchment scale [*Eder et al.*, 2014].

*Williams* [1989] put forth a detailed record of the different hysteresis loop shapes that have been observed, characteristic features and postulated reasons for appearance within the context of suspended sediment. Five classes of loop behavior were outlined, including (1) straight (no loop), (2) clockwise loop, (3) counterclockwise loop, (4) straight plus a loop and (5) figure-eight. Within each of these classes exists a wide range of secondary characteristics, such as curvature (i.e., asymmetry) and the degree of loop openness. *Williams* [1989] and other prior work related to suspended sediment hysteresis is very relevant to turbidity [e.g., *Terajima et al.*, 1997; *Nistor and Church*, 2005]. The strong correlation between measured turbidity and suspended sediment concentration is widely leveraged in suspended sediment research and, in fact, turbidity is by far the most commonly used surrogate [*Gray and Glysson*, 2003; *Jones et al.*, 2011].

### **2.2.1 Background**

Several models to characterize individual event water quality have been developed. Empirical rating curves are by far the most widely used approach to model individual event concentration as a function of discharge. Simple rating curves are based on the assumption that concentration,  $C(t)$ , is a unique function of discharge,  $Q(t)$ . (In our case, turbidity is considered a “concentration”.) This relationship has been in use at

least since *Campbell and Bauder* [1940] and often takes the form of a power law function, where  $a$  and  $b$  are fitting parameters.

$$C(t) = aQ(t)^b \quad (2.1)$$

Parameter  $b$  is dimensionless, and parameter  $a$  has dimensions that are dependent on  $b$ ,  $[M/L^3][T/L^3]^b$ , assuming dimensions of  $[M/L^3]$  and  $[L^3/T]$  for  $C(t)$  and  $Q(t)$ , respectively [*Syvitski et al.*, 2000].<sup>3</sup> Despite the widespread observation that concentration is not a unique function of discharge and that the power law model cannot represent looped behavior, basic rating curves are still commonly used.

*VanSickle and Beschta* [1983] recognized that, together with streamflow, sediment supply exerts a dominant control on catchment sediment yield patterns. They presented a modification of the basic rating curve (equation (2.2)) by the inclusion of a washout function,  $g[S(t)]$ , which is a function of sediment supply,  $S(t)$  (tracked using equation (2.4)). This allows for a looped relationship between discharge and concentration as the supply decreases (washes-out) throughout the event. Their chosen function for  $g[S(t)]$  was a simple capacitor-like relationship (equation (2.3)), although they noted that other functions could be used, and the exact form is probably not critical.

$$C(t) = aQ(t)^b \cdot g[S(t)] \quad (2.2)$$

$$\text{where, } g[S(t)] = p \cdot \exp\left(r \frac{S(t)}{S_0}\right) \quad (2.3)$$

$$\text{and, } \frac{dS(t)}{dt} = -Q(t) \cdot C(t) \quad (2.4)$$

---

<sup>3</sup> where M = mass, L = length and T = time.

This model contains five fitting parameters ( $a, b, p, r, S_0$ ) and involves iteratively solving for the initial sediment supply ( $S_0$ ) for each hydrologic event. Parameter  $a$  again has dimensions of  $[M/L^3][T/L^3]^b$ ,  $S_0$  has dimensions of  $[M]$ , and  $b, p$  and  $r$  are dimensionless. This model was further explored in several studies [e.g., *Asselman, 1999; Vongvixay et al., 2010*], but uncertainty in how to deal with the initial sediment supply term remains a problem.

Several studies have modeled stream hysteresis of dissolved solutes [*House and Warwick, 1998*], suspended sediment [*Krueger et al., 2009; Eder et al., 2010*] and sediment associated compounds [*Bowes et al., 2005*] by including an event model term related to the local slope of the hydrograph (first time-derivative of discharge). This falls in line with concept of “first-flush”, allowing rapid resuspension and transport of sediment on the rising limb of the hydrograph in response to flow and energy conditions. *Krueger et al. [2009]* utilized this concept for sediment-discharge hysteresis modeling using:

$$C(t) = aQ(t)^b + c \frac{dQ(t)}{dt} \quad (2.5)$$

in which  $a, b$ , and  $c$  are fitting parameters corresponding to the slope, curvature and size/direction of the hysteresis loop, respectively. Parameter  $c$  has dimensions of  $[M/L^3][T^2/L^3]$ . This model performed well for the small hysteresis loops in *Krueger et al. [2009]*, but was unable to fully capture the extent of rapid flushing at the onset of some events.

## **2.2.2 Study Objectives**

The number of locations where co-located discharge and high temporal resolution turbidity data are collected has increased in recent years. These data allow both a detailed analysis of each turbidity event and the comparison of the response characteristics of many different catchments across the U.S. The purpose of this study was to evaluate the skill of several existing loop models and to improve the performance by combining elements of existing models into a new model for individual event loop shape characterization for catchments of varying size and characteristics. To achieve this, hydrologic events and resulting turbidity-discharge loops were identified using data collected between 2007 and 2013 from 20 Mid-Atlantic catchments. These turbidity-discharge loops were fit to several event models with varying numbers of fitting parameters. Additionally, we set out to test the utility of a power law (rating curve) parameter decorrelation method that we hypothesized would assist in physical interpretation of fitting parameter values.

## **2.3 Methods**

### **2.3.1 Loop Modeling**

Existing loop models have generally used either a supply-based or a flow/energy-based approach to fit loop shapes; however, as discussed above, turbidity response is a combination of both these processes that varies by catchment and over time. This study utilizes both the local slope of the hydrograph and event supply dynamics to develop a loop model with more universal applicability and improved accuracy of fit. Table 2.1 shows the five models explored in this study. Model M1 is the basic power law rating curve (included here for comparison even though it cannot represent a loop), M2 is the *Krueger et al.* [2009] hydrograph slope-based model and M3 is an adaptation of the

*VanSickle and Beschta* [1983] supply-based model. Models M4 and M5 are new loop models introduced in this study. They include both slope and supply terms and differ only in the manner in which the supply term is applied.

Models M3, M4 and M5 contain an  $e^{-rV}$  term, in which  $r$  is the rate at which materials contributing to turbidity (i.e., the “supply pool”) is depleted and is constrained to be greater than zero. The  $e^{-rV}$  term decreases from a value of unity towards zero throughout the event in proportion to the volume of water discharged. It does not directly contain information on the magnitude of initial supply, although this could be added at the expense of an additional fitting parameter (as in the original *VanSickle and Beschta* [1983] model formulation).

**Table 2.1. Loop models.**

Model	Model Equation	$a$ - $b$ Decorrelated Model Equation
M1	$TURB(t) = a_1 Q(t)^{b_1}$	$TURB(t) = a_{1,0} \left( \frac{Q(t)}{Q_{1,0}} \right)^{b_1}$
M2	$TURB(t) = a_2 Q(t)^{b_2} + c_2 \frac{dQ(t)}{dt}$	$TURB(t) = a_{2,0} \left( \frac{Q(t)}{Q_{2,0}} \right)^{b_2} + c_2 \frac{dQ(t)}{dt}$
M3	$TURB(t) = a_3 Q(t)^{b_3} e^{-r_3 V(t)}$	$TURB(t) = a_{3,0} \left( \frac{Q(t)}{Q_{3,0}} \right)^{b_3} e^{-r_3 V(t)}$
M4	$TURB(t) = a_4 Q(t)^{b_4} e^{-r_4 V(t)} + c_4 \frac{dQ(t)}{dt}$	$TURB(t) = a_{4,0} \left( \frac{Q(t)}{Q_{4,0}} \right)^{b_4} e^{-r_4 V(t)} + c_4 \frac{dQ(t)}{dt}$
M5	$TURB(t) = \left( a_5 Q(t)^{b_5} + c_5 \frac{dQ(t)}{dt} \right) e^{-r_5 V(t)}$	$TURB(t) = \left( a_{5,0} \left( \frac{Q(t)}{Q_{5,0}} \right)^{b_5} + c_5 \frac{dQ(t)}{dt} \right) e^{-r_5 V(t)}$
	where, $V(t) = \int_0^t Q(t) dt \cong \sum_{i=0}^T Q_i \Delta t$	where, $a_{n,0} = a_n (Q_{n,0})^{b_n}$

Note:  $Q$  = discharge;  $TURB$  = turbidity;  $a_n, b_n, c_n, r_n$  = fitting parameters for model  $n$ ;  $Q_{n,0}$  = decorrelation scaling factor for model  $n$ ;  $a_{n,0}$  = decorrelated parameter  $a_n$ .

In this study, discharge is normalized to drainage area prior to model fitting, resulting in discharge units of  $(m/s)$  (here as  $((m^3/s)/km^2)$  for convenience), and turbidity is treated as a “concentration”, having units of  $(FNU)$ . The resulting units for  $a$ ,

$b$ ,  $c$  and  $r$  are  $(FNU(s/m)^b)$  (here as  $(FNU(km^2(s/m^3))^b)$ ), unitless,  $(FNU(s^2/m))$  and  $(m^{-1})$ , respectively.

### 2.3.2 Parameter Decorrelation

All models used in this study include a power law term of the form  $aQ^b$  and consequently, because  $Q$  is dimensional, the dimension of  $a$  is dependent on the value of  $b$ . In fact, an inherent correlation between  $a$  and  $b$  has been observed in many studies in the hydrologic literature [Thomas, 1988; Asselman, 2000; Syvitski et al., 2000] as well as for power law relationships in other fields (e.g., fatigue crack growth, [Cortie, 1991; Zilberstein, 1992]; blood viscosity [Hussain et al., 1999a, 1999b]). This correlation is problematic for identifying the underlying behavior of each parameter individually and limits the physical interpretation of parameter values.

Here we utilize the methods of Bergner and Zouhar [2000] to transform  $aQ^b$  into a dimensionally simpler form through the use of a  $Q$ -dimensioned scaling factor,  $Q_0$ . Our reasons for doing this are twofold—to remove the parameter-dependent dimensions of  $a$  and to decorrelate the values of  $a$  and  $b$ .

$$TURB = a_0 \left( \frac{Q}{Q_0} \right)^b \quad (2.6)$$

$$\text{where, } a_0 = a(Q_0)^b \quad (2.7)$$

Parameter  $a$ , originally having units of  $(FNU(km^2(s/m^3))^b)$ , is thus transformed into  $a_0$ , with units of  $(FNU)$ , which are independent of the value of  $b$ . The key aspect of this transformation that results in the decorrelation of  $a$  and  $b$  is choosing the value of the scaling factor,  $Q_0$ , such that the correlation between  $a_0$  and  $b$  will vanish ( $r = 0$ ). Using

the relationship for the correlation coefficient,  $r$ , *Bergner and Zouhar* [2000] derived an explicit equation to calculate the necessary value of  $Q_0$  (equation (2.8)). In our application of this approach,  $Q_0$  is calculated for each gage for each model using:

$$Q_0 = 10^{-\left(\frac{\sum_{j=1}^k (b_j - \bar{b}) \log(a_j / a_g)}{\sum_{j=1}^k (b_j - \bar{b})^2}\right)} \quad (2.8)$$

$$\text{where, } a_g = \left(\prod_{j=1}^k a_j\right)^{\frac{1}{k}} \quad (2.9)$$

$j$  is the gage event index,  $k$  is the total number of individual events for the gage,  $\bar{b}$  is the mean of the  $b$  values for the gage and  $a_g$  is the geometric mean of the  $a$  values for the gage. For each model, this requires first fitting all of the events using the untransformed model, then using the fitted values of  $a$  and  $b$  to calculate  $Q_0$ , which is then used to transform the values of  $a$  into  $a_0$ . The actual model fit to the observations and all other parameter values (including  $b$ ) remain unchanged. Table 2.1 shows the model equations following power law parameter decorrelation.

### 2.3.3 Data and Hydrologic Event Identification

This study utilizes discharge and turbidity data available from the U.S. Geological Survey (USGS, parameter codes 00060 and 63680, respectively). These datasets are typically recorded at 15-minute or 30-minute intervals. Discharge and turbidity data were aligned to a synchronous timescale and linear interpolation was used to fill small data gaps. The local slope of the hydrograph ( $dQ/dt$ ) was calculated using the central difference method (equation (2.10)) and smoothed using a Savitzky-Golay filter [*Savitzky and Golay, 1964*].

$$\frac{dQ_i}{dt} = \frac{Q_{i+1} - Q_{i-1}}{2\Delta t} \quad (2.10)$$

Hydrologic events were identified as peaks in the discharge time-series using an automated peak detector program within LabVIEW (2012 SP1, National Instruments). The event window associated with each peak was assumed to extend from  $0.5N$  days prior to the time of discharge peak to  $N$  days following the peak. The value of  $N$  was calculated according to equation (2.11) [Sloto and Crouse, 1996], where  $A$  is the drainage area in square miles and  $N$  is the duration of surface runoff in days.

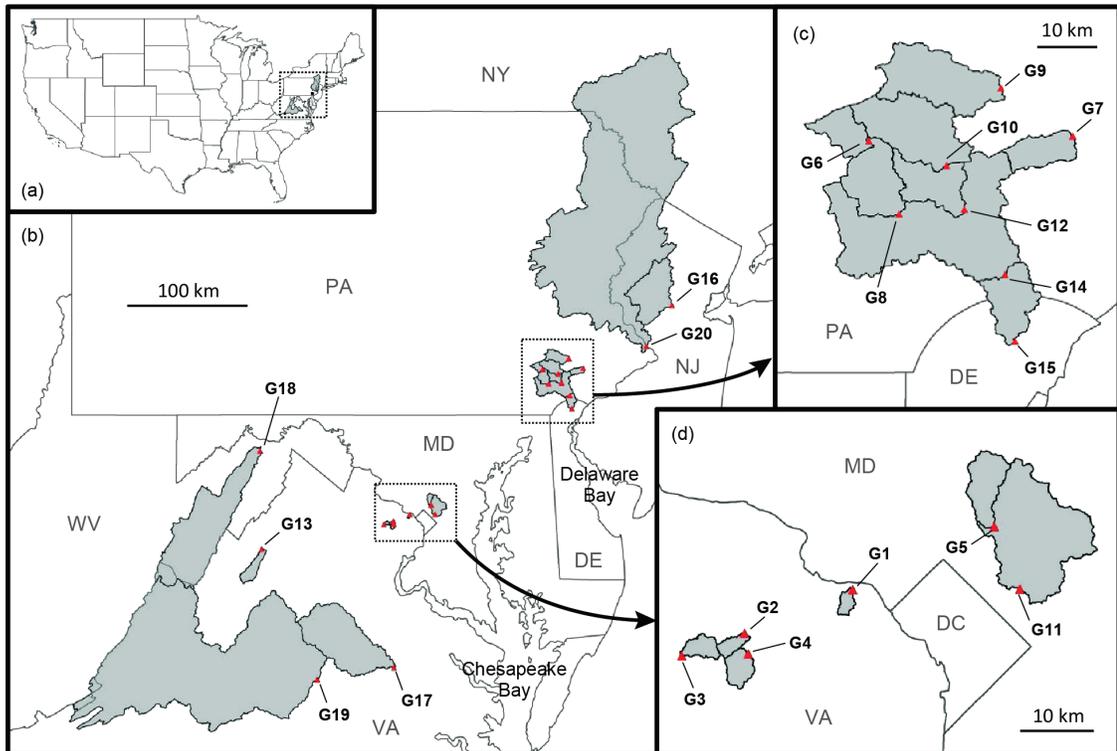
$$N = A^{0.2} \quad (2.11)$$

This window typically accommodated the range from prior baseflow conditions of discharge and turbidity, the full turbidity response (which often reached a maximum prior to the peak in discharge), and the return to baseflow conditions for both discharge and turbidity following the hydrologic event. The exception within this study is the Pamunkey River (gage G17, Table 2.2), which required a larger window ( $2 \times 0.5N = N$ ) on the leading side of the discharge peak to capture the full discharge and turbidity event response. Events were individually screened to eliminate those for which part of the data record was incomplete.

#### 2.3.4 Study Sites

Twenty gages from the Mid-Atlantic region of the United States were used in this study (Figure 2.1 and Table 2.2). The Mid-Atlantic region is characterized by several large rivers draining to estuaries (i.e., Chesapeake Bay and Delaware Bay) with adjacent dense urban areas. Also present are expansive rural areas of predominantly deciduous forests and areas of agricultural activity. Water quality issues, including suspended

sediment and excessive nutrients, are present in this region [Phillips, 2002]. Several of the catchments are nested and some are subject to flow regulation. The period of record for the discharge and turbidity data used for the 20 gages varied from approximately 2 to 5 years within the 2007-2013 time period (Table A 1). The gages were numbered from G1 to G20 in order of smallest to largest drainage area.



**Figure 2.1. Map of gage locations (triangles) and catchment outlines (shaded areas), showing (a) catchment locations within the U.S., (b) gages and catchments relative to each other, and (c-d) gage and catchment detail.**

There are some caveats of this set of gages that should be considered in the interpretation of the study results: (1) similar sized catchments tended to be grouped geographically (particularly the small catchments in this set) and (2) land cover characteristics tended to vary with catchment size (i.e., the smallest catchments are in

predominantly urban areas, while the largest catchments are mostly forested). Detailed information on topographic, land cover, soil and flow characteristics of these catchments can be found in the Table A 1 of Appendix A.

**Table 2.2. Study gages.**

Gage	USGS ID	Stream	Drainage Area ( $km^2$ )
G1	01646305	Dead Run	5.31
G2	01645762	S F Little Difficult Run	7.02
G3	01656903	Flatlick Br.	10.9
G4	01645704	Difficult Run	14.2
G5	01649190	Paint Br.	33.9
G6	01480300	W Br. Brandywine C.	48.4
G7	01473169	Valley C.	53.8
G8	01480617	W Br. Brandywine C.	143
G9	01472157	French C.	153
G10	01480700	E Br. Brandywine C.	157
G11	01649500	NE Br. Anacostia R.	188
G12	01480870	E Br. Brandywine C.	233
G13	01632900	Smith C.	242
G14	01481000	Brandywine C.	743
G15	01481500	Brandywine C.	813
G16	01400500	Raritan R.	1269
G17	01673000	Pamunkey R.	2799
G18	01608500	S Br. Potomac R.	3847
G19	02035000	James R.	16199
G20	01463500	Delaware R.	17553

Note: River, R.; Creek, C.; Branch, Br.;

### 2.3.5 Model Fitting and Evaluation

For each hydrologic event, models M1-M5 were fit independently using non-linear curve fitting with the trust-region dogleg (TRDL) algorithm and least-squares minimization (LabVIEW 2012 SP1, National Instruments). Values of  $r$  in M3-M5 were constrained to be greater than or equal to zero, but all other fitting parameters were free to take all values.

The primary metric of evaluation and comparison used in this study is the Nash-Sutcliffe Efficiency ( $NSE$ ) [Nash and Sutcliffe, 1970]:

$$NSE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.12)$$

where  $y_i$  is the observed value,  $\hat{y}_i$  is the predicted value and  $\bar{y}$  is the mean of the observed values for the event. The *NSE* is dimensionless and ranges from  $-\infty$  to 1, with values greater than zero indicating that the fit is better than the mean of the observations.

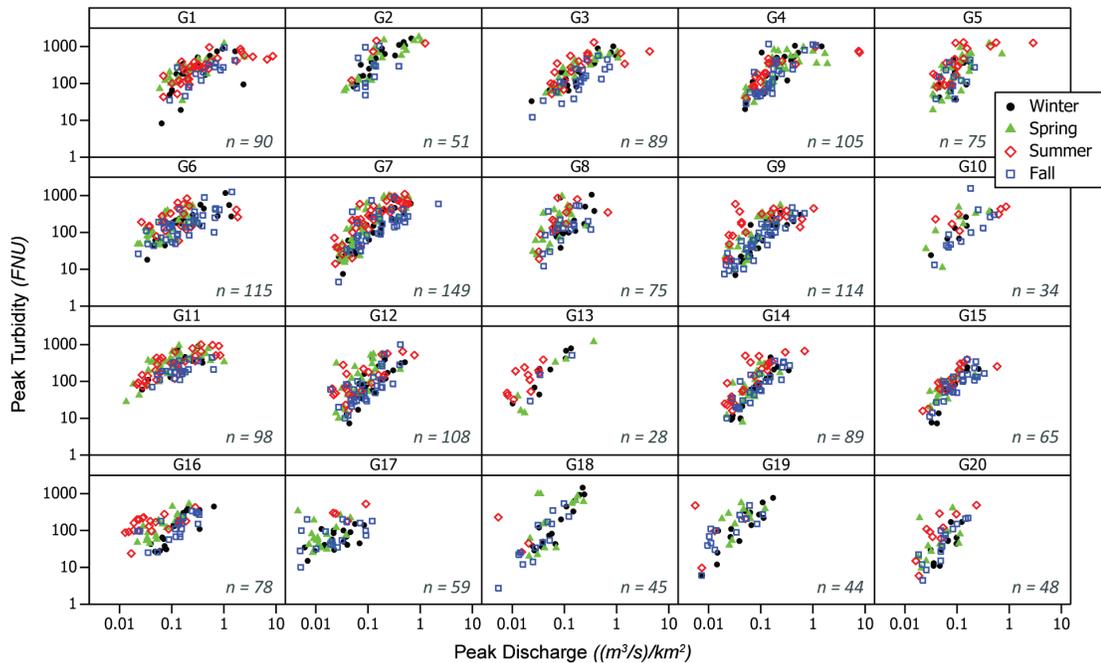
## 2.4 Results

### 2.4.1 Event Turbidity Magnitude and Timing

A total of 1559 turbidity events from the 20 gaged stations were identified and retained for further analysis. The numbers of events for individual gages ranged from 28 to 149 and are shown in Figure 2.2. Although the primary goal of this research was to examine looped turbidity-discharge behavior, the general character of the event turbidity behavior observed at these gages is summarized here using simple metrics of magnitude and timing.

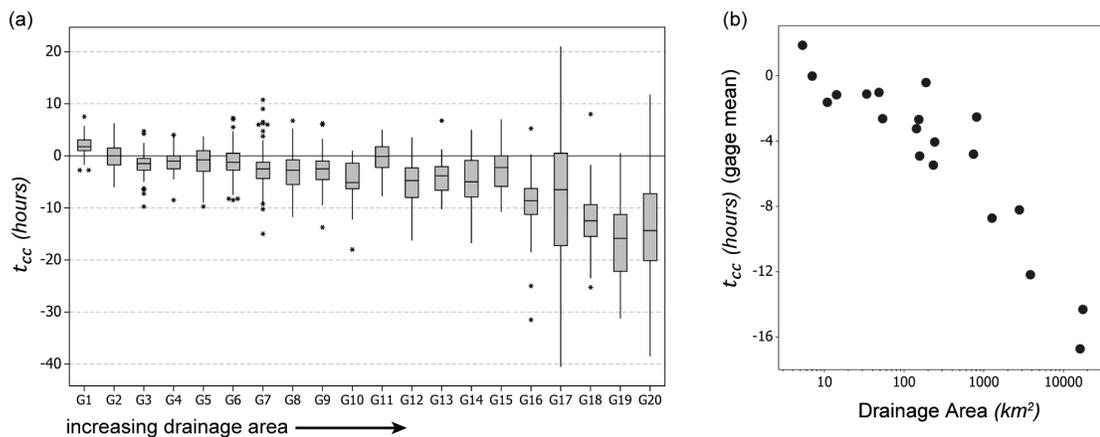
Individual event peak turbidity and peak discharge (scaled by drainage area) by gage and season are shown in Figure 2.2. Peak discharge generally ranged from  $\sim 0.01$ – $1 (m^3/s)/km^2$  and peak turbidity from  $\sim 10$ – $1000 FNU$ . The values of peak turbidity show some leveling off around  $1000 FNU$ , which is the approximate upper limit of the sensor. Thus, it is likely a few of the highest peak turbidities for several gages are underestimated. Figure 2.2 shows there are generally strong correlations between peak turbidity and peak discharge for all gages, with individual gages displaying more or less scatter in the relationship. The correlation is particularly good for Winter (Jan/Feb/Mar; mean correlation 0.82) and Fall (Oct/Nov/Dec; mean correlation 0.73) events, whereas events from Spring (Apr/May/June; mean correlation 0.70) and Summer (Jul/Aug/Sep;

mean correlation 0.68) are often more scattered, particularly towards higher peak turbidity values. This is consistent with the observations of *Lawler et al.* [2006], who found that the correlation between event peak turbidity and peak discharge shifted in slope and intercept through the course of a season (March – May). Seasonal changes in this relationship may be a reflection of differences in sediment sources, availability and transport, which may be related to seasonal changes in precipitation (e.g., intensity), antecedent moisture conditions and vegetation influences (e.g., attenuation and stabilization) [*Lana-Renault and Regiès, 2009; Giménez et al., 2012; Eder et al., 2014*]. Overall, the results shown in Figure 2.2 confirm that there is a strong relation between discharge and turbidity event magnitude for these gages and support the use of discharge in event turbidity modeling.



**Figure 2.2.** Scatterplots of event peak turbidity and peak discharge (scaled by drainage area) with individual gages shown on separate panels. Each data point corresponds to an individual event with the total number of events per gage ( $n$ ) indicated. Event markers indicate season in which the event occurred; Winter (Jan/Feb/Mar), Spring (Apr/May/Jun), Summer (Jul/Aug/Sep) and Fall (Oct/Nov/Dec).

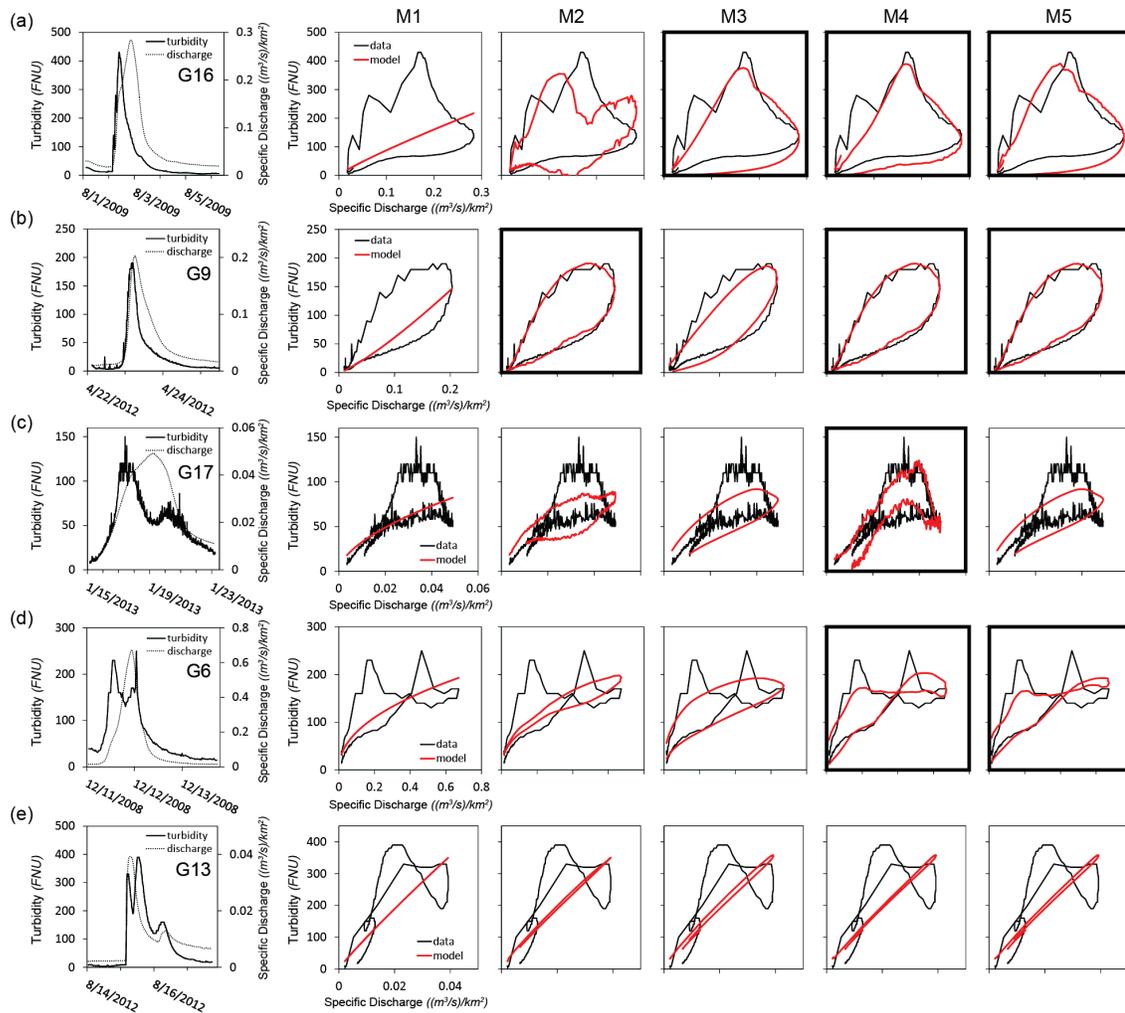
The discharge-turbidity centroid-to-centroid lag time (where here centroid is equivalent to the center of mass), represented here as  $t_{cc}$  (hours), was used to characterize the timing of the event turbidity response. Positive values of  $t_{cc}$  indicate how much the turbidity centroid follows, or lags, the discharge centroid. The dominance of negative values of  $t_{cc}$  in Figure 2.3a shows that for these gages, the event turbidity primarily leads the discharge. As catchment drainage area increases (from left to right across the figure—gages are numbered by increasing drainage area), the typical lead time generally increases ( $t_{cc}$  becomes more negative) and a greater range of timing behavior is observed (overall box and whisker length). A strong correlation exists when gage mean  $t_{cc}$  is plotted against catchment drainage area (Figure 2.3b), indicating that while there is event-to-event variability in  $t_{cc}$  (Figure 2.3a), the mean catchment behavior is predictable in terms of drainage area.



**Figure 2.3.** (a) Boxplots of event centroid-to-centroid discharge-turbidity lag ( $t_{cc}$ ) by gage; negative values indicate the centroid of turbidity leads the centroid of discharge. Boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside of the upper and lower whiskers. (b) Gage mean  $t_{cc}$  as related to catchment drainage area. Each point represents a gage.

## 2.4.2 Turbidity-Discharge Loop Modeling

Several example loop events and model fits are shown in Figure 2.4. Although these are just a few events out of the total 1559 shown, they were chosen to summarize many aspects of the general ability of the loop models to fit the data. For relatively simple discharge and turbidity loops (Figure 2.4a and Figure 2.4b), a generally good fit can be achieved with a 3-parameter model (M2 or M3). Particular catchments are better fit using the supply-based model (M3), as in Figure 2.4a, while the dominant loop shape of some catchments is more suited for hydrograph slope-based model (M2), as in Figure 2.4b. For simple events there is little gain in fitting ability by increasing model complexity to a 4-parameter model (M4 and M5). However, the use of a 4-parameter model allows a broader range of events to be fitted with a single model (both Figure 2.4a and Figure 2.4b loops for example). For other events (Figure 2.4c and Figure 2.4d), the combination of both the supply-based and hydrograph slope-based terms in the model (both M4 and M5) allows a better fit than either term alone (as in M2 and M3). Specifically in these examples, M4 and M5 allow the representation of a loop shape resulting from an event with an unequal number of discharge and turbidity peaks. Despite the advances in fitting capability offered by the 4-parameter models M4 and M5, there remain some complex events (Figure 2.4e for example), accounting for ~10% of the dataset, whose looped shape cannot be mimicked by any of the models evaluated here.

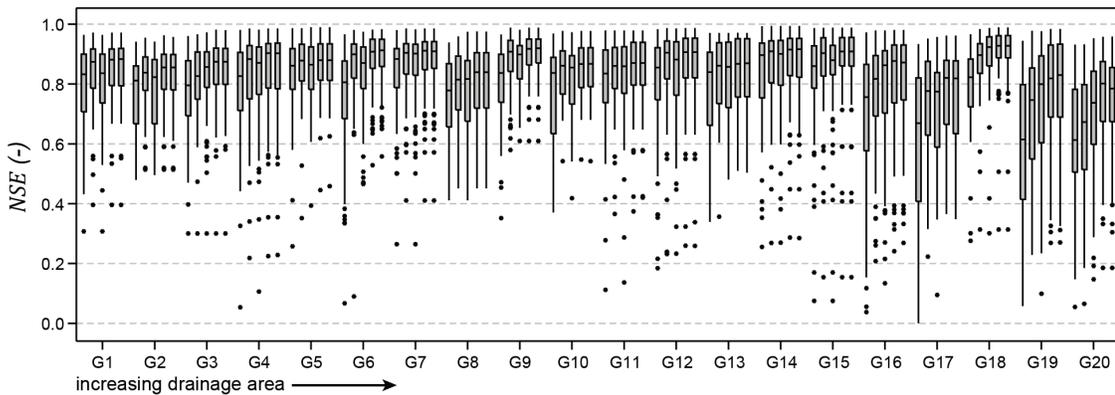


**Figure 2.4. Example event data and model fits. (a) Raritan R. (G16) event #20, (b) French C. (G9) event #100, (c) Pamunkey R. (G17) event #57, (d) WB Brandywine C. (G6) event #34, and (e) Smith C. (G13) event #20. Bold outline indicates the best model(s) for each event shown.**

The loop models were quantitatively evaluated and compared using the Nash-Sutcliffe Efficiency ( $NSE$ ) (equation (2.12)). Figure 2.5 shows individual event  $NSE$  for each model and each gage in the form of boxplots. In general, the larger catchments (towards the right on Figure 2.5) have lower median model  $NSE$  and greater event  $NSE$  variability (overall box and whisker length), the strong exception being G18. Larger catchments also display some of the largest increases in median  $NSE$  with increasing model complexity (moving from left to right within each cluster of five boxplots for a

gage). This is in part due to the fact that events from the larger catchment gages nearly always displayed strong looped behavior, whereas the small and mid-sized catchments often had narrower loops and occasional non-looped events. Therefore, going from the simple power law model (M1) to models increasingly capable of representing loops offered the greatest improvement for these gages.

Comparing the 3-parameter models (M2 and M3) to the simple rating curve (M1), we see that model M2 has a higher median NSE than M3 for most of the gages from G1 – G15; whereas, the opposite ( $M3 > M2$ ) is generally true for gages G16 – G20. This suggests that for the gages included in this study, the inclusion of a supply-based term in the model better reflects the turbidity dynamics of the larger catchments; whereas, the addition of a hydrograph slope-based term is superior for the mid-sized and smaller drainage areas.



**Figure 2.5. Boxplot of *NSE* for each model (M1-M5) by gage. Each cluster of five boxes corresponds to models M1-M5, in order from left to right, for the gage shown below the cluster. Boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside of the upper and lower whiskers.**

Models M4 and M5 (the 4-parameter models) have a similar *NSE* median and variability for most gages. The median model *NSE* of the 4-parameter models, by gage, is 0.001 – 0.065 higher than the best 3-parameter model (M2 or M3). The gages that show the greatest improvement in using a 4-parameter model over a 3-parameter model are several of the smallest (G2, G3 & G4) and largest (G17, G19 & G20) catchments included in this study. These gages had a generally higher occurrence of complex, multi-peaked events. For the small, urban catchments precipitation is transformed into discharge with little dampening or smoothing, resulting in increased event discharge and turbidity complexity. The effects of spatially non-uniform precipitation and channel routing become more important in large catchments, influencing the number of peaks within a single event. In both of these cases, the 4-parameter models offered the greatest improvement over the 3-parameter models.

### 2.4.3 Loop Model Parameters

Values of parameter  $a$  were decorrelated from parameter  $b$  for each catchment and each model. Figure 2.6 shows the results of the parameter decorrelation for all catchments and model M4; the results for the other models were similar. Initial model parameters  $a_4$  and  $b_4$  show a strong correlation (Figure 2.6a), with  $a_4$  varying by over 12 orders of magnitude across the 20 gages. Through the use of the  $Q_{4,0}$  scaling factor (calculated from values of  $a_4$  and  $b_4$ ),  $a_4$  is transformed into  $a_{4,0}$  with units of (*FNU*) and only 4 orders of magnitude variation across the 20 gages (Figure 2.6b).

Figure 2.7 shows the calculated values of  $Q_{n,0}$  for each model and each of the study gages. As catchment size increases (left to right across Figure 2.7), the value of

$Q_{n,0}$  tends to decrease. For individual gages, the values of  $Q_{n,0}$  for the five models are generally consistent. This largely model-independent behavior of the calculated scaling factor suggests that for the turbidity-discharge relationships examined here,  $Q_{n,0}$  behaves as an intrinsic characteristic of the catchment.

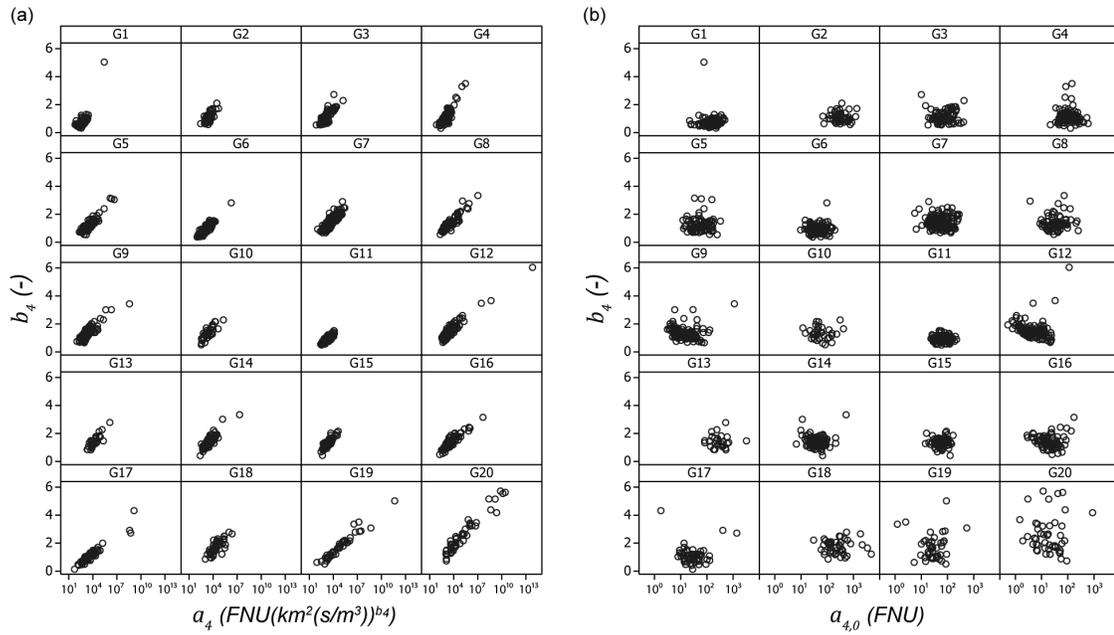


Figure 2.6. Scatterplots by gage of (a)  $b_4 (-)$  vs.  $a_4 (FNU(km^2(s/m^3))^{b_4})$  and (b)  $b_4 (-)$  vs.  $a_{4,0} (FNU)$ . Each marker represents an individual hydrologic event.

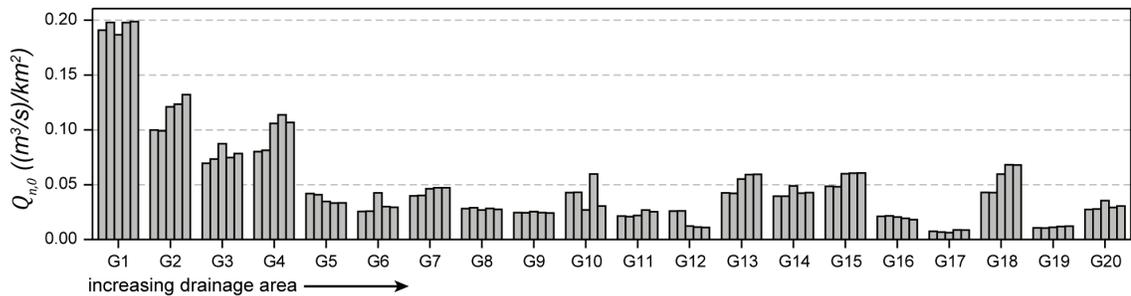


Figure 2.7. Calculated values of  $Q_{n,0}$  for each model and gage. Each cluster of 5 bars corresponds to models M1-M5, in order from left to right.

The decorrelation of  $a$  from  $b$  should allow the underlying relationships between the power law coefficient,  $a_0$ , and catchment and/or event characteristics to be better understood. Figure 2.8 shows a simple comparison of model M4 seasonal behavior between original and decorrelated parameters  $a_4$  and  $a_{4,0}$  for gages G16 and G9. Weak seasonality of  $a_4$  is visible in these example gages with a peak around April-May (not considering sparse high values). However,  $a_{4,0}$  displays more distinct seasonal behavior with a strong peak in July. The scatter present in Figure 2.8 shows that, in addition to the apparent seasonal pattern, there are other factors (e.g., event magnitude, antecedent moisture conditions, etc.) influencing the parameter values. While further investigation is required to determine the extent of the improvement in power law (rating curve) parameter interpretation provided by decorrelation, these results suggest that such an investigation is worth undertaking.

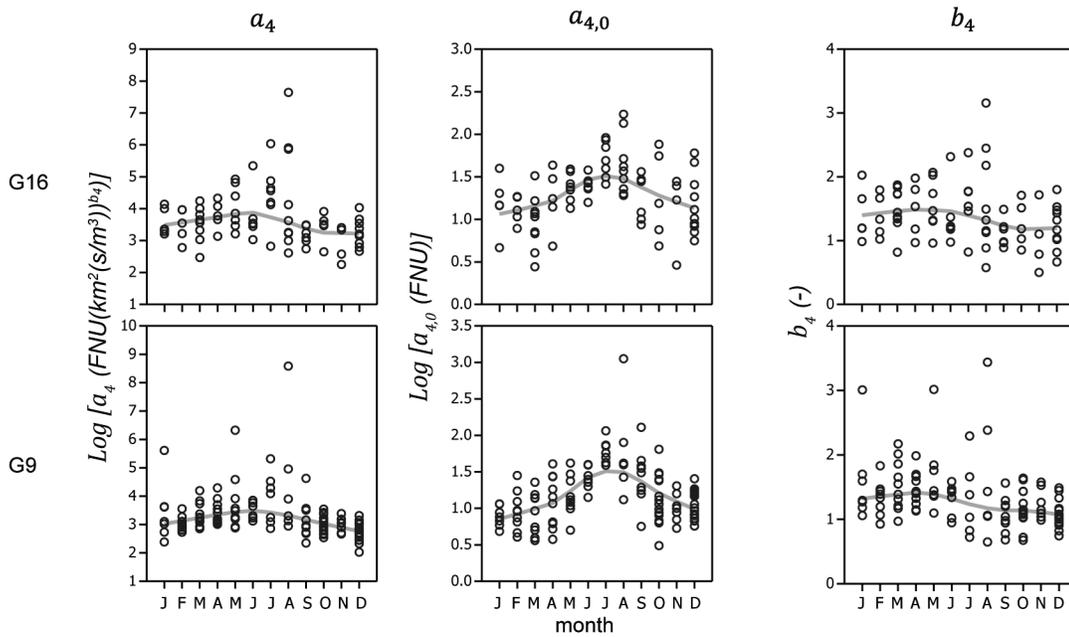


Figure 2.8. M4 model parameters  $a_{4,0}$ ,  $a_4$ , and  $b_4$  by month for gages G16 and G9. Data points are individual events and the trendline is a LOWESS smoother.

## 2.5 Discussion and Conclusions

The increasing availability of high temporal resolution, co-located turbidity and discharge data provides an opportunity for more-detailed examination of the factors that control individual event turbidity behavior. Our objective here was to evaluate the skill of several existing loop models and to improve performance by combining elements of existing models into a new model for individual event loop shape characterization for catchments of varying size and characteristics. To accomplish this, the primary drivers of turbidity behavior, namely source material supply and event-driven transport, were considered. Our research suggests that a more general model, including both terms, allows characterization of a wide range of simple and complex events. This increased applicability provides a foundation for characterizing observed loop shapes, supporting the larger goal of understanding the physical controls on turbidity-discharge hysteretic behavior.

Previous turbidity-discharge studies have generally used one of two types of mathematical loop models, one based on the local slope of the hydrograph [*Krueger et al.*, 2009, 2012; *Eder et al.*, 2010] and one based on the evolution of the “supply pool” throughout the event [*VanSickle and Beschta*, 1983; *Asselman*, 1999]. However, in these studies the models were tested primarily on very small ( $\sim 10^0 \text{ km}^2$ ) catchments [*VanSickle and Beschta*, 1983; *Eder et al.*, 2010] and field-scale ( $\sim 10^2 \text{ km}^2$ ) sites [*Krueger et al.*, 2009; *Eder et al.*, 2010], with the exception of supply-based modeling for events on the Rhine River ( $\sim 10^5 \text{ km}^2$ ) [*Asselman*, 1999]. The current study evaluated these two loop models using a group of 20 gages with catchments ranging in size from  $10^0$  to  $10^4 \text{ km}^2$ . These models were able to represent the main features of various simple

looped events, with the hydrograph slope-based model capable of producing both clockwise and counter-clockwise loops, while the supply-based model could only form clockwise loops (when parameter  $r$  is constrained to be greater than zero). In general, it was found that the supply-based model was better able to fit the event shapes from the larger catchments included in the study ( $\sim 10^3 - 10^4 \text{ km}^2$ ), whereas, the hydrograph slope-based model was more appropriate for the small and mid-sized catchments ( $\sim 10^0 - 10^3 \text{ km}^2$ ).

We formulated two new models, incorporating the primary aspects of the existing models. These new loop models included both the hydrograph slope term ( $cdQ/dt$ ) and the exponential supply term ( $e^{-rV}$ ) and differed only in the manner in which terms were combined (i.e., in one model the supply term was applied only to the discharge portion of the equation, while in the other it was applied to both the discharge and hydrograph slope portions of the model). The results show that for relatively simple loops, resulting from a single discharge and turbidity peak, the new models were not significantly better than the existing supply-based and hydrograph slope-based models. However, by including both supply and hydrograph slope terms, the new models were able to represent a much broader range of simple loop behaviors (for example, both the tear-drop shape of Figure 2.4a and the triangular loop of Figure 2.4b). Additionally, the new models showed superior capabilities for fitting many more complex shapes (Figure 2.4c and Figure 2.4d). In general, applying the supply term to the hydrograph slope portion of the model ( $c(dQ/dt)e^{-rV}$ ), as in model M5, generally did not improve model performance over model M4, therefore the simpler model (M4) is suggested as the preferred approach. Model M4 can be summarized by the following equations:

$$TURB(t) = a_0 \left( \frac{Q(t)}{Q_0} \right)^b e^{-rV(t)} + c \frac{dQ(t)}{dt} \quad (2.13)$$

$$\text{where, } V(t) = \int_0^T Q(t)dt \cong \sum_{i=0}^T Q_i \Delta t \quad (2.14)$$

$$\text{and, } a_0 = a(Q_0)^b \quad (2.15)$$

Although this model contains an additional fitting parameter as compared to either of the two existing models alone, it does not contain any “new” parameters, but instead leverages those that have been used previously. Both the supply and event-driven transport are believed to be primary drivers of turbidity behavior, and as a result, both terms are needed in equation (2.13) to represent the complexity of the processes involved.

Some complex event shapes proved challenging for all of the models tested here, despite the improved capabilities of the new models. These events typically had differing numbers of discharge and turbidity peaks. Multi-peaked events with different numbers sediment and discharge peaks have previously proved challenging in loop modeling studies [Eder *et al.*, 2010]. With some events (e.g., Figure 2.4c and Figure 2.4d) the new model was able to produce at least moderately skilled fitting; however, with other events this was not the case (e.g., Figure 2.4e). It was observed with the data used in this study that multiple turbidity peaks were often associated with multiple rainfall periods during a single discharge event. It is likely that including precipitation in addition to discharge in the event model would improve model fits for these complex cases. Another possible enhancement to the model is the explicit addition of a “lag” in the arrival of event turbidity response. The data in Figure 2.3b indicate that lag time correlates well with drainage area ( $r^2 \cong 0.9$ ), which suggests that the addition of this term might represent an

approach to improve model performance; however, there is significant scatter in lag time from event-to-event (Figure 2.3a) and additional work will be required to assess the potential of this approach.

This study employed a power law parameter decorrelation method using a discharge-dimensioned scaling factor ( $Q_0$ ) unique to each catchment. In doing this, we were able to transform the power law coefficient,  $a$  with units of  $(FNU(km^2(s/m^3))^b)$ , into  $a_0$ , which has simpler units of  $(FNU)$  and is independent from the power law exponent,  $b$ . Furthermore, this method allowed the decorrelation of parameters  $a$  and  $b$ , which has been suggested [Bergner and Zouhar, 2000] to improve their physically-based interpretation. Within this study, we compared the seasonal trends of the power law parameters (Figure 2.8), and the results show that  $a_0$  and  $b$  display distinct, and independent, seasonal behavior. While further investigation is required to determine the full extent of benefits power law parameter decorrelation provides, these results are encouraging. In addition, for the catchments examined here, the values of  $Q_0$  were quite constant among all of the models, and thus it is hoped that they can be interpreted as an intrinsic characteristic of the catchment.

While no significant attempt towards physical interpretation of model parameters was presented here, the model resulting from this study (M4) lends itself to future work in this regard. Turbidity behavior is diverse, varying largely in loop size, shape and direction across time and space. Existing loop models fail to differentiate between various nuances of loop shape and thus fall short of extracting some of the information of the loop. With the model used here, the loop is described using two terms,  $c$  and  $r$ , while

the overall position on the turbidity-discharge plot is described by  $a$  and  $b$ . The combination of both  $c$  and  $r$  provides greater discrimination and description of loop shape than either term alone. Future research is expected to link the magnitude of these parameters with catchment and hydrologic controls on turbidity sources and behavior based on catchment and event-based hydrologic characteristics.

# CHAPTER 3

---

## **Characterization and Exploratory Analysis of Stream Turbidity during 5928 Events from 110 U.S. Gages**

*Amanda L. Mather and Richard L. Johnson*

### **3.1 Abstract**

Stream turbidity behavior during hydrologic events varies from stream-to-stream as well as from event-to-event for a particular stream. To better understand this variability, stream turbidity during 5928 hydrologic events from 110 U.S. gages was modeled. The resulting fitted model parameters were used to explore relationships between stream turbidity and catchment characteristics, including geographic location. In general, land cover characteristics and stream baseflow index had the best correlations with model parameters. Cluster analysis was used to identify groups of gages with similar model parameter values. This elucidated some regional patterns, particularly dividing the west coast gages from the gages in the eastern half of the U.S. The results of this research suggest that prediction of typical event turbidity behavior using catchment characteristics may be possible.

## 3.2 Introduction

Stream turbidity during hydrologic events (i.e., the turbidigraph) is a reflection of active catchment and hydrologic processes and their relative degrees of influence [Seeger *et al.*, 2004; Mukundan *et al.*, 2013b]. These processes vary over time and space (i.e., both within a catchment and between catchments) and result in a diverse array of observed turbidigraph shapes [Seeger *et al.*, 2004; Lawler *et al.*, 2006; Eder *et al.*, 2010; Mukundan *et al.*, 2013b]. Figure 1.1 in Chapter 1 provides an example of observed event-to-event variability in stream turbidity for a single stream gage site. Efforts to examine the factors influencing turbidigraph shape have generally been limited to <30 events from at most a few catchments in a given study [e.g., Seeger *et al.*, 2004; Lawler *et al.*, 2006; Eder *et al.*, 2010; Mukundan *et al.*, 2013], and a systematic evaluation of many events from a large sample of catchments to explore the factors driving turbidigraph variability and dominant catchment behavior has not yet been undertaken (in the spirit of Gupta *et al.* [2014]).

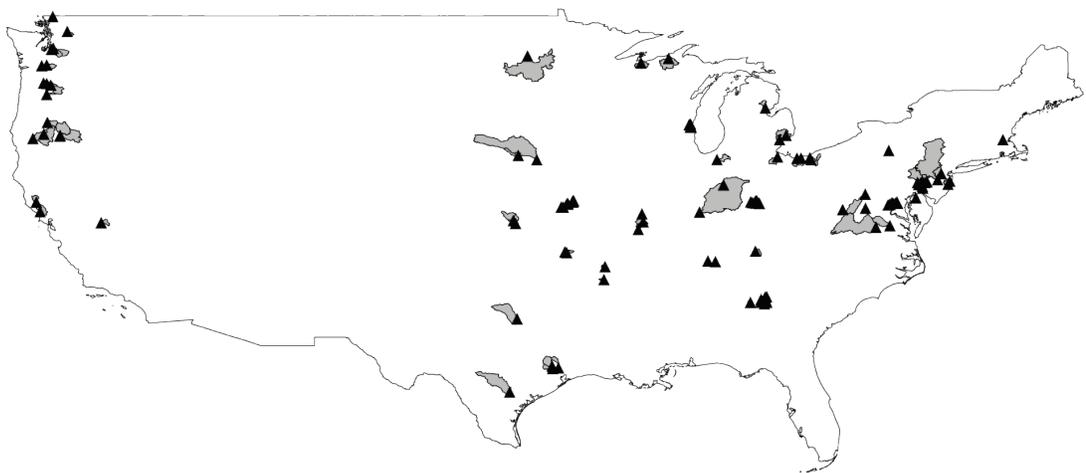
The objective of this study was to explore how turbidigraph shapes vary across U.S. streams. To accomplish this we utilized an empirical model to describe observed turbidity behavior during hydrologic events. This approach allowed the quantitative characterization of turbidigraph shape in the context of the hydrograph shape, which was important due to the strong linkage between streamflow and stream turbidity. Additionally, utilizing an event-based model supported the comparison of a large number of events. By applying the event turbidity model to many streams with diverse characteristics, we were able to explore the relationships that exist between the resulting

model parameters and the observable catchment characteristics. Additionally, model parameters were mapped to observe any regional patterns that may be present.

### 3.3 Methods

#### 3.3.1 Study Gages and Datasets

A total of 110 gages located across the contiguous U.S. were used in this study (Figure 3.1 and Table B 1 of Appendix B). Selection of these gages was made by considering all gages for which U.S. Geological Survey (USGS) discharge (i.e., streamflow) and turbidity data were available for the same time period for at least one full year. Several gages were then discarded due to extensive flow regulation or withdrawal. Some multigated and nested catchments were included, accounting for ~20% of the dataset. There were no suitable gages in the western half of the country outside of the west coast region, resulting in the large area lacking study gages (Figure 3.1). A few urban areas have clusters of closely spaced, small drainage area catchments.



**Figure 3.1. Location of 110 gages (triangles) and corresponding catchment outlines (shaded areas).**

Sub-hourly discharge and turbidity data (codes 00060 and 63680, respectively) were obtained from the USGS. The datasets were typically recorded at 15- or 30-minute intervals and were interpolated in this study to a common 15-minute timestep. The duration of record used for each gage varied from 1 to 6 years, based on data availability, within the time period of 2007-2013. Specific details regarding the data time periods used for each gage can be found in Table B 1.

Hydrologic (i.e., discharge) events were identified using an automated peak detection program (LabVIEW, National Instruments). Each event window was assumed to extend from  $0.5N$  days prior to the discharge peak to  $N$  days following the peak, where  $N = A^{0.2}$  is the duration of surface runoff in days and  $A$  is the drainage area in square miles [Sloto and Crouse, 1996]. For 11 gages, the size of the event window prior to the discharge peak had to be extended by a factor of 2-3 to capture the full event discharge and turbidity response. This modification of event window size was made for all events from affected gages (i.e., not on an event-by-event basis). The gages where this modification was used are identified in Table B 1.

A total of 29 topographic, morphologic, land cover, soil, climate, impoundment and discharge characteristics of the catchment draining to each gage were calculated. These characteristics were selected either due to their nature as basic catchment physical descriptors or because they are known or believed to be related to sources, sinks and/or transport of turbidity-causing materials (e.g., suspended sediment). A description of the data sources and methods used in the calculation of these characteristics can be found in Section B.2 of Appendix B. A summary of the catchment characteristics for the study gages is shown in Table 3.1.

**Table 3.1. Summary of the catchment characteristics for the 110 study gages.**

Catchment Characteristic	Minimum	Mean	Median	Maximum
<i>Topography &amp; Morphology</i>				
Gage latitude, <i>LAT</i> (degrees)	28.8487	39.6562	39.0868	48.9748
Gage longitude, <i>LONG</i> (degrees)	-123.318	-91.2646	-86.6598	-71.2703
Drainage area, <i>DA</i> (km <sup>2</sup> )	1.24	1804	261	29280
Shape factor, <i>SF</i> (-)	0.11	0.36	0.33	1.31
Drainage density, <i>DD</i> (km/km <sup>2</sup> )	0.57	1.51	1.29	4.70
Mean catchment slope, <i>SLP</i> (degrees)	0.28	4.28	2.99	20.70
Mean catchment elevation, <i>ELEV</i> (m)	17	387	280	2847
Relief, <i>RELIEF</i> (m)	19	527	183	4384
Relief ratio, <i>RELIEF RATIO</i> (-)	0.00084	0.014	0.0075	0.073
<i>Land Cover</i>				
Fraction of forest land cover, <i>FOREST</i> (%)	0.2	37.5	33.7	96.7
Fraction of urban land cover, <i>URBAN</i> (%)	0	28.5	13.3	99.6
Fraction of agriculture land cover, <i>AG</i> (%)	0	21.5	11.4	88.1
Fraction of wetland or water land cover, <i>WETWAT</i> (%)	0	3.5	1.8	21.8
<i>Climate &amp; Soils</i>				
Mean annual temperature, <i>MAT</i> (°C)	4.4	12.2	12.2	21.7
Mean annual precipitation, <i>MAP</i> (m)	0.26	1.15	1.10	2.66
Precipitation seasonality index, <i>PSI</i> (-)	0.13	0.31	0.28	0.77
Aridity index, <i>AI</i> (-)	0.34	1.09	1.00	4.17
Fraction of type A SCS hydrologic soil group, <i>SOILA</i> (%)	0	4.2	0.5	47.8
Fraction of type B SCS hydrologic soil group, <i>SOILB</i> (%)	0	42.3	42.1	99.2
Fraction of type C SCS hydrologic soil group, <i>SOILC</i> (%)	0	34.4	30.5	100
Fraction of type D SCS hydrologic soil group, <i>SOILD</i> (%)	0	18.8	15.5	77.8
<i>Impoundments</i>				
Impoundment density, <i>ID</i> (number/km <sup>2</sup> )	0	1.15	0.84	5.17
Fraction of impoundment area, <i>IA</i> (%)	0	1.01	0.65	8.68
Fraction of largest impoundment area, <i>IMA</i> (%)	0	0.29	0.11	4.53
<i>Discharge</i>				
Runoff ratio, <i>RR</i> (-)	0.033	0.434	0.413	1.500
Mean annual runoff, <i>MAR</i> (m)	0.025	0.530	0.448	2.614
Streamflow variability, <i>SV</i> (-)	0.495	2.147	2.030	5.336
Skewness in daily flows, <i>SKEW</i> (-)	1.106	2.652	2.109	9.079
Baseflow index, <i>BFI</i> (-)	0.293	0.575	0.568	0.920

Note: Details regarding source data and calculation methods can be found in Section B.2 of Appendix B.

### 3.3.2 Turbidity-Discharge Event Modeling

Turbidity (*TURB*) during each hydrologic event was modeled using a power-law based event model with supply and discharge (*Q*) slope terms [*Mather and Johnson, 2014*]:

$$TURB(t) = aQ(t)^b e^{-rV(t)} + c \frac{dQ(t)}{dt} \quad (3.1)$$

$$\text{where, } V(t) = \int_0^T Q(t)dt \cong \sum_{i=0}^T Q_i \Delta t \quad (3.2)$$

and *a*, *b*, *c* and *r* are fitting parameters. Parameter *r* was not permitted to take negative values. The local slope of the hydrograph (*dQ/dt*) was calculated using the central difference method:

$$\frac{dQ_i}{dt} = \frac{Q_{i+1} - Q_{i-1}}{2\Delta t} \quad (3.3)$$

and was smoothed using a Savitzky-Golay filter [*Savitzky and Golay, 1964*]. Discharge was scaled to catchment drainage area and therefore had units of  $(m^3/s)/km^2$ . The turbidity data used here have units of (*FNU*) [*Anderson, 2005*] and are treated as a “concentration”. Model parameters *a*, *b*, *c* and *r* have units of  $(FNU(km^2(s/m^3)))^b$ ,  $(-)$ ,  $(FNU(km^2s)(s/m^3))$  and  $(km^2/m^3)$ , respectively.

The event turbidity model used here is, by design, fairly complex (as measured by the degrees of freedom) and is intended to be applicable for varied turbidigraph shapes resulting from diverse catchment characteristics (e.g., size, climate, land cover) and individual event hydrologic characteristics (e.g., season, precipitation, antecedent conditions). However, the full complexity of the event model may not be warranted for

all catchments or each individual event. Here we simplify the event model structure on an event-by-event basis by removing irrelevant model terms to obtain an appropriately parsimonious model [Linsley, 1982; Gupta et al., 2014]. For each event, the initial model (equation (3.1)) was fit to the event data with non-linear curve fitting using the trust-region dogleg (TRDL) algorithm and least-squares minimization (LabVIEW 2012, National Instruments). Model parameters  $c$  and/or  $r$  were removed from the model (by setting the parameter value to zero) if the 95% confidence interval for parameter  $c$  and/or  $r$  contained zero. The model was then refit with the remaining model terms to obtain the final event model.

After obtaining the final model parameters for each event, parameter  $a$  was transformed into  $a_0$  using a discharge-dimensioned, gage-specific scaling factor,  $Q_0$  [Bergner and Zouhar, 2000; Mather and Johnson, 2014] using the expression:

$$a_0 = a(Q_0)^b \quad (3.4)$$

$$\text{where, } Q_0 = 10^{-\left(\frac{\sum_{j=1}^k (b_j - \bar{b}) \log(a_j/a_g)}{\sum_{j=1}^k (b_j - \bar{b})^2}\right)} \quad (3.5)$$

$$\text{and, } a_g = \left(\prod_{j=1}^k a_j\right)^{\frac{1}{k}} \quad (\text{geometric mean}) \quad (3.6)$$

$j$  is the gage event index,  $k$  is the total number of individual events for the gage,  $\bar{b}$  is the mean of the  $b$  values for the gage and  $a_g$  is the geometric mean of the  $a$  values for the gage. This transformation decorrelates the power law parameters  $a$  and  $b$ , and simplifies the units of parameter  $a$  from  $(FNU(km^2(s/m^3))^b)$  to  $(FNU)$ , which is not dependent on parameter  $b$ . The method requires that the event model (equation (3.1)) be fitted to all

of the events first, then the fitted values of  $a$  and  $b$  are used to calculate  $Q_0$  (equation (3.5)), which is used to transform the values of  $a$  into  $a_0$  using equation (3.4). The actual model fit to the observations, and the other parameter values ( $b$ ,  $c$  and  $r$ ), remain unchanged.

The Nash-Sutcliffe Efficiency ( $NSE$ ) [Nash and Sutcliffe, 1970] was used to evaluate the goodness-of-fit for each event:

$$NSE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.7)$$

where  $y_i$  is the observed value,  $\hat{y}_i$  is the predicted value and  $\bar{y}$  is the mean of the observed values for the event.  $NSE$  is unitless and ranges from  $-\infty$  to 1; values greater than zero indicate that the fit is better than the mean of the observations.

### 3.3.3 Cluster Analysis

Cluster analysis was used to group gages based on their typical turbidigraph shape as indicated by the event model parameter values. The input data used for cluster analysis were the median parameter values (i.e., median  $a_0$ ,  $b$ ,  $c$  and  $r$ ) for all events from each gage. Clusters were determined using an agglomerative hierarchical algorithm in which each gage is originally its own cluster and then the individual-gage clusters are repeatedly combined into larger clusters based on input data similarity. The input data were standardized, and gage median  $a_0$  was log-transformed. Euclidean distance was used as a distance metric, and Ward's method was used as the linkage criterion [e.g., Olden *et al.*, 2012].

## 3.4 Results and Discussion

### 3.4.1 Model Fit

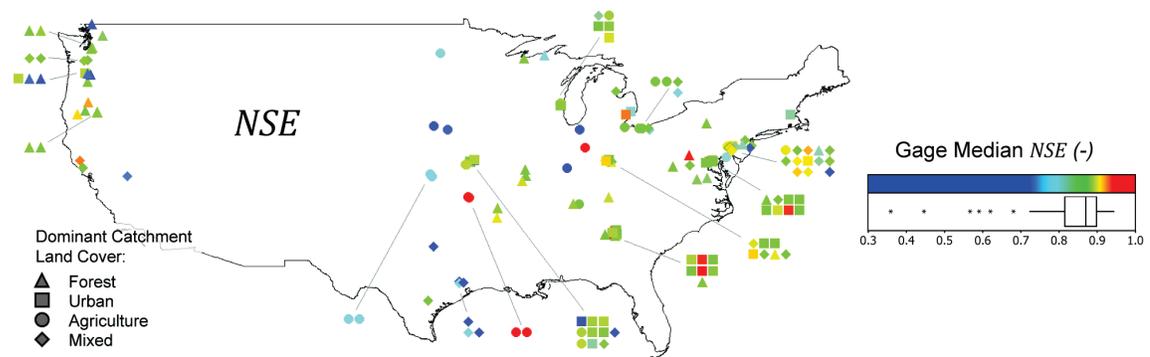
A total of 5928 events from the 110 gages were used in this study, with individual gages having 8 to 195 events (median of 54) (Figure B 1). The largest drainage area catchments in the study set tended to have the fewest number of events, likely because larger weather systems were required to produce an “event” discharge response.

The event model *NSE* was calculated for each event and the results, grouped by gage, are shown in Figure 3.2. *NSE* distributions tend to be slightly tailed with some outliers towards low values. The outlier events were usually complex, non-typical events that were not expected to be well-fit by an event model. Overall, as drainage area increases (left to right across Figure 3.2), *NSE* decreases and becomes more variable (i.e., longer boxplot).

The median model *NSE* for each gage is mapped in Figure 3.3, showing >75% of the gages have a median *NSE* greater than 0.80 and ~25% greater than 0.90. Distinct spatial patterns are not apparent in Figure 3.3; however, a portion of the middle region of the U.S. is characterized by gages having a generally lower median *NSE*.



The Spearman's rank correlation coefficient between gage median *NSE* and various catchment characteristics is shown in Table 3.2. Drainage area (*DA*) and wetland/water land cover (*WETWAT*) had the highest absolute value correlation and both are negatively correlated with median *NSE*. Water bodies (reservoirs and lakes) and wetlands both act to reduce and delay the turbidity event response (possibly causing the turbidity peak to occur substantially after the discharge peak) and can cause lengthy “tailing” on the turbidigraph. These characteristics are more challenging for the event model to fit than the more-typical case where the turbidity peak leads discharge with subsequent exhaustion [Mather and Johnson, 2014]. The correlation with drainage area is likely related to the increased effects of spatial non-uniformity in precipitation and active turbidity-causing material sources across larger catchments.



**Figure 3.3. Median event *NSE* by gage.<sup>4</sup> Key to the right of map shows a boxplot of gage median event *NSE* and color key for mapped markers. Gradient on color key extends from lower (blue) to upper (red) whisker, centered on the median (green). Low outliers are blue and high outliers are red. Mapped marker shape indicates dominant catchment land cover: forest (triangle), urban (square), agriculture (circle) or mixed (diamond). A particular land cover category is considered dominant if it represents at least 60% of the catchment area.**

<sup>4</sup> For Figure 3.3, and for the spatial maps in the remainder of this chapter, markers for densely gaged areas are on the map both at their coordinate-determined locations and also off to the side of the map in a grid pattern where colors and/or shapes assigned to each gage are readily visible. The general spatial arrangement (i.e., cardinal directions) of the gages with respect to adjacent gages was maintained as much as possible in the grid arrangement to the side of the map.

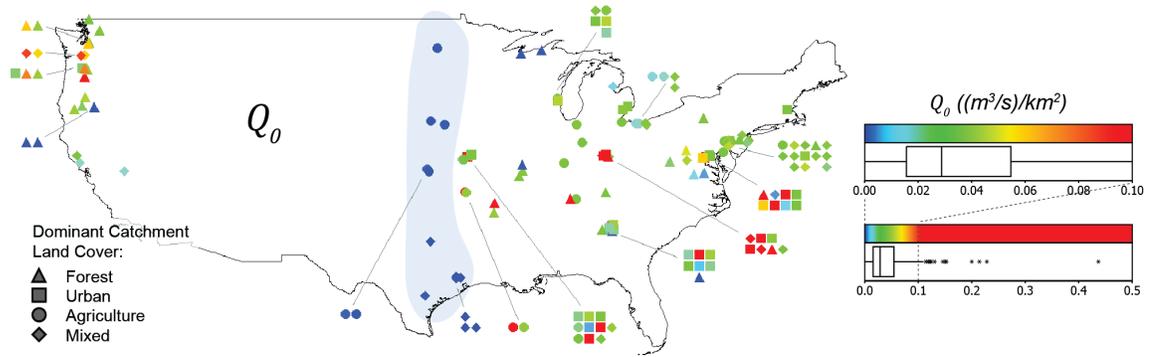
**Table 3.2. Spearman's rank correlation coefficients.**

Catchment Characteristics	Median $NSE$	Median $a_0$	Median $b$	Median $c$	Median $r$	$Q_0$
<i>LAT</i>			<b>0.46</b>			
<i>LONG</i>						
<i>DA</i>	<b>-0.30</b>		<b>0.41</b>		0.23	<b>-0.43</b>
<i>SF</i>	0.19					0.24
<i>DD</i>						
<i>SLP</i>	0.24	-0.24	<b>0.35</b>			<b>0.43</b>
<i>ELEV</i>			<b>0.41</b>		0.23	
<i>RELIEF</i>		<b>-0.38</b>	<b>0.62</b>		0.19	
<i>RELIEF RATIO</i>	<b>0.23</b>	-0.24	<b>0.28</b>			<b>0.50</b>
<i>FOREST</i>		<b>-0.44</b>	<b>0.46</b>		0.22	
<i>URBAN</i>	0.19	<b>0.30</b>	<b>-0.48</b>		<b>-0.25</b>	
<i>AG</i>		<b>0.39</b>				
<i>WETWAT</i>	<b>-0.33</b>		0.23		0.21	<b>-0.40</b>
<i>MAT</i>		0.20	<b>-0.55</b>		-0.23	
<i>MAP</i>		<b>-0.40</b>	0.21			<b>0.30</b>
<i>PSI</i>			<b>0.28</b>			
<i>AI</i>		<b>0.38</b>	<b>-0.32</b>			<b>-0.36</b>
<i>SOILA</i>	-0.20	-0.22	<b>0.36</b>			-0.21
<i>SOILB</i>			0.20			
<i>SOILC</i>				0.19		
<i>SOILD</i>						<b>-0.36</b>
<i>ID</i>		<b>0.39</b>	<b>-0.37</b>			
<i>IA</i>					<b>0.30</b>	
<i>IMA</i>						
<i>RR</i>		-0.22	0.23			<b>0.47</b>
<i>MAR</i>	<b>0.23</b>	<b>-0.29</b>	0.24		-0.21	<b>0.50</b>
<i>SV</i>		<b>0.34</b>	<b>-0.71</b>			
<i>SKEW</i>		<b>0.26</b>	<b>-0.72</b>			
<i>BFI</i>	-0.21	<b>-0.32</b>	<b>0.85</b>		0.23	

Note: Only values significant at  $p \leq 0.05$  shown; values in bold indicate  $p \leq 0.01$ . Catchment characteristic abbreviations can be found in Table 3.1.

### 3.4.2 Power Law Decorrelation Scaling Factor

As part of the power law parameter decorrelation method, a gage-specific scaling factor,  $Q_0$ , was calculated, and the resulting values for this scaling factor are shown in Figure 3.4. Spatially, values are low through the middle of the U.S. and for several (mainly forest-dominated) catchments in other regions. The highest values of  $Q_0$  within this dataset are for several small, urban-dominated catchments on the eastern side of the U.S. (although not all gages with these characteristics have high  $Q_0$ ).



**Figure 3.4.** Decorrelation scaling factor,  $Q_0$ , for each gage. Key to the right of map shows a boxplot of  $Q_0$  and color key for mapped markers. For full description of marker colors and shapes, see Figure 3.3 caption.

Table 3.2 shows that  $Q_0$  is correlated with several catchment characteristics, including drainage area ( $DA$ ), topography ( $SLP$  and  $RELIEF\ RATIO$ ), soils ( $SOILD$ ) and numerous terms related to water balance and storage ( $WETWAT$ ,  $MAP$ ,  $AI$ ,  $RR$  and  $MAR$ ). Multiple linear regression was used to examine the extent to which  $Q_0$  can be predicted from observable catchment characteristics. Best subset analysis was used, and an equation with three catchment characteristics was selected:

$$\text{Log } Q_0 = -1.58 - 0.185 \text{ Log } DA - 0.110 \text{ WETWAT}^{0.5} + 1.46 RR \quad R^2 = 52.9\%$$

The terms present in the regression equation are related to scale ( $DA$ ), storage ( $WETWAT$ ) and water balance ( $RR$ ) and are not strongly correlated to one another. These results suggest that observable catchment characteristics may provide a basis for predicting  $Q_0$  for streams where turbidity data are not available; however, this topic is beyond the scope of the current work.

### 3.4.3 Model parameter $a_0$

Values of model parameter  $a_0$  span nearly 9 orders of magnitude over the total 5928 events, from a minimum of  $2.69 \times 10^{-3}$  to a maximum of  $1.84 \times 10^6$  *FNU*. This parameter acts as a scaling-type factor in the model; therefore, high values of  $a_0$  are generally associated with overall high levels of turbidity in proportion to the event specific discharge. The values of  $a_0$  for the events from each gage are summarized in Figure 3.2 and are generally log-normally distributed over a 1-2 order of magnitude range. There is not a clear trend either in magnitude or variance of  $a_0$  with drainage area (increasing left to right across Figure 3.2). The range of  $a_0$  varies significantly between gages. This suggests that, although the model parameter varies by event, the range of  $a_0$  values for a particular gage is controlled by catchment-specific characteristics.

Figure 3.5 shows the spatial mapping of median  $a_0$  values for each gage, which vary by  $\sim 3$  orders of magnitude from 2.74 to 2713 *FNU*. The highest values of median  $a_0$  (red markers) in this dataset are mostly associated with agriculture- or urban-dominated catchments on the eastern half of the U.S. Gages on the west coast of the U.S. generally had the lowest values of median  $a_0$  (blue markers). There are several scattered gages with low median  $a_0$  in the eastern half of the U.S. and most are associated with forest-dominated catchments (triangle markers). This, along with the generally low  $a_0$  values of western forested catchments, suggests that forested catchments may tend to have lower values of  $a_0$ .

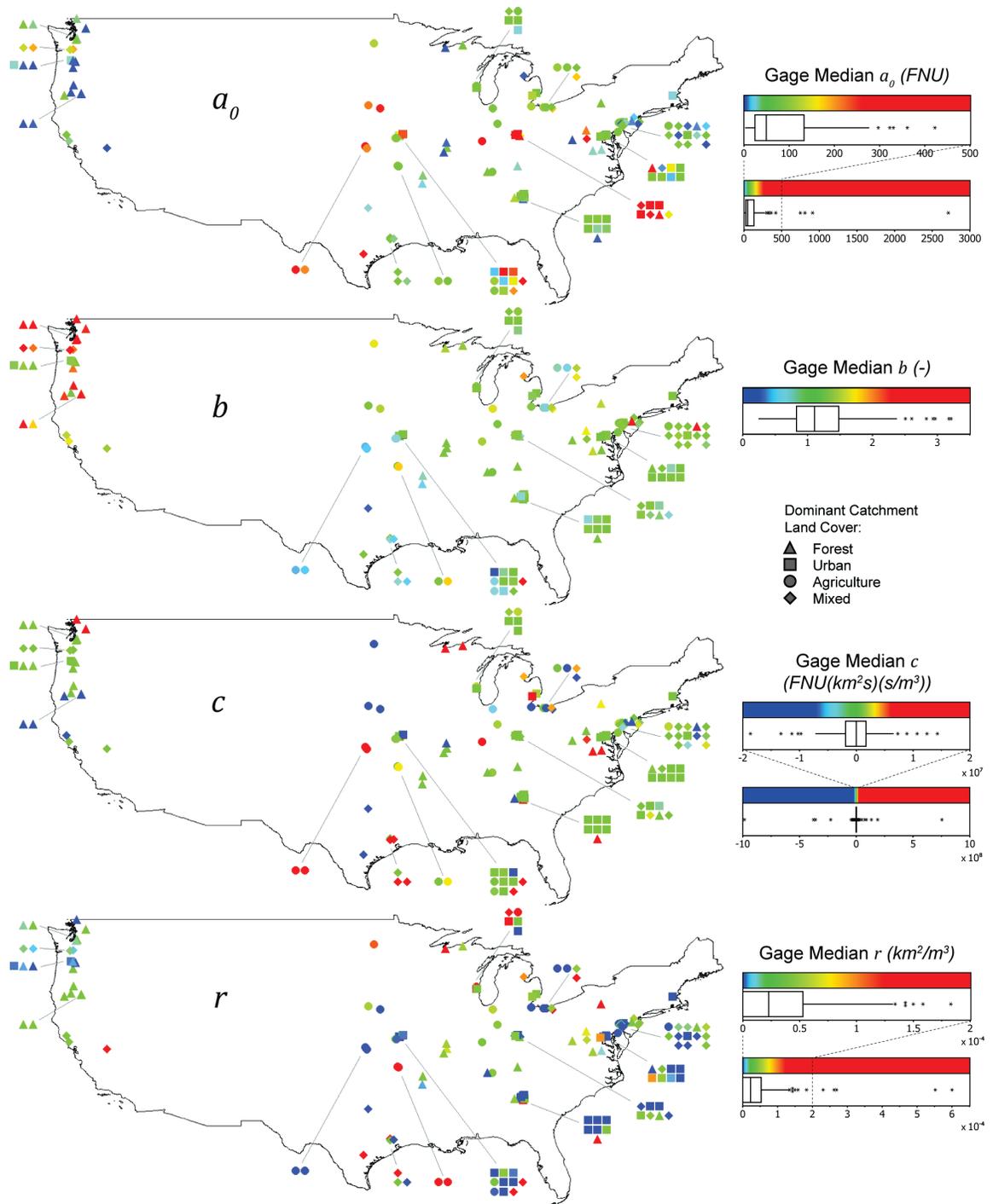


Figure 3.5. Median model parameter values for all events from each gage. Key to the right of each map shows a boxplot of gage median parameter values and color key for mapped markers. For full description of marker colors and shapes, see Figure 3.3 caption.

Gage median  $a_0$  has several significant correlations with catchment characteristics (Table 3.2), the largest magnitude correlations being with: topography (*RELIEF*), land cover (*FOREST*, *URBAN* and *AG*), climate (*MAP* and *AI*), impoundment density (*ID*) and discharge characteristics (*MAR*, *SV*, *SKEW*, *BFI*). Catchments with a high percentage of forest land cover and abundant water (low *ID*, *AI* and high *MAP*, *BFI*) tend to have the lowest values of  $a_0$ .

#### **3.4.4 Model parameter $b$**

Model parameter  $b$  ranges from an overall minimum of 0.0538 to a maximum of 8.83 over the 5928 study events. Values of  $b$  greater than one indicate that the output (turbidity) is a positively accelerating function of the input (specific discharge) (i.e., the rate of turbidity rise increases as the magnitude of specific discharge increases). Figure 3.2 shows that gages display distinct ranges of  $b$  values, however with  $b$  there is more overlap among gages than for  $a_0$ . The distribution of model parameter  $b$  for each gage varies from normal to lognormal, and the range of  $b$  for each gage varied from  $\sim 0.5$  to as large as  $\sim 6$  (ignoring star symbol outliers). Model parameter  $b$  displays a small increase in gage median and variance with increasing drainage area.

Figure 3.5 shows the spatial mapping of gage median  $b$  values, which range from 0.24 to 3.21. The spatial arrangement of median  $b$  displays two clear features, (1) nearly all of the highest median  $b$  gages are located on the west coast, and (2) the eastern half of the U.S. displays a general gradient in values, increasing from south (Texas) to north (Great Lakes).

Gage median values of parameter  $b$  are well correlated with many catchment characteristics (Table 3.2), the highest being baseflow index ( $BFI$ ), followed by discharge variability characteristics ( $SV$  and  $SKEW$ ), relief ( $RELIEF$ ) and climate ( $MAT$ ). In fact, out of the catchment characteristics considered here, 21 of the 29 have significant Spearman's rank correlation with gage median  $b$ . In general, high gage median  $b$  is associated with catchments having high baseflow index, low discharge variability, large catchment relief and low mean annual temperature.

### 3.4.5 Model parameter $c$

Model parameter  $c$  displays diverse behavior across the 110 gages. Values of model parameter  $c$  range from an overall minimum of  $-7.12 \times 10^9$  to a maximum of  $5.32 \times 10^9 \text{ FNU}(km^2s)(s/m^3)$ , with 18% of the total 5928 events having an event  $c$  of zero. The distribution of  $c$  for each gage varies widely and does not consistently conform to a standard distribution (Figure 3.2). For gages with drainage area below  $\sim 10^3 \text{ km}^2$ ,  $c$  typically has a small range and a median near zero. However, larger catchments display a much greater range of  $c$  values, both positive and negative, and their median value of  $c$  is often considerably different from zero. Positive values of model parameter  $c$  typically indicate that the turbidigraph peak occurs prior to the discharge peak, while negative values are typically associated with turbidity peaks occurring after the peak discharge.

Figure 3.5 shows the spatial mapping of gage median  $c$  values, which range from  $-9.94 \times 10^8$  to  $7.54 \times 10^8 \text{ FNU}(km^2s)(s/m^3)$ , with 29 of the 110 gages having a median value of zero. The spatial arrangement of median  $c$  does not display strong patterns; however, it does appear that there is a dominance of low values through the middle of the U.S. (Texas and northward) and possibly an increasing trend from south to north along

the west coast. The urban-dominated catchments (square markers), which also tend to be small drainage area catchments, typically have median  $c$  near zero. Agriculture-dominated catchments (circle marker) often have either a very high or a very low value of median  $c$ .

Median values of model parameter  $c$  showed the least correlation with catchment characteristics of the four parameters in the event turbidity model (Table 3.2). A weak correlation between gage median  $c$  and type C soils (*SOILC*) was found, however this was only significant correlation for the catchment characteristics considered here. One explanation for this could be that gage median  $c$  is related to characteristics local to the gage or adjacent to the stream (i.e., along stream corridor) and therefore poorly represented by the catchment-wide characteristics computed here.

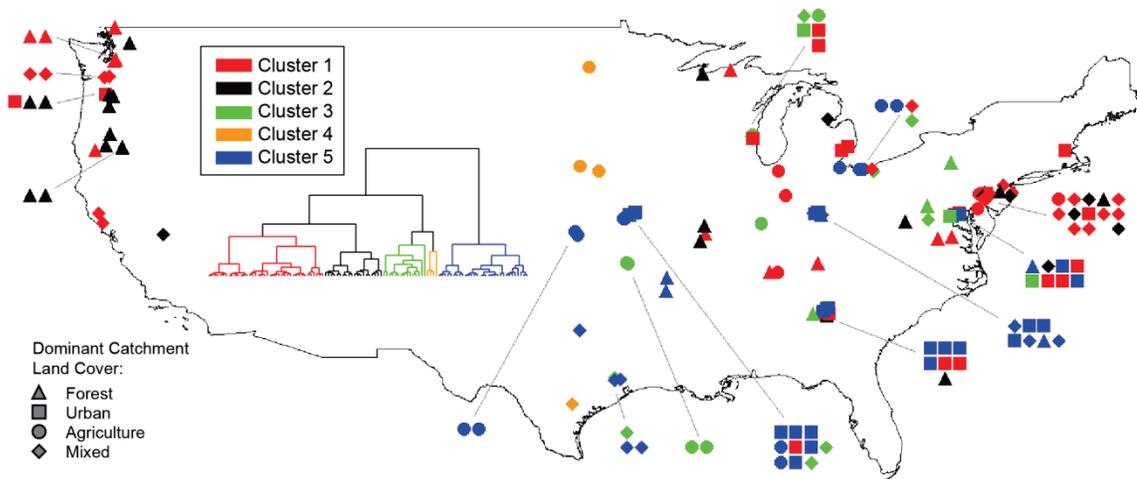
#### **3.4.6 Model parameter $r$**

Individual event values of model parameter  $r$  range from an overall minimum of 0 to a maximum of  $6.81 \times 10^{-2} \text{ km}^2/\text{m}^3$ , with 45.3% of the total 5928 events having a  $r$ -value of zero. Higher values of model parameter  $r$  indicate a more-rapid exhaustion of turbidity-causing materials (i.e. supply) during the event, which causes the turbidity peak to occur prior to the discharge peak. When event values of parameter  $r$  are equal to zero, the turbidity peak usually occurs after or together-with the discharge peak. The distribution of  $r$  for each gage is highly skewed with positive outliers (Figure 3.2). There is not a clear trend either in magnitude or variance in event values of  $r$  with drainage area.

Figure 3.5 shows the spatial mapping of gage median  $r$  values, which range from 0 to  $6 \times 10^{-4} \text{ km}^2/\text{m}^3$ , with 37 of the 110 gages having a median value of zero. The west coast has generally mid- to low-values of median  $r$ , and urban-dominated catchments (square markers) throughout the U.S. typically have low values. The gages with high median  $r$  are mostly agricultural and mixed land cover catchments. Gage median  $r$  has several significant correlations with catchment characteristics (Table 3.2), the most notable being a negative correlation with urban land cover (*URBAN*) and a positive correlation with impoundment area (*IA*).

#### **3.4.7 Gage Similarity Based on Median Parameter Values**

The 110 gages were clustered based on similarity in median values for  $a_0$ ,  $b$ ,  $c$  and  $r$ ; the resulting five clusters (C1-C5) are shown in Figure 3.6. The gages in C1 (red) are mostly either forested or mixed land cover and are located throughout the northern U.S. The median parameter values of  $c$  and  $r$  for the gages in this cluster (Table 3.3) suggest that the turbidigraph peak typically occurs slightly before or around the same time as the hydrograph peak. The value of  $a_0$  indicates that the overall event turbidity is moderately low as a function of specific discharge; however,  $b$  is greater than one indicating turbidity is a positively accelerating function of specific discharge.



**Figure 3.6. Study gages clustered by gage median model parameter values.**

C2 gages (black) are nearly all forested catchments, and all of the gages on the west coast are either from C2 or C1 (Figure 3.6). The value of  $a_0$  for C2 is the lowest, and  $b$  is the highest of the five clusters (Table 3.3). These parameter values suggest that for lower specific discharges, the event turbidity is fairly low; however, for higher specific discharge there is a large increase in turbidity for incremental increases in specific discharge. Based on the values of  $c$  and  $r$ , the turbidigraph peak typically occurs prior to the hydrograph peak for these gages.

The gages in C3 (green) are located in the eastern half of the U.S. and mostly have either agriculture or mixed land covers (Figure 3.6). Values of  $a_0$  and  $b$  for this cluster (Table 3.3) suggest moderate to high event turbidity, accelerating in rate of increase with specific discharge magnitude. Parameters  $c$  and  $r$  for this cluster are highest among the five clusters and indicate that the turbidigraph peak typically occurs substantially before the hydrograph peak for these gages.

**Table 3.3. Median of the gage median parameter values for each cluster.**

	Number of gages	$a_0$	$b$	$c$	$r$
Cluster 1 (red)	40	40.64	1.22	0	1.60E-05
Cluster 2 (black)	20	8.86	1.45	1.57E+05	3.10E-05
Cluster 3 (green)	15	106.00	1.39	1.65E+06	1.32E-04
Cluster 4 (orange)	4	225.50	1.27	-3.68E+08	8.30E+05
Cluster 5 (blue)	31	120.20	0.71	0	0.00

Note: Units are as follows:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)), and  $r$  (km<sup>2</sup>/m<sup>3</sup>).

There are only four gages in C4 (orange) and all are located in the central region of the U.S. (Figure 3.6). Three of the four gages are agricultural and all gages correspond to fairly large drainage area catchments (>700 km<sup>2</sup>). The value of  $a_0$  for this cluster indicates that the overall turbidity during events is high (Table 3.3). This cluster is the only one where parameter  $c$  is large and negative, which suggests the turbidigraph peak typically lags the hydrograph peak. The value of  $r$  is not zero and therefore acts to counter the negative value of  $c$ ; however, based on the relative magnitudes of  $c$  and  $r$ , it is likely that the typical turbidigraph behavior is dominated by parameter  $c$ , and  $r$  simply acts to modify the shape to match the event complexity.

C5 (blue) gages are located in the eastern half of the U.S. and are commonly urban catchments (Figure 3.6) and many are smaller in size. Gages from this cluster have a distinctive set of parameter values (Table 3.3). This is the only cluster having median  $b$  less than one, which specifies that turbidity behaves as a negatively accelerating function of specific discharge. Parameters  $c$  and  $r$  both have a median of zero, indicating that the turbidigraph generally occurs synchronous with the hydrograph.

### 3.5 Conclusions

This study used an empirical event model to characterize turbidigraph shape within the context of the hydrograph for 5928 events from streams across the U.S. The fitted event model parameters were used to examine spatial and catchment-characteristic relationships. Because many factors contribute to stream turbidity, it was not surprising that there was considerable scatter in all of the parameters observed here. At the same time, using the event model, together with parameter decorrelation, allowed a number of relationships between model parameters and both geographic location and catchment characteristics to be identified. In particular, land cover and baseflow index were well correlated with parameter  $a_0$  and  $b$ , respectively. These relationships were understandable in the context of known hydrologic and catchment processes, and in the way the parameters affect model behavior (i.e., based on the form of the model terms associated the parameters). The model parameters that proved to be the most challenging to understand relate to the timing of the turbidigraph peak with respect to the hydrograph peak (i.e.,  $c$  and  $r$ ). This characteristic of the turbidigraph may be better understood in future studies that include catchment characteristics local to the gage and/or adjacent to the stream (i.e., along the stream corridor) [Buck *et al.*, 2004] in addition to catchment-wide characteristics.

A number of the relationships warrant further analysis. The correlations with geographic location and catchment characteristics suggest that there may be applications for use in investigations of “ungaged” streams. In this context, ungaged streams would likely include streams for which there is a discharge dataset but no sub-hourly turbidity data. Additionally, analysis of the variability of turbidity model parameters over time for

a particular stream and relationships that may exist with event characteristics, such as antecedent moisture conditions, would further clarify the role of hydrologic and catchment processes in the determination of turbidigraph shape.

# CHAPTER 4

---

## Event-based Prediction of Stream Turbidity using Regression and Classification Tree Approaches<sup>5</sup>

*Amanda L. Mather and Richard L. Johnson*

### 4.1 Abstract

Stream turbidity typically increases along with streamflow during hydrologic events. However, similar event hydrographs can produce markedly different event turbidity due to the many factors influencing turbidity in addition to streamflow, including antecedent moisture conditions, season and supply of turbidity-causing materials. Event turbidity models can be used to quantify the event-to-event variability in stream turbidity behavior through event model parameter values. Here we examine the extent to which we can predict stream turbidity through the prediction of event model parameters. Using three mid-sized streams from the Mid-Atlantic region of the U.S., we show that event turbidity can be predicted using both multiple linear regression and classification tree approaches. The classification tree approach produced slightly better

---

<sup>5</sup> In review with Journal of Hydrology, reproduced here as submitted with edits.

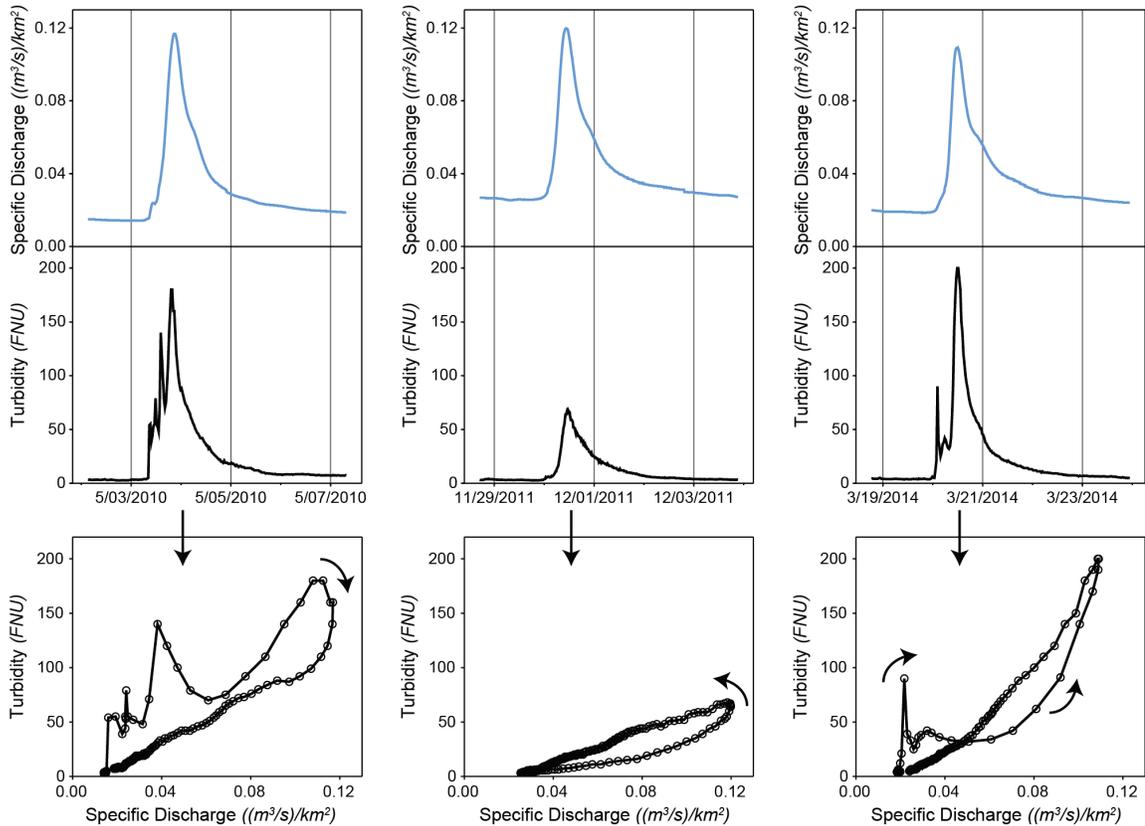
predictions; however, both approaches had a higher success rate than prediction using catchment-specific fixed-parameter values. Prediction success rates for all approaches were highest for mid-sized events. The results from both the regression and classification tree approaches suggest that catchment antecedent moisture is an important factor in the prediction of event turbidity. This research demonstrates prediction of stream turbidity using several approaches and serves as a foundation for subsequent studies of stream turbidity forecasting, which we anticipate will be key for future improvements in management of water resources.

## **4.2 Introduction**

Turbidity is an important physical and visual characteristic of surface waters that varies in general accordance with streamflow (i.e., discharge). It is also a versatile stream water quality surrogate that is relatively simple and inexpensive to monitor continuously in-stream [Davies-Colley and Smith, 2001]. Turbidity is often correlated with suspended sediment [e.g., Gray and Glysson, 2003], agricultural chemicals [e.g., Hickman, 2004], pathogens [e.g., Christensen *et al.*, 2000] and heavy metals [e.g., Miller, 1997], and is therefore an important component of water quality monitoring and regulation [Gray and Glysson, 2003].

Hydrologic, or discharge, events typically cause increases in stream turbidity that are influenced by many factors, including the hydrograph shape (e.g., baseflow contribution; Bača, 2008), antecedent moisture conditions [Seeger *et al.*, 2004; Giménez *et al.*, 2012], season [Steege *et al.*, 2000; Mather and Johnson, 2014] and supply of turbidity-causing materials [Brasington and Richards, 2000; Doomen *et al.*, 2008; Rodríguez-Blanco *et al.*, 2010]. Many of these factors vary on an event-by-event basis

and, as shown in Figure 4.1, similar event hydrographs may be accompanied by notably different event turbidity. Event characteristics, including antecedent moisture, hydrologic and meteorologic characteristics, provide insight into these controlling factors, and through these event characteristics, it may be possible to predict the event turbidity.



**Figure 4.1. Three example hydrologic events from the Raritan River with similar hydrographs (blue) and different event turbidity (black). The resulting turbidity-discharge loops are shown below the time series plots. Data points are only shown on the loop plots and the loop direction is as shown.**

However, prediction of stream turbidity has been mostly related to prediction of event turbidity load/yield [Mukundan *et al.*, 2013b], mean event turbidity [Mukundan *et al.*, 2013b] or prediction of daily turbidity [Mukundan *et al.*, 2013a], often within the context of supporting reservoir turbidity models [e.g., Samal *et al.*, 2013]. Recent

research has shown some success with modeling sub-hourly stream turbidity during hydrologic events [Eder *et al.*, 2010; Mather and Johnson, 2014] and has shown that event model parameters vary on an event-by-event basis. A logical next step is to examine the extent to which we can predict the values of these turbidity event model parameters from the event characteristics. Multiple linear regression (MLR) has frequently been used for prediction of model parameters [e.g., Heng and Suetsugi, 2014] and yields information on which predictive variables influence model parameter values [Holder, 1985]. Another prediction approach that is increasingly used in hydrology is classification and regression trees (CART) [e.g., Laaha and Blöschl, 2006]. In many applications, input data are first clustered into discrete classes using cluster analysis [e.g., Mukundan *et al.*, 2013b; Sawicz *et al.*, 2014]. A particular advantage of the classification tree approach is the ability to implicitly deal with non-linearity and interactions among predictor variables [Reidy Liermann *et al.*, 2012].

The purpose of this study is to compare several approaches, including multiple linear regression and classification trees, for prediction of event-based stream turbidity. We compare the success rates for the prediction approaches and also examine which event characteristics are the most useful for prediction of stream turbidity. The knowledge gained from this study is expected to benefit future studies of operational prediction and forecasting of stream turbidity.

## 4.3 Methods

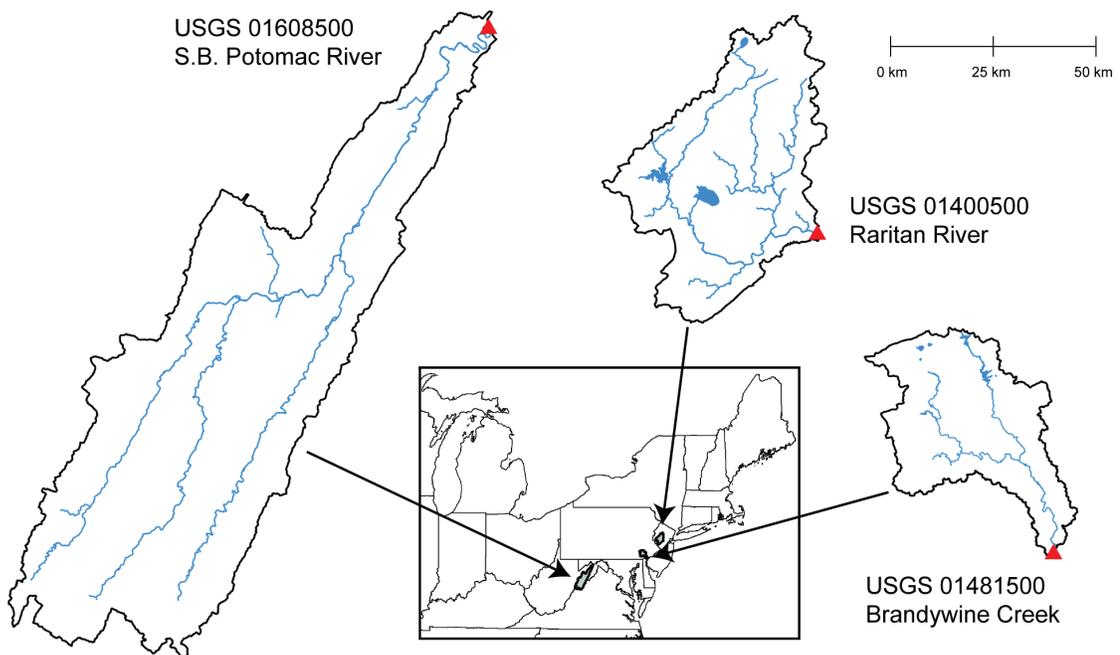
The general methodology used in this study was to first fit a turbidity event model to many hydrologic events for each stream (Sections 4.3.1 and 4.3.2). The fitted event model parameter values, along with antecedent moisture, hydrologic and meteorologic event characteristics (Section 4.3.3), were then used to build predictive relationships for event turbidity using several approaches (Section 4.3.4).

### 4.3.1 Study Catchments

This study uses data from three catchments within the Mid-Atlantic hydrologic region of the U.S. [Seaber *et al.*, 1987], as shown in Figure 4.2. This region receives an annual rainfall of approximately 1 m and precipitation has low seasonality [Neff *et al.*, 2000; Pryor and Schoof, 2008]; however, significant seasonality of streamflow is observed due to variations in evapotranspiration. The mean annual temperature over the last century was approximately 11°C [Polisky *et al.*, 2000]. The Raritan River (U.S. Geological Survey (USGS) 01400500) catchment has an area of 1269 km<sup>2</sup> and rises from 6 m elevation at the gage to a maximum elevation of 378 m with a mean slope of 5.5% [Falcone *et al.*, 2010]. This catchment is approximately half within the Piedmont physiographic province and half in the New England province [Fenneman and Johnson, 1946]. The dominant land cover classes for the Raritan River catchment are forest (40%, mostly deciduous), agriculture (28%, about equal parts pasture/hay and cultivated crops) and urban (22%) [Fry *et al.*, 2011].

The Brandywine Creek (USGS 01481500) catchment area is 813 km<sup>2</sup>, and it lies almost entirely within the Piedmont province. This catchment has a minimum elevation of 22 m, a maximum of 320 m and a mean slope of 5.6% [Falcone *et al.*, 2010]. The dominant land cover classes for the Brandywine Creek catchment are agriculture (34%,

about two-thirds pasture/hay and one-third cultivated crops), forest (33%, mostly deciduous) and urban (23%) [Fry *et al.*, 2011]. The South Branch Potomac River (USGS 01608500) catchment is larger than the other two study catchments at 3784 km<sup>2</sup>. It is positioned primarily in the Valley and Ridge province and rises in elevation from 175 m at the gage to 1480 m with a mean slope of 20.9% [Falcone *et al.*, 2010]. The dominant land cover classes for the S.B. Potomac River catchment are forest (81%, dominantly deciduous), agriculture (15%, almost all as pasture/hay) and urban (4%) [Fry *et al.*, 2011].



**Figure 4.2.** The three study catchments are located in the Mid-Atlantic hydrologic region of the U.S.; red triangles indicate location of gaging stations. Catchments are all shown at the same scale.

### 4.3.2 Event Turbidity Modeling

Discharge and turbidity data were obtained from the USGS (data codes 00060 and 63680, respectively). These sub-hourly datasets (15-min or 30-min data) were interpolated to a common 15-min time step. The data from the three catchments were divided into training and validation sets. The duration of record used for the training dataset was about five years and from the 2007-2013 time period. The validation dataset was about one year in length and was from 2013-2014. Specific details of the study data for each gage can be found in Table 4.1.

**Table 4.1. Study catchments and dataset details.**

USGS Gage	Drainage Area ( $km^2$ )	NCDC Station	Study Dataset	Events <sup>a</sup>	Date Range
USGS 01400500 Raritan River	1269	USC00283029	Training	70	2008-04-15 to 2013-04-22
			Validation	17	2013-05-01 to 2014-08-19
USGS 01481500 Brandywine Creek	813	USC00369464	Training	62	2007-10-01 to 2013-04-22
			Validation	28	2013-05-01 to 2014-08-19
USGS 01608500 South Branch Potomac River	3784	USC00467730	Training	41	2007-10-27 to 2013-04-22
			Validation	10	2013-05-01 to 2014-08-19

Note: USGS = U.S. Geological Survey; NCDC = National Climatic Data Center.

<sup>a</sup> Number of retained events. Does not include events from training datasets not used due to low *NSE*: Raritan River (8 events), Brandywine Creek (3 events) and S.B. Potomac River (3 events). See Section 4.4.1 for discussion.

Hydrologic events were identified by searching for peaks in the discharge time series using an automated peak detection program (LabVIEW, National Instruments). The event window in time was assumed to extend from  $0.5N$  days prior to the discharge peak to  $N$  days following the peak, where

$$N = A^{0.2} \quad (4.1)$$

and  $A$  is the drainage area in square miles and  $N$  is the duration of surface runoff in days [Sloto and Crouse, 1996].

A power-law based event model with supply and discharge slope terms was used to model the event turbidity as a function of the event discharge [*Mather and Johnson, 2014*]:

$$TURB(t) = aQ(t)^b e^{-rV(t)} + c \frac{dQ(t)}{dt} \quad (4.2)$$

$$\text{where, } V(t) = \int_0^t Q(t)dt \cong \sum_{i=0}^t Q_i \Delta t \quad (4.3)$$

and  $a$ ,  $b$ ,  $c$  and  $r$  are event model parameters (described below). This model, specifically the terms associated with  $c$  and  $r$ , allow the turbidity values on the rising and falling limbs of the discharge hydrograph to be different, representing the looped behavior often observed between turbidity and discharge (Figure 4.1). Prior to fitting the event model, the discharge was scaled to catchment drainage area, resulting in discharge units of  $(m^3/s)/km^2$ . Turbidity is considered here as a “concentration” and has units of (FNU) [*Anderson, 2005*]. The resulting units for event model parameters  $a$ ,  $b$ ,  $c$  and  $r$  are  $(FNU(km^2(s/m^3))^b)$ ,  $(-)$ ,  $(FNU(km^2s)(s/m^3))$  and  $(km^2/m^3)$ , respectively. Event model parameter  $r$  was not permitted to take negative values. The local slope of the hydrograph ( $dQ/dt$ ) was calculated using the central difference method and was smoothed using a Savitzky-Golay filter [*Savitzky and Golay, 1964*].

In this study we simplified the event model structure on an event-by-event basis by eliminating irrelevant model terms to arrive at an appropriately-parsimonious model [*Linsley, 1982; Gupta et al., 2014*]. To accomplish this, the following procedure was performed for each event individually. The initial event model (equation (4.2)) was fit to the event data with non-linear curve fitting using the trust-region dogleg (TRDL)

algorithm and least-squares minimization (LabVIEW, National Instruments). If the 95% confidence interval for event model parameter  $c$  and/or  $r$  contained zero, then  $c$  and/or  $r$  were removed from the model (by setting the parameter value to zero), and the model was refit with the remaining model terms to obtain the final event model.

Once the final model for each event was determined,  $a$  was transformed into  $a_0$  using a discharge-dimensioned, gage-specific scaling factor,  $Q_0$ , following the method of *Mather and Johnson* [2014].

$$a_0 = aQ_0^b \quad (4.4)$$

This transformation decorrelates the power law parameters  $a$  and  $b$ , and simplifies the units of  $a$  from  $(FNU(km^2(s/m^3))^b)$  to  $(FNU)$  (i.e., the units of  $a$  are no longer dependent on  $b$ ).

The Nash-Sutcliffe Efficiency (*NSE*) [*Nash and Sutcliffe*, 1970] was used to evaluate the goodness-of-fit of the model for each event. The *NSE* is unitless and ranges from  $-\infty$  to 1, with values greater than zero indicating that the fit is better than the mean of the observations, and a value of 1 indicates a perfect fit.

### 4.3.3 Event Characteristics

Daily discharge, precipitation and temperature datasets were used to calculate the antecedent moisture, hydrologic and meteorologic characteristics for each event (collectively referred to herein as the “event characteristics”). Daily mean discharge data was obtained from the USGS, and daily meteorologic data (precipitation and temperature) was obtained from the NOAA National Climatic Data Center (NCDC) from

the available station closest to the centroid of the catchment with a suitable dataset. The NCDC station used for each catchment is listed in Table 4.1. A total of 32 antecedent moisture, hydrologic and meteorologic characteristics were calculated for each event and are described here.

Catchment antecedent moisture conditions (AMCs) are widely considered to exert considerable control on event-to-event variability in catchment response to precipitation events, specifically in runoff generation [*James and Roulet, 2009*]. AMC has been acknowledged as significant in discerning the direction and extent of turbidity/sediment-discharge loops in nearly all studies in which it was considered [e.g., *Seeger et al., 2004; Bowes et al., 2005; Eder et al., 2010; cf., Oeurng et al., 2010*]. There is, however, less agreement with regard to the most suitable surrogate for AMC (e.g., does it vary between catchments or is a single surrogate sufficient) [*Ali and Roy, 2010*]. For this reason, we compute four types of AMC indicators in this study:

1. specific discharge at the beginning of the event ( $Q_{beg}, (m^3/s)/km^2$ ),
2. precipitation and discharge over a set time period preceding event start:
  - a. sum of the daily precipitation for 1, 2, 5 and 10 days prior to event start ( $AP1, AP2, AP5$  and  $AP10, mm$ ),
  - b. average daily specific discharge for 1, 2, 5, 10 and 30 days prior to event start ( $AQ1, AQ2, AQ5, AQ10$  and  $AQ30, (m^3/s)/km^2$ ),
  - c. maximum daily specific discharge for 5, 10 and 30 days prior to event start ( $MQ5, MQ10$  and  $MQ30, (m^3/s)/km^2$ ),
3. days since the occurrence of a particular precipitation or discharge threshold:
  - a. days since measureable precipitation ( $DSP, days$ ),

- b. days since daily precipitation of at least 10, 20 and 50 mm ( $DSP10$ ,  $DSP20$  and  $DSP50$ , days),
  - c. days since daily discharge of at least 1, 2, 3, 4 and 5 times the mean annual daily discharge ( $DSQ1$ ,  $DSQ2$ ,  $DSQ3$ ,  $DSQ4$  and  $DSQ5$ , days),
4. continuous accounting of precipitation and discharge decaying in influence over time, as shown in the equations:

$$API_i = kAPI_{i-1} + P_i \quad (4.5)$$

$$AQI_i = kAQI_{i-1} + Q_i \quad (4.6)$$

where  $API_i$  and  $AQI_i$  represent the antecedent precipitation and discharge indices on day  $i$ , and  $P_i$  and  $Q_i$  represent the daily precipitation and mean daily discharge on day  $i$ , respectively [Kohler and Linsley, 1951; Lewis et al., 2001]. In this study, day  $i$  was taken as the day preceding event start. The values of parameter  $k$  used were 0.87055 and 0.93303, corresponding to a half-life of 5 and 10 days, respectively ( $API05$ ,  $API10$ ,  $mm$  and  $AQI05$ ,  $AQI10$ ,  $(m^3/s)/km^2$ ).

Event peak specific discharge ( $Q_p$ ,  $(m^3/s)/km^2$ ) is a basic descriptor of overall event magnitude. It has been hypothesized that, because peak discharge depends on both the intensity and amount of rainfall, it can serve as a surrogate for rainfall energy [Rankl, 2004; Duvert et al., 2012], which is expected to relate to runoff-associated turbidity. The relationship between the specific discharge at the beginning of the event and the peak specific discharge ( $Q_{beg}/Q_p$ ) was also included. The event characteristics  $Q_p$  and  $Q_{beg}$  were determined using sub-hourly data, while all other event characteristics were determined from daily datasets.

Precipitation ( $PRCP$ ,  $mm$ ) was calculated as the sum of the daily precipitation from the day of event start to the day of peak discharge (which could be the same day). Ideally, sub-hourly or hourly precipitation would have been used and resulted in higher quality knowledge of event precipitation. However, these more time-resolved datasets were found to have frequent missing portions during events, and were sparser in spatial coverage (e.g., sometimes the nearest station would have been outside of the catchment). The mean air temperature ( $TEMP$ ,  $^{\circ}C$ ) for the day of peak discharge was included primarily to assess seasonal effects [Guy, 1964] and was estimated here as the mean of the daily minimum and daily maximum temperature. Finally,  $YEAR$  was included to account for any basic trending in time.

#### **4.3.4 Turbidity Prediction Approaches**

Four different approaches with increasing complexity (both in computation and model form) were used to predict turbidity event. The simplest approach used here was the basic power law rating curve (referred to as the “Power Law” approach) where the values of  $c$  and  $r$  in equation (4.2) are zero. For this approach, the median values of  $a_0$  and  $b$  from fitting the model to each event in the training dataset for each stream were used for prediction of the validation events (i.e., the same event model parameter values were used for all predicted events for each stream). The second approach is similar, however the full equation (4.2) event model was used (referred to as the “Median Value” approach). In this approach, the median values of  $a_0$ ,  $b$ ,  $c$  and  $r$  from the training events for each stream were used to predict turbidity of the validation events. The remaining two approaches, referred to as “Cluster/Classification” and “Regression” are described next.

#### 4.3.4.1 Cluster Analysis and Classification Trees

Cluster analysis was used to group events based on their modeled turbidity-discharge behavior (i.e., event parameters  $a_0$ ,  $b$ ,  $c$  and  $r$ ). An agglomerative hierarchical algorithm was used, and the results are presented in a dendrogram showing the hierarchy of clusters (Minitab 16 Statistical Software). The number of clusters (i.e., the cutting threshold) was determined using a common similarity level for all three streams [Olden *et al.*, 2012]. Clustering was performed for each stream separately, and the input data used were the fitted values for  $a_0$ ,  $b$ ,  $c$  and  $r$  for each event in the training dataset. The input data were standardized, and model parameter  $a_0$  was log-transformed. Euclidean distance was used as a distance metric, and Ward's method was used as the linkage criterion [Monk *et al.*, 2007; Sawicz, 2009; Olden *et al.*, 2012].

In order to predict the cluster membership for new (i.e., validation) events, classification trees were constructed for each stream. Classification trees are a method for building predictive models from data [e.g., Olden *et al.*, 2008; Reidy Liermann *et al.*, 2012] and are used here to predict the relevant cluster, and corresponding model parameter values, for an event using the event characteristics. The cluster membership for the training events and the event characteristics associated with these events were used as the input data for constructing the trees. The trees were developed using binary recursive partitioning, which involves repeatedly splitting the dataset (i.e., training events) into mutually exclusive groups based on a single event characteristic at each step (MATLAB R2013b, MathWorks, Inc.). The size of each classification tree was controlled by the constraint that each parent node of the tree was not permitted to have less than 10 observations (i.e., events). Potential interactions of event characteristics are handled

implicitly with this method and results are unaffected by monotonic transformations of event characteristics [De'ath and Fabricius, 2000; Olden et al., 2008; Reidy Liermann et al., 2012].

#### **4.3.4.2 Regression Analysis**

Multiple linear regression was used to predict values of  $a_0$ ,  $b$ ,  $c$  and  $r$  on an event-by-event basis using the event characteristics (Section 4.3.3). The event characteristics were transformed to follow a normal distribution where appropriate, this resulted in the use of  $\text{Log}(1 + x)$  transformation for  $AP1$ ,  $AP2$ ,  $AP5$  and  $AP10$ , no transformation for  $TEMP$  and  $YEAR$ , and  $\text{Log}(x)$  transformation for all remaining event characteristics. Only events from the training datasets were used to generate regression models.

Stepwise regression was used to generate regression models by adding or removing significant event characteristics. The regression models were then screened for multicollinearity using the Variance Inflation Factor (VIF) [Helsel and Hirsch, 2002]. If any model term had a VIF greater than five, the model was revised either by removal of an event characteristic or replacement of an event characteristic with one less correlated. An additional metric used here during regression model development was  $R_{pred}^2$ , for which each observation is systematically removed and the regression model recalculated (with the same event characteristics present) to determine how well the model predicts the removed observation (Minitab 16 Statistical Software). It is used here to indicate how well the regression model may predict turbidity event model parameters for new observations (i.e., validation events). Each event characteristic that was added to the

regression equation as part of the stepwise process was only added if it also resulted in at least a 1% increase in the model  $R_{pred}^2$ .

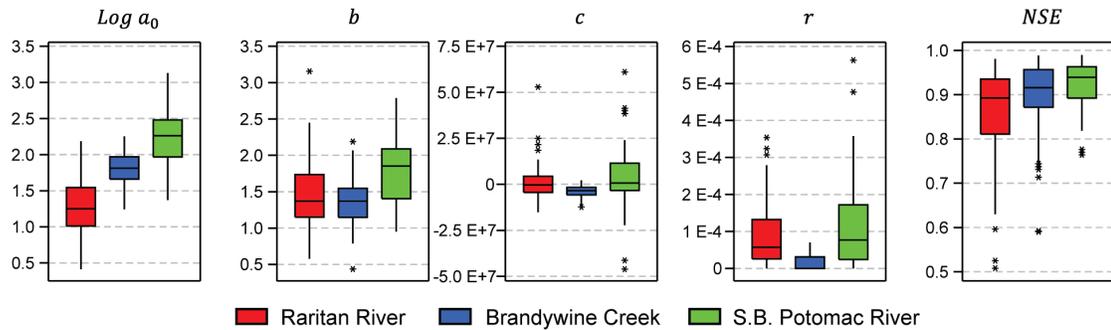
## 4.4 Results and Discussion

### 4.4.1 Study Events

Hydrologic events were identified in both the training and validation datasets for each stream. As described above, the event model (equation (4.2)) was fit to each event individually (both training and validation), resulting in a unique best-fit event model parameter set of  $a_0$ ,  $b$ ,  $c$ ,  $r$  and goodness-of-fit metric  $NSE$  for each event from each stream. The number of events obtained from the training and validation datasets for each stream are shown in Table 4.1. Events from the training dataset with  $NSE$  lower than 0.5 were removed from the study as these events were considered to not have a useful predictive contribution (this resulted in removal of 14 of the original 187 training events, see Table 4.1 footnote).

Figure 4.3 summarizes the best-fit values of the event model parameters and the  $NSE$  for the training events. Model fit was generally very good, with a median event  $NSE$  of 0.893, 0.916 and 0.940 for Raritan River, Brandywine Creek and S.B. Potomac River, respectively. Event model parameter  $a_0$  follows a lognormal distribution and was log-transformed, while  $b$  is normally distributed and is shown in original form. Event model parameters  $c$  and  $r$ , however, are more problematic in distribution because both are heavily tailed and can take a value of zero. Additionally, event model parameter  $c$  can take both positive and negative values. Due to these properties, no normalizing transformation was used for  $c$  and  $r$ . The median values of the event model parameters,

as used in the Median Value prediction approach, and the median event model parameter values for the Power Law simplification are shown in Table 4.2.



**Figure 4.3. Boxplots of best-fit event model parameter values for the events in the training datasets:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU( $\text{km}^2\text{s}$ )( $\text{s}/\text{m}^3$ )),  $r$  ( $\text{km}^2/\text{m}^3$ ) and  $NSE$  (-). Boxplot boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to the nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside the upper and lower whiskers (outliers); boxplot colors indicate stream.**

Raritan River had the lowest range of event  $NSE$  values, which is likely due in part to the complexity of the event turbidity in this stream. Raritan River events often displayed early, short-duration spikes in turbidity followed by a more typical rise and fall in turbidity (for example, see third event (right) on Figure 4.1). These brief, high values of turbidity were rarely modeled well; however, the remainder of the event was often well-fit by the models. Brandywine Creek events, when viewed as loops, were typically narrower than the other two streams and were often counterclockwise in direction (i.e., specific discharge increasing prior to turbidity, which is less common). This is consistent with small, negative values of  $c$  and near-zero values of  $r$ , as shown in Figure 4.3. The S.B. Potomac River had the greatest range of  $c$  and  $r$  values. This may result from its larger catchment size, which causes spatially non-uniform precipitation to produce a

broad range of event shapes. Boxplots for the validation events are shown in Figure C 1 of Appendix C.

**Table 4.2. Turbidity event model parameter values used for Power Law and Median Value prediction approaches.**

Stream		$Q_0$	$a_0$	$b$	$c$	$r$
Raritan River	Power Law	0.0201	9.58	1.476		
	Median Values	0.0186	14.92	1.335	-1657720	0.000055
Brandywine Creek	Power Law	0.0502	35.19	1.350		
	Median Values	0.0636	54.85	1.345	-3233070	0
S.B. Potomac River	Power Law	0.0328	27.28	2.057		
	Median Values	0.0608	182.60	1.850	668100	0.000077

Note: Units are as follows:  $Q_0$  ( $(m^3/s)/km^2$ ),  $a_0$  (FNU),  $b$  (-),  $c$  (FNU( $km^2s$ )( $s/m^3$ )), and  $r$  ( $km^2/m^3$ ).

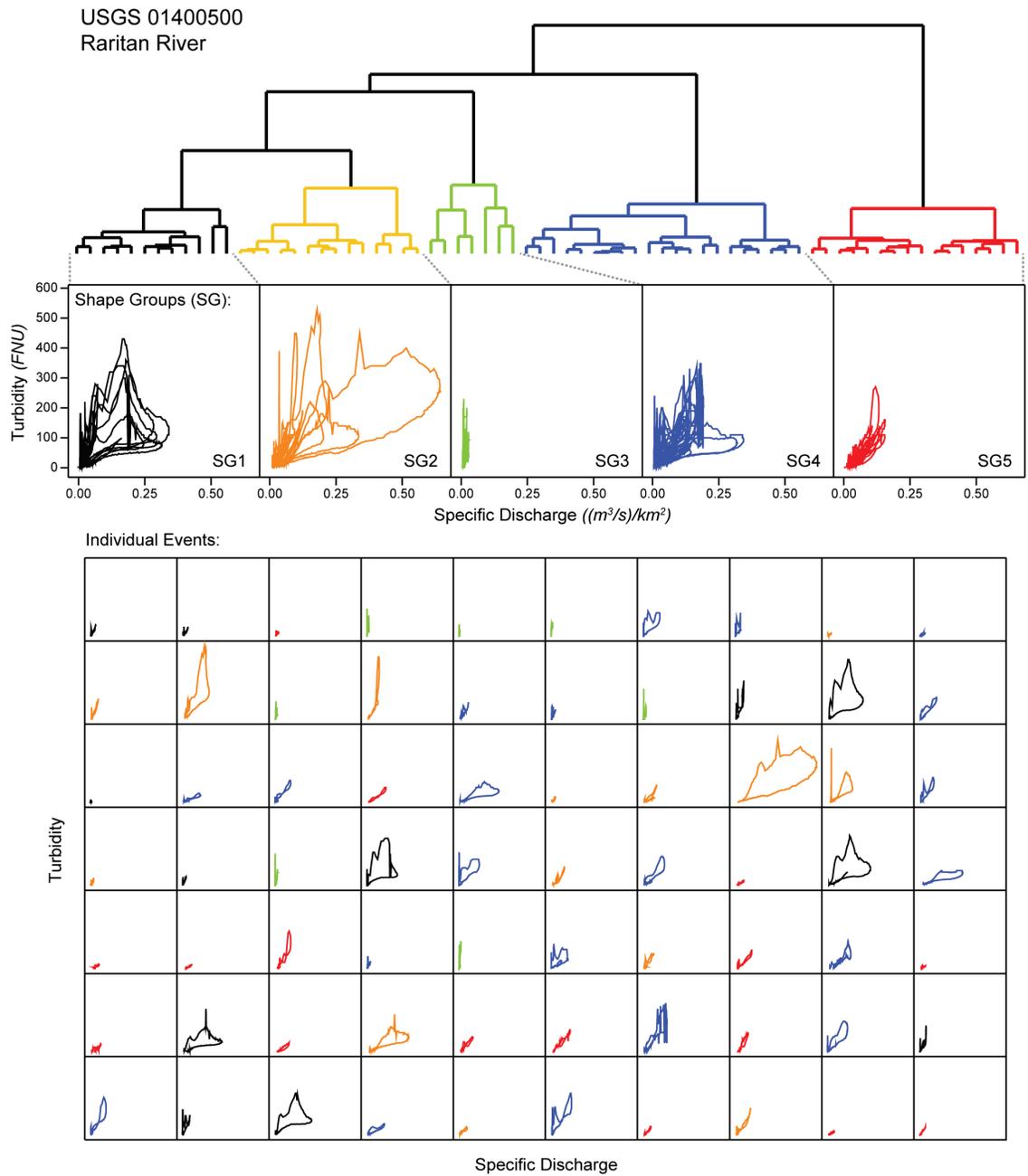
## 4.4.2 Turbidity Prediction using Cluster Analysis and Classification Trees

### 4.4.2.1 Cluster Analysis

The fitted parameter values shown in Figure 4.3 (with sets of  $a_0$ ,  $b$ ,  $c$  and  $r$  collectively describing each single event) were used to cluster the events into shape groups with similar event model parameter values. The results for cluster analysis of Raritan River training events are shown in Figure 4.4. The results for Brandywine Creek and S.B. Potomac River are shown in Figure C 2 and Figure C 3 located in Appendix C.

Clustering of the 70 training events from the Raritan River produced five visually distinct shape groups (SGs) (Figure 4.4). The mean model parameter values associated with each shape group are shown in Table 4.3. SG1 (black) and SG2 (orange) are the most similar of the five clusters (as indicated by the shortness of the vertical distance on the dendrogram to the line above connecting the two clusters). Numerous events from these two groups show considerable time-separation between the points of maximum turbidity and maximum discharge and, together with the starting/ending points, results in the loop having a triangle shape. SG3 (green) has the fewest number of events, but the shape of the events in this group

is the most visually distinct of the five groups. Events from SG3 have a small discharge magnitude but a comparably large turbidity response. The within-cluster similarity for events from SG3 is the overall lowest of the five groups, as indicated by the height of the within-cluster branches (horizontal lines) connecting individual events on the dendrogram. The small and mid-sized events from SG4 (blue) and most events from SG5 (red) share many visual characteristics (e.g., many have a “figure-eight-like” shapes); however, they are actually quite different. Events from SG4 rarely display a true figure-eight behavior (i.e., the upper and lower loop lines cross over, making part of the loop a clockwise loop and part counterclockwise); rather, the upper line drops towards and sometimes touches the lower line, but both parts of the loop are clockwise in direction. In contrast, most SG5 events are true figure-eight shapes, typically beginning in a clockwise direction, then transitioning to counterclockwise, and occasionally switching to clockwise again forming three sub-loops during a single event.



**Figure 4.4. Dendrogram showing cluster analysis results and the events associated with each cluster (shape group) for Raritan River. The events are shown as turbidity vs. discharge loops and are colored by shape group. The small plots of individual events all have the same axis scaling.**

For the Cluster/Classification prediction approach used here, each shape group was assigned a set of event model parameters. We assigned these as the mean value of the event model parameters for the training events that comprise the shape group, as shown in Table 4.3.

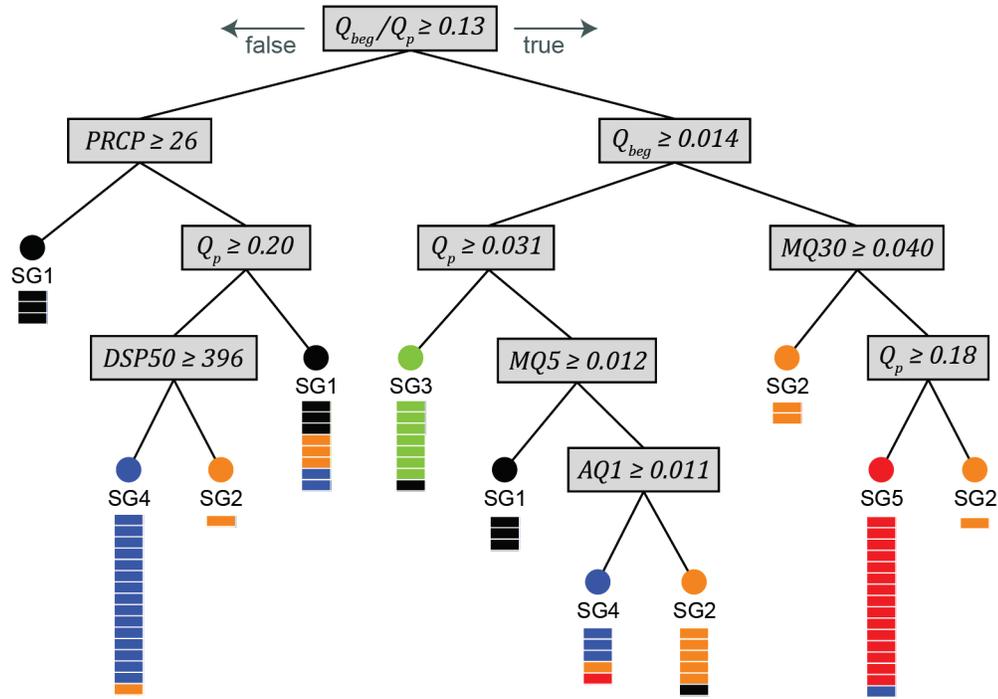
**Table 4.3. Turbidity event model parameter values used for the Cluster/Classification prediction approach.**

Stream	Shape Group (SG)	$a_0$	$b$	$c$	$r$
Raritan River $Q_0 = 0.0186 (m^3/s)/km^2$	SG1	39.33	1.203	-2355977	0.000201
	SG2	15.75	1.652	-5540576	0.000119
	SG3	74.10	2.265	17170149	0.000101
	SG4	23.73	1.044	4840730	0.000054
	SG5	6.45	1.627	-1813195	0.000022
Brandywine Creek $Q_0 = 0.0636 (m^3/s)/km^2$	SG1	96.70	1.129	-4057215	0
	SG2	88.21	1.518	-2196542	0.000003
	SG3	65.34	1.180	-9468748	0.000003
	SG4	100.46	2.050	-8100507	0
	SG5	34.85	1.207	-2893783	0.000003
	SG6	35.62	1.664	-5437535	0.000005
	SG7	82.07	1.281	-1861779	0.000045
S.B. Potomac River $Q_0 = 0.0608 (m^3/s)/km^2$	SG1	549.00	2.286	3455956	0.000363
	SG2	345.70	1.641	-16984660	0.000197
	SG3	123.50	2.119	863836	0.000046
	SG4	66.53	2.073	46938232	0
	SG5	167.00	1.388	9719963	0.000068

Note: Units are as follows:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)), and  $r$  (km<sup>2</sup>/m<sup>3</sup>).

#### 4.4.2.2 Classification Trees

The results from cluster analysis of the training events from each stream were used along with individual event characteristics to build predictive models for each stream using classification trees. The results for Raritan River are shown in Figure 4.5 and the results for Brandywine Creek and the S.B. Potomac River are shown in Figure C 4 and Figure C 5 in Appendix C. The classification tree results for Raritan River are discussed in depth here, and the reader is directed to Appendix C for comparison to the classification trees for Brandywine Creek and S.B. Potomac River.



**Figure 4.5. Classification Tree for Raritan River. Terminal node (circles) class identity colored by shape group, as shown in Figure 4.4. Membership for training events shown with rectangles representing individual events under each terminal node. Only 69 of the total 70 events are shown due to one event at the *PRCP* branch having missing precipitation data.**

Branching points of the Raritan River classification tree are based on event characteristics from several categories, including: discharge magnitude, AMC (from both discharge and precipitation) and event precipitation. The first split on the tree is based on  $Q_{beg}/Q_p$ , which serves as a basic surrogate measure of event baseflow index. At this tree split, SG3 (green) and SG5 (red) are entirely on the right branch with higher values of  $Q_{beg}/Q_p$ , while most of SG4 (blue) events are to the left with lower values. The next branch on the right side of the tree is based on the discharge at the beginning of the event ( $Q_{beg}$ ) and completely divides the SG3 (green) events (lower values) from SG5 (red) (higher  $Q_{beg}$ ). Subsequent branching on both sides of the tree is mostly related to isolating events from SG1 (black) and SG2 (orange).

Events from SG1 (black) and SG2 (orange) proved to be more difficult to correctly classify, and accounted for 7 of the 11 misclassifications for Raritan River. As discussed above, SG1 and SG2 are the most similar amongst the five shape groups, and these results suggest there may be insufficient information in the event characteristics dataset to fully account for the division in event membership between these two shape groups. Overall, the classification trees correctly classified 84% of the training events for Raritan River, 83% for Brandywine Creek and 85% for S.B. Potomac River (calculated based on re-substitution of the event characteristics for the training events into classification tree (MATLAB R2013b, MathWorks, Inc.)).

#### **4.4.3 Turbidity Prediction using Regression Analysis**

Event turbidity was also predicted using a multiple linear regression approach. The results of this analysis are shown in Table 4.4. The number of event characteristics in the regression equations varied from one to four. Raritan River and Brandywine Creek have many similarities in the event characteristics present, particularly in the regression equations for event model parameters  $a_0$  and  $b$ . The S.B. Potomac River regression equations are largely based AMC and do not include  $TEMP$ , indicating that if seasonality is a significant factor in this catchment, the AMC terms must be sufficient to account for it.

The most commonly appearing event characteristics, in decreasing order, were  $DSQx$ ,  $TEMP$ ,  $Q_p$  and  $Q_{beg}/Q_p$  ( $DSQx$  includes  $DSQ1 - DSQ5$ ). A total of 10 of the 12 equations (4 event model parameters x 3 streams) have at least one AMC-related event characteristic present. This suggests the strongest overall predictive event characteristics for the event model parameter values were antecedent moisture conditions, season and

event discharge. Comparison plots of fitted versus predicted values for each event model parameter and each stream are shown in Figure C 6 – Figure C 8.

**Table 4.4. Regression equations for turbidity event model parameters.**

Stream	Equation	$R^2$	$R^2_{PRED}$
Raritan River $Q_0 = 0.0186 (m^3/s)/km^2$	$Log a_0 = -0.986 - 0.847 Log Q_{beg} + 0.136 TEMP$ $+ 0.230 Log Q_p - 0.410 Log AQ30$	74.2%	69.9%
	$b = 3.457 + 0.414 Log DSQ3 - 0.756 Log Q_p$ $+ 1.829 Log Q_{beg} + 0.0120 TEMP$	55.4%	47.7%
	$c = -9.143x10^5 - 7.253x10^6 Log Q_p + 1.010x10^7 Log DSQ1$ $+ 1.109x10^7 Log MQ10 + 2.736x10^5 TEMP$	37.2%	24.1%
	$r = -7.806x10^{-6} + 6.178x10^{-5} Log DSQ3$ $+ 3.307x10^{-5} Log (1 + AP2)$	19.9%	12.1%
Brandywine Creek $Q_0 = 0.0636 (m^3/s)/km^2$	$Log a_0 = 90.698 + 0.0103 TEMP - 0.195 Log Q_{beg}/Q_p$ $+ 0.155 Log DSP50 - 0.0445 YEAR$	62.2%	54.7%
	$b = 1.581 + 0.00811 TEMP + 0.707 Log Q_{beg}/Q_p$ $+ 0.178 Log DSP50$	56.8%	48.8%
	$c = -1.502 x10^7 - 7.557x10^6 Log AQ30$ $- 2.208x10^6 Log DSQ2$	17.6%	11.1%
	$r = -1.601x10^{-5} + 2.052x10^{-5} Log PRCP$	12.9%	6.5%
S.B. Potomac River $Q_0 = 0.0608 (m^3/s)/km^2$	$Log a_0 = 1.750 + 0.321 Log DSQ2 + 0.214 Log (1 + AP2)$	49.8%	40.9%
	$b = 2.474 + 0.847 Log Q_{beg}/Q_p$	45.3%	40.6%
	$c = 3.881 x10^7 - 1.050x10^7 Log DSQ2 - 2.091x10^7 Log API05$	27.3%	10.0%
	$r = -2.568x10^{-5} + 0.000103 Log DSQ1$ $+ 6.416x10^{-5} Log DSP20$	35.1%	25.1%

Note: Variance Inflation Factor (VIF) values for all equation terms were less than 5. The event characteristics  $AP5$  and  $AP10$  were not used as candidates for the S.B. Potomac River equations due to missing values in the dataset. See Section 4.3.4.2 for a description of  $R^2_{PRED}$ .

Some parallels can be seen between each stream’s regression equations and classification tree. Brandywine Creek for example, has  $PRCP$  at the top of the classification tree (Figure C 4) and this divides the high- $r$  events (SG7, Table 4.3) to the right side of the tree (high- $PRCP$ ), and similarly,  $PRCP$  is the event characteristic present in the regression equation for event model parameter  $r$  (Table 4.4). Likewise, the tree for the S.B. Potomac River (Figure C 5) begins with a division based on  $Q_{beg}/Q_p$  which splits the low- $b$  groups (SG2 and SG5) from the high- $b$  groups (SG1, SG3 and SG4),

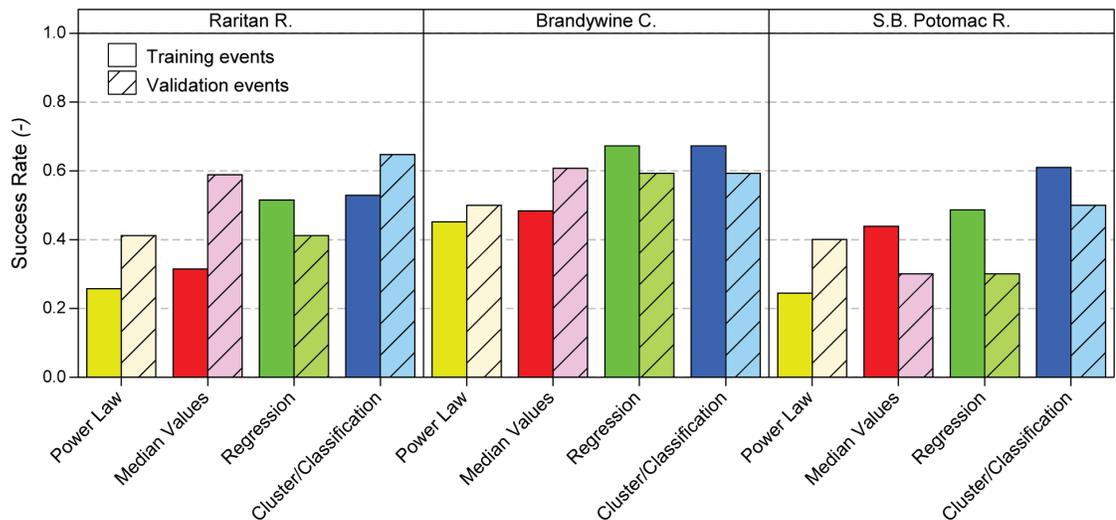
coinciding with the regression equation for event model parameter  $b$ . For the Raritan River, the similarity between the classification tree and regression equations is less clear, likely because this stream does not have any single-predictor regression equations, and therefore, did not have any individual event model parameters highly related to a single branch (i.e., event characteristic) on the classification tree.

#### **4.4.4 Comparison of Turbidity Prediction Approaches**

To evaluate the overall skill of the event turbidity prediction approaches, the “success rate” of each approach was calculated. Success rate is defined here as the number of acceptable predictions divided by the total number of predictions [Arsenault and Brissette, 2014], and for this study, an event prediction was considered acceptable if the  $NSE$  of the predicted turbidity was at least 70% of the  $NSE$  resulting from directly fitting the model to the observed turbidity. Figure 4.6 shows the success rate for the training and validation events for each turbidity prediction approach.

The Power Law and Median Value approaches used fixed values of the event model parameters and, as expected, have lower success rates than the Regression and Cluster/Classification approaches, which determined event model parameters on an event-by-event basis. The Cluster/Classification approach generally had the highest success rate, although for Brandywine Creek, the Regression approach was equally good. The fixed parameter value prediction approaches (Power Law and Median Value) were nearly as successful for Brandywine Creek as the approaches that vary parameter values by event (Regression and Cluster/Classification). This is likely because Brandywine Creek displayed the least looped behavior and the lowest event-to-event model parameter variability of the three streams examined here (see Figure 4.3). The success rates for the

S.B. Potomac River validation events were generally much lower than for the training events. This is likely influenced by the small number of validation events from this stream (10 events). The Power Law and Median Value approaches for the Raritan River had substantially higher success rates for the validation events over the training events. This is likely because there proportionally fewer number of small discharge events in the validation dataset compared to the training dataset (not shown, but see discussion that follows).



**Figure 4.6. Success Rate for training and validation events using each prediction approach.**

Figure 4.7 shows the success rate by discharge event size ( $Q_p$ ) and season for all events (training and validation combined). In general, the relationship between success rate and event size (Figure 4.7a) is more prominent than between success rate and season (Figure 4.7b) (although the two are related due to seasonality in streamflow). For Raritan River and Brandywine Creek, success rates for all prediction approaches are highest for events with “middle-range”  $Q_p$ , and show significant decreases in success rate for the

larger and smaller discharge events. The success rate for S.B. Potomac River appears less sensitive to  $Q_p$ . There are a greater number of events associated with each of the middle-range  $Q_p$  points on the plot (see Figure 4.7), so the values of these points are likely more robust than the largest and smallest  $Q_p$  bins. Additionally, and likely more importantly, the high population of events in the middle-range  $Q_p$  means the event model parameter prediction (for all approaches) is heavily controlled by the behavior of “average-sized” events. Thus, we would expect the predictions to be more successful for events of this size. Overall, the Cluster/Classification approach appears to be less sensitive to  $Q_p$  than the other prediction approaches.

Figure 4.8 uses Taylor diagrams [Taylor, 2001; Nijssen *et al.*, 2003] to compare the correlation coefficient, normalized standard deviation and root-mean-square deviation (RMSD) measures of the four prediction approaches for the validation events. The normalized standard deviation is the standard deviation of the predicted values of all the data points within the event divided by the standard deviation of the observed turbidity values for the data points, and is represented on the diagram as the radial distance from the origin. The correlation between observed and predicted event turbidity data points is plotted along the arc of the diagram, while the RMSD is proportional to the distance between the prediction data marker and the observed. The observed event turbidity marker (OBS) is positioned at a unit distance from the origin, along the horizontal axis.

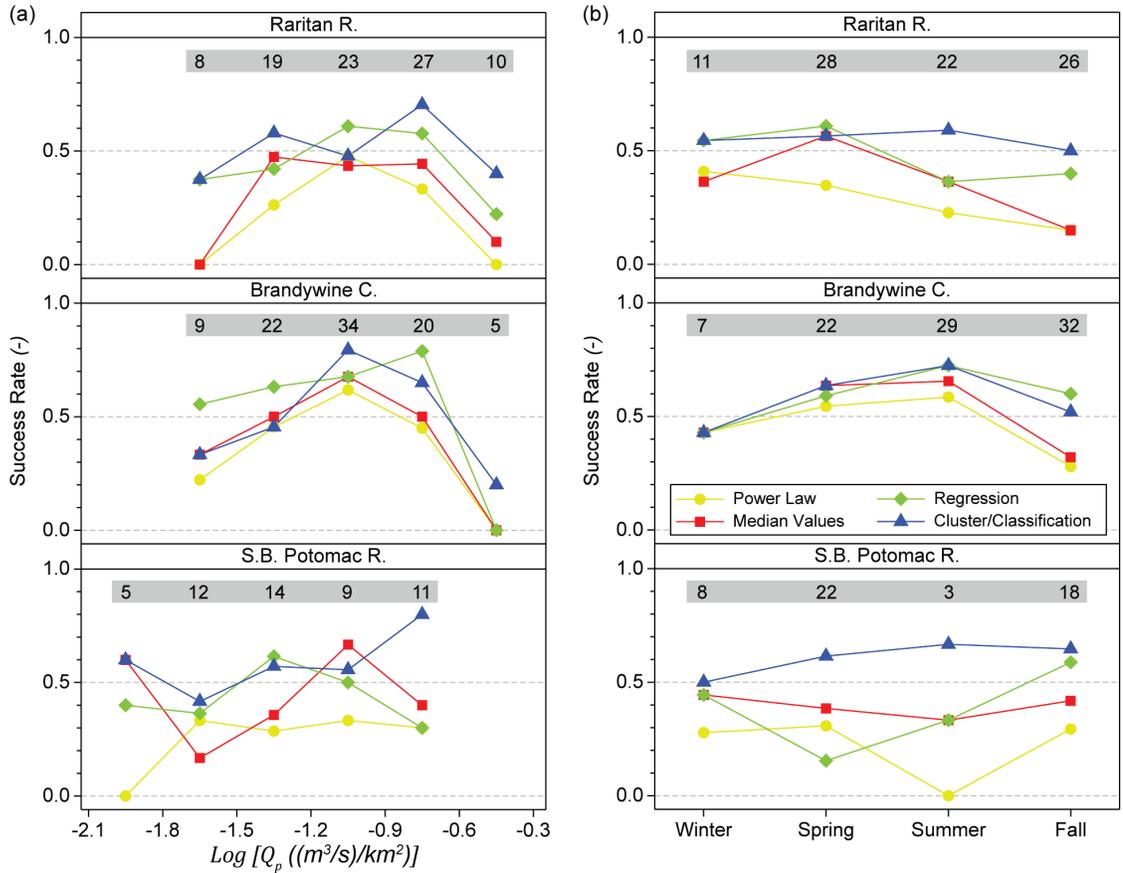
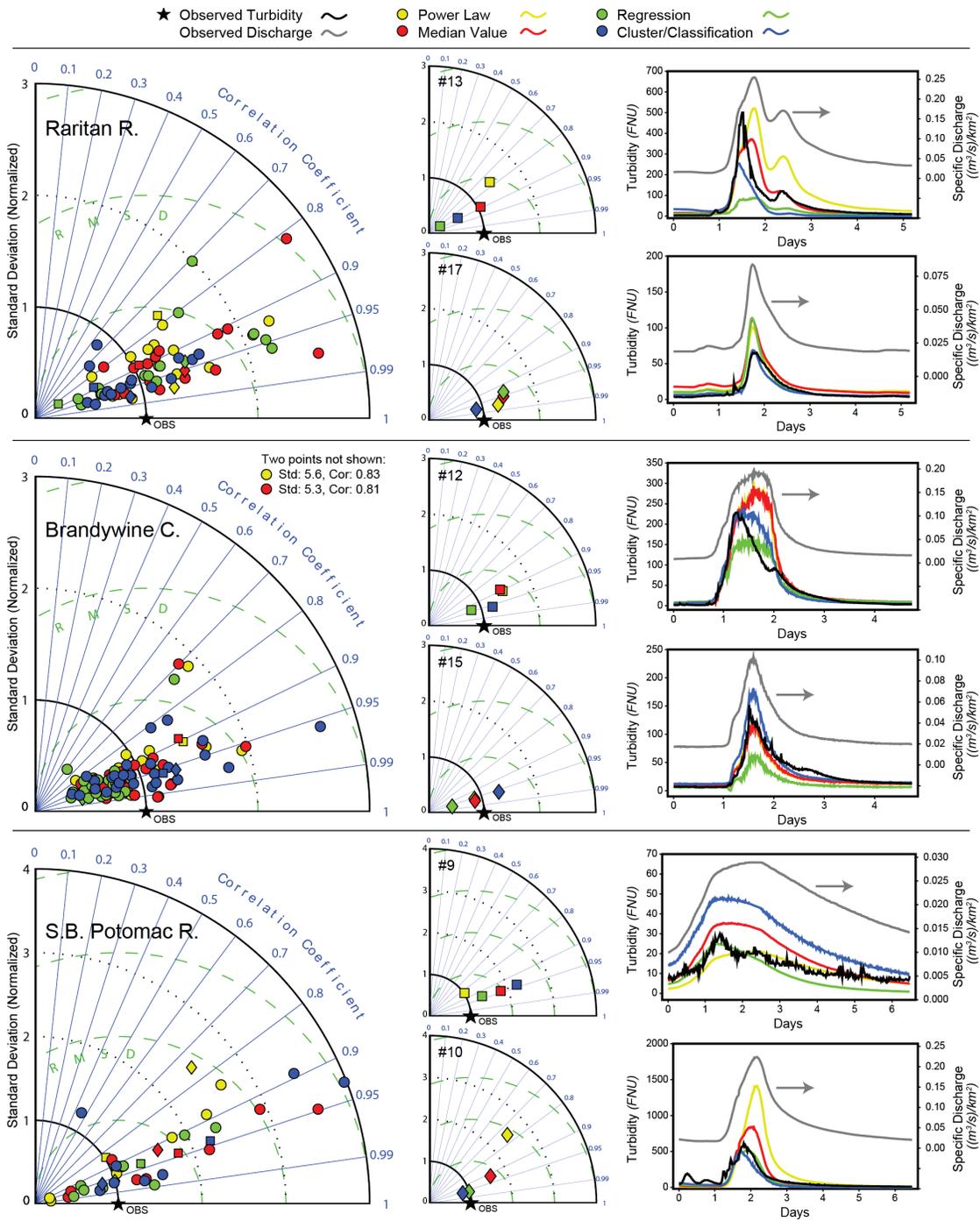


Figure 4.7. Success rate by prediction approach for the full dataset (training and validation combined). (a) Relationship with event peak specific discharge ( $Q_p$ ); events are binned into 0.3-intervals on the x-axis, centered on the data markers shown. (b) Relationship with season event occurred; Winter is Jan/Feb/Mar, Spring is Apr/May/Jun, Summer is Jul/Aug/Sep and Fall is Oct/Nov/Dec. Numbers in gray bars indicate the number of individual events represented by each point below.



**Figure 4.8. Taylor diagrams of the validation events using all four prediction approaches. Each data point represents an event predicted using the approach indicated by the marker color. Left diagrams for each stream show all validation events, middle diagrams are two example events from each stream and time series plots on the right correspond to the two example events shown in the middle. Square and diamond markers on left Taylor diagrams indicate the events used as examples in middle and right. RMSD is the root-mean-square error and OBS is the observed event turbidity reference point.**

For all three streams and four prediction approaches, most events sit in the wedge between a correlation of 0.8 and 0.99. For Raritan River, predicted event turbidity using the Cluster/Classification approach (blue) is mostly positioned along the 0.95 correlation line and is clearly grouped closer to the observed (OBS) than any of the other prediction approaches. The Regression approach for Brandywine Creek resulted in nearly all predictions having a standard deviation less than the observed (less than one). This indicates that the predictions using this approach did not capture the full turbidity behavior throughout the event. For all events and prediction approaches, the match between predicted and observed turbidity during the baseflow portions of the event (i.e., at the beginning and during the late recession) was quite good. As a result, a standard deviation less than the observed can only occur if the peak portion of an event is under-predicted. Upon further examination, it was found that there was a decreasing trend in values of event model parameter  $a_0$  for Brandywine Creek over the training dataset time period; however, this trend leveled off over the validation time period. The predictions of event model parameter  $a_0$  were too low and caused the under-predictions produced by the Regression approach (see also Figure C 7). It is interesting to note that the Regression and Cluster/Classification approaches were similarly successful for Brandywine Creek when measured by the success rate (Figure 4.6). However, as discussed above and shown in Figure 4.8, the Regression approach consistently under-predicts turbidity whereas the Cluster/Classification approach is balanced between over- and under-predicting. The failure of the success rate to detect this bias in the predictions is due to the fact that *NSE* does not differentiate the direction (sign) of deviation from the observed.

## 4.5 Conclusions

For most catchments, including the Raritan River and S.B. Potomac River in this study, it is expected that satisfactory prediction of event turbidity will require an approach where event model parameter values are adjusted on an event-by-event basis in response to the antecedent moisture, hydrologic and meteorologic characteristics of the event. The results from both the Regression and Cluster/Classification approaches suggest that the antecedent moisture conditions (AMC) of the catchment are important in the prediction of turbidity event model parameters. AMC-related event characteristics, particularly those derived from antecedent discharge, were common in both the regression equations and classification trees for all streams studied here. This is in agreement with *Eder et al.* [2010] who suggested that event model parameters may be “predefined” by initial soil water content (antecedent moisture) and rainfall.

Turbidity prediction success rates were highest for mid-sized hydrologic events. This underscores the common problem in hydrology of underrepresentation of extreme events in datasets. Turbidity predictive relationships were controlled by the behavior of “average-sized” events and we believe this resulted in decreased prediction success for very small and large hydrologic events.

For the streams studied here, the Cluster/Classification approach produced slightly better predictions than the Regression approach. However, because the Cluster/Classification approach involves predicting a “set” of event model parameters that do not vary continuously with event characteristics, small changes in the event characteristics can cause large changes in the value of one or more event model parameters, and therefore, the event turbidity prediction. This occurs because the event

may “jump” into a different class (shape group) and associated event model parameter set. While this was not a problem for the current study, it could represent a limitation of the Cluster/Classification approach in an operational forecasting mode, and in that setting the regression approach may be more robust.

The study described here was undertaken both as a heuristic evaluation of event turbidity models and, more importantly, as an exploration and foundation for turbidity-based forecasting of surface-water quality. Approximately two-thirds of the U.S. population relies on surface waters as a source of drinking water and of these, about two-thirds of surface water intakes are in rivers [*Wickham et al.*, 2011]. As a consequence of this and other stressors, rivers will be increasingly challenged in the future to provide an adequate quantity and quality of public drinking water [*Levin et al.*, 2002]. As a result, prediction—and future forecasting—of turbidity may be key components for insuring safe drinking water and the overall management of water resources.

# CHAPTER 5

---

## Forecasting Turbidity during Streamflow Events for Two Mid-Atlantic U.S. Streams<sup>6</sup>

*Amanda L. Mather and Richard L. Johnson*

### 5.1 Abstract

Despite the widespread short-term forecasting of streamflow, routine operational forecasting of water quality remains relatively limited. Of the stream water quality constituents of common interest, turbidity is particularly important to drinking water providers receiving source waters from streams, both because it impacts operational costs and is a surrogate for other water quality constituents. In this study, turbidity forecasts for two Mid-Atlantic U.S. streams (118 individual forecasts) were produced. These 3-day duration, 6-hr resolution turbidity forecasts were based on a combination of observed and forecasted streamflow, precipitation and air temperature. Turbidity forecast errors were examined, as well as the uncertainty resulting from streamflow forecast errors. The results show that the turbidity forecasts were useful compared to a persistence reference,

---

<sup>6</sup> In preparation for the Journal of Environmental Engineering.

except at the shortest lead times, and turbidity forecast error was relatively constant with lead time. The results demonstrate that currently-available observed and forecast inputs support useful forecasting of stream turbidity and may allow the forecasting of other water quality constituents.

## **5.2 Introduction**

While streamflow forecasts are important aspects of modern water management, extension of routine forecasting to include stream water quality is not yet commonplace. Of the stream water quality constituents of common interest, turbidity stands out in that it is widely monitored, can often be used to estimate other water quality constituents and serves as an overall indicator of the environmental health of streams. Turbidity is of particular interest to drinking water providers receiving source water from streams because drinking water treatment systems bear an estimated 0.25% increase in treatment cost per 1% increase in turbidity [*Dearmont et al.*, 1998], and episodes of elevated turbidity may require use of alternative raw water sources [e.g., *Portland Water Bureau*, 2011] or temporary shutdown of treatment systems [e.g., *Duncan and Grant*, 2003]. Forecasting stream turbidity for 1-3 days in the future could, for example, allow drinking water providers to anticipate future turbidity levels and make corresponding management decisions.

Prior work related to stream turbidity forecasting generally does not address day-to-day management decisions. For example, *Towler et al.* [2010] used seasonal probabilistic precipitation forecasts to create probabilistic forecasts of the number of seasonal turbidity threshold exceedances using streamflow ensembles. While useful for long-term decision making, these seasonal forecasts are not well suited for making short-

term operational decisions. *Yang et al.* [2014] established a turbidity early-warning forecasting system based on observed upstream rainfall accumulations. This forecasting system identified if the turbidity at a downstream station was expected to surpass a particular threshold (500 or 1000 *NTU*) during a given rainfall/streamflow event. While certainly useful for some management decisions, these approaches do not provide time-resolved forecasts of stream turbidity several days into the future.

The objective of the current study was to demonstrate and evaluate the utility of short-term (up to 3-day) operational stream turbidity forecasting based on routine streamflow, precipitation and temperature forecasts. Using two streams from the Mid-Atlantic region of the U.S., sources of error in the turbidity forecasts and the effect of streamflow forecast uncertainty on turbidity forecast uncertainty were assessed. The value of turbidity forecasting and the foundation this provides for the forecasting of other stream water quality constituents is also discussed.

### **5.2.1 Study Sites**

The Raritan River at Manville, NJ (USGS 01400500) is a 1269  $km^2$  watershed providing the drinking water source for a large portion of the ~1.2 million people living in the watershed [*New Jersey Water Supply Authority*, 2002]. Water quality in the Raritan River watershed is most notably degraded by phosphorus and fecal coliform, with other contaminants posing concerns in localized areas of the watershed [*New Jersey Water Supply Authority*, 2002]. The upper portion of the watershed is in the New England physiographic province while the lower region is in the Piedmont province [*Fenneman and Johnson*, 1946]. The land cover within the Raritan River watershed is 40% forest, 28% agriculture, 22% urban and 10% other classes [*Fry et al.*, 2011].

The Brandywine Creek at Wilmington, DE (USGS 01481500) has a watershed area of  $813 \text{ km}^2$  and is located predominantly in Pennsylvania, with the lowest reaches in Delaware. The Brandywine Creek watershed serves as the drinking water source for 200,000+ people [Crockett Consulting, 2010]. Water quality impairments in the Brandywine Creek watershed are commonly related to agricultural activities (siltation, nutrients and pathogens) [Crockett Consulting, 2010]. The watershed is mostly within the Piedmont province [Fenneman and Johnson, 1946], and the land cover is 34% agriculture, 33% forest, 23% urban and 10% other classes [Fry et al., 2011]. The Brandywine Creek watershed is located about 70 km from the Raritan River watershed.

## **5.3 Methods**

### **5.3.1 Turbidity Forecast Input Data**

Turbidity forecasts were made as if in an operational setting (i.e., based only on information known at the original time the streamflow/turbidity forecasts were issued).

The data assumed to be available for each turbidity forecast were:

- (1) A streamflow forecast,
- (2) a basin-average quantitative precipitation forecast (QPF),
- (3) a daily minimum and maximum air temperature forecast,
- (4) observed sub-hourly discharge prior to the forecast issue time,
- (5) observed precipitation prior to the forecast issue time, and
- (6) observed temperature prior to the forecast issue time.

The specific details for these input data are described in the sections below. In this study, it is assumed that observed discharge, observed precipitation and observed temperature are available up to one hour prior to the turbidity forecast issue time.

### **5.3.1.1 Streamflow Forecasts**

The stage and discharge forecasts (“streamflow forecasts”) used in this study were exactly as issued by the National Oceanic and Atmospheric Administration’s (NOAA) National Weather Service (NWS). Forecasts originally issued between 3/17/2014 and 5/18/2014 by the Middle Atlantic River Forecast Center (RFC) were archived on the day they were issued. To address data gaps, additional archived forecasts were obtained from the Service Records Retention System (SRRS) Bulletins located within the Hierarchical Data Storage System (HDSS; available through NOAA’s National Climatic Data Center (NCDC)). The station ID and bulletin ID used to retrieve the forecasts was KRHA and FGUS5, respectively. The NWS station ID is MNVN4 for the Raritan River and WMND1 for the Brandywine Creek location.

Streamflow forecasts for the two streams are issued “routinely”, meaning that at least one forecast is issued every day (typically in the morning local time), regardless of flow levels. During some periods, particularly when flow is high, additional forecasts may be issued throughout the day, and this was the case for several of the events included in this study. A “typical” streamflow forecast for the two streams used here includes stage and discharge at 6-hr time steps (i.e., 0:00, 6:00, 12:00 and 18:00 UTC) for 3 days into the future (~12 forecast points). A detailed description of the NWS streamflow forecast process can be found in *Welles et al.* [2007].

### **5.3.1.2 Precipitation Forecasts**

Basin-average quantitative precipitation forecasts (QPF) are routinely issued by the Middle Atlantic River Forecast Center multiple times a day and include up to 48 hours into the future with 6-hr accumulation time periods. Relevant QPFs were archived

on the day they were issued. A QPF forecast is issued for the basin draining to the Raritan River location used in this study (MNVN4). The WMND1 Brandywine Creek location does not have a corresponding QPF forecast, so an upstream station, CDFP1, was used as a substitute.

### ***5.3.1.3 Temperature Forecasts***

Daily minimum and maximum air temperature forecasts for 7-days into the future are routinely issued multiple times a day by the Philadelphia/Mt Holly NWS forecast office. Archived temperature forecasts were obtained from the NCDC HDSS system using the station ID KPHI and the bulletin ID FPUS6. Temperature forecasts for Somerville, NJ and Wilmington, DE were used in this study for Raritan River and Brandywine Creek, respectively; both cities are located near the stream's gaging station.

### ***5.3.1.4 Prior Observed Discharge, Precipitation and Temperature***

Sub-hourly observed discharge is available for these two streams from the USGS (gage numbers 01400500 and 01481500 for Raritan River and Brandywine Creek, respectively). To our knowledge, basin-average observed precipitation data are not currently published by the NWS. In this study we use a single station located centrally in each watershed and available through NOAA NCDC as input data for prior observed precipitation. The stations are USC00283029 and USC00369464 for Raritan River and Brandywine Creek, respectively. In true operational forecasting, the forecaster may choose to use several stations in the area to generate an estimate of basin-average observed precipitation; however this was not pursued in this study. Observed temperature was obtained from the same stations as the observed precipitation.

### 5.3.2 Turbidity Forecasting

The turbidity model and parameter estimation methods used in this study are described in Chapter 4, so only a brief description is presented here. Sub-hourly turbidity ( $TURB$ ) and discharge ( $Q$ ) data were obtained from the USGS for a 3+ year period (10/1/2009 to 4/22/2013) prior to the time period for which forecasts were made in this study and were used for model calibration. Hydrologic, or streamflow, events were identified in the calibration dataset and each event was individually fit to the following turbidity model using non-linear curve-fitting:

$$TURB(t) = aQ(t)^b e^{-rV(t)} + c \frac{dQ(t)}{dt} \quad (5.1)$$

$$\text{where, } V(t) = \int_0^T Q(t)dt \cong \sum_{i=0}^T Q_i \Delta t \quad (5.2)$$

and  $a$ ,  $b$ ,  $c$  and  $r$  are event model parameters [Mather and Johnson, 2014]. Parameter  $r$  was not permitted to take negative values. After fitting the model to all calibration events, fitted values of parameter  $a$  were transformed into  $a_0$  using a discharge-dimensioned, gage-specific scaling factor,  $Q_0$ , and the relationship:  $a_0 = aQ_0^b$  [Mather and Johnson, 2014]. This process decorrelated the power law parameters  $a$  and  $b$ , and simplified the units of  $a$  such that they were not dependent on  $b$ . Discharge was divided by catchment drainage area, resulting in specific discharge units of  $((m^3/s)/km^2)$ , and the turbidity data used here have units of ( $FNU$ ) [Anderson, 2005]. Therefore, the resulting units for event model parameters  $a_0$ ,  $b$ ,  $c$  and  $r$  are ( $FNU$ ),  $(-)$ ,  $(FNU(km^2s)(s/m^3))$  and  $(km^2/m^3)$ , respectively.

Prior observed and forecasted discharge, precipitation and temperature were used to estimate (i.e., forecast) the values of turbidity model parameters  $a_0$ ,  $b$ ,  $c$  and  $r$  using a “regression” approach (see Chapter 4 and Figure D 1 in Appendix D). The estimated model parameters  $a_0$ ,  $b$ ,  $c$  and  $r$  were used along with equations (5.1) and (5.2) to produce a turbidity forecast. The cumulative discharge throughout the forecast ( $V(t)$ ; equation (5.2)) was determined using either the forecast discharge or a combination of the prior observed and forecast discharge, depending on the timing of the turbidity forecast with respect to the timing of the streamflow event (see Figure D 2 for more details).

### 5.3.3 Evaluation of Forecasts

The mean error ( $ME$ ) and mean absolute error ( $MAE$ ) were used to evaluate the forecast accuracy of both the NWS streamflow forecasts and the turbidity forecasts made in this study.  $ME$  identifies if the forecasts tend to be too high or too low, while the  $MAE$  assesses how close the forecast is to the observed, regardless of the direction of the error. The “error” is defined here as the forecast minus the observed value at each forecast data point. The formulas used to calculate  $ME$  and  $MAE$  are:

$$ME = \frac{1}{N} \sum_{i=1}^N (f_i - o_i) \quad (5.3)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |f_i - o_i| \quad (5.4)$$

where  $N$  is the number of forecast/observation pairs,  $f_i$  is a forecast point and  $o_i$  is the corresponding observation.

To evaluate the “usefulness” of the forecasts and to aid in the interpretation of the calculated error metrics, we used a “persistence reference”, where the observation at the

forecast issue time is repeated for all lead times in the forecast period [Welles *et al.*, 2007]. The persistence reference (i.e., assuming that the future will be the same as the present) provides an objective baseline for minimum useful forecast accuracy. If a forecast has less error than the persistence reference, it is implied that the forecast is “useful”.

#### **5.3.4 Uncertainty Intervals**

NWS streamflow forecasts (and the turbidity forecasts made in this study) are deterministic (i.e., point) forecasts that do not give any indication of how accurate the forecast is likely to be. However, by comparing the forecast discharge to the subsequent actual observed discharge we can compute the empirical discharge uncertainty, which can then be propagated into the turbidity forecasting space. This provides a basis for visualizing the turbidity forecast uncertainty as well as the streamflow forecast uncertainty from which it was derived. A detailed description of the methods used to accomplish this can be found in Appendix D Section D.3; however, a brief description is provided below.

The deterministic streamflow forecasts used in this study and the corresponding observed discharge (i.e., forecast/observed data pairs) were binned into equal sized, logarithmic intervals based on forecast discharge. The relative forecast discharge error (observed divided by forecast, “*obs/fcst*”) was calculated for each data pair. Rather than assuming a specific distribution for the *obs/fcst* within each bin, the empirical distribution of each bin was used. For each originally issued deterministic streamflow forecast, 10,000 “new” streamflow forecasts were made by randomly sampling from the empirical *obs/fcst* distribution corresponding to the forecast discharge bin of the

originally issued forecast. Turbidity forecasts corresponding to each new streamflow forecast were then produced. Finally, the 10,000 new values associated with each forecast point were used to determine the 70% and 90% uncertainty intervals.

## 5.4 Results and Discussion

### 5.4.1 Streamflow Events and Forecasts

To demonstrate and evaluate the utility of turbidity forecasts, six streamflow events from 3/19/2014 to 5/21/2014 for both locations were used in this study. The details of these streamflow events are shown in Table 5.1. A total of 63 streamflow forecasts were issued in association with the six streamflow events for Raritan River and 55 for Brandywine Creek (Figure 5.1). Any routine streamflow forecasts issued between these events (i.e., during baseflow periods) were not considered here. Figure D 3 and Figure D 4 show the streamflow forecasts associated with each event in more detail.

**Table 5.1. Streamflow events used in this study.**

Stream	Event ID	Day of Event Peak Discharge	Observed Peak Specific Discharge $((m^3/s)/km^2)$	Observed Peak Turbidity (FNU)	Forecast Period of Interest
Raritan River	A	March 20, 2014	0.109	200	March 17-21
USGS: 01400500	B	March 30, 2014 <sup>a</sup>	0.254	500	March 26-April 1
NWS: MNVN4	C	April 8, 2014	0.044	17	April 5-9
	D	April 16, 2014	0.187	280 <sup>d</sup>	April 12-17
	E	May 1, 2014	0.565 <sup>b</sup>	280 <sup>d</sup>	April 27-May 3
	F	May 17, 2014	0.084	65	May 14-18
Brandywine Creek	A	March 20, 2014	0.055	72.1	March 17-20
USGS: 01481500	B	March 31, 2014 <sup>a</sup>	0.122	162	March 26-31
NWS: WMND1	C	April 8, 2014	0.036	9.7	April 5-8
	D	April 17, 2014	0.178	370	April 12-16
	E	May 1, 2014	0.794 <sup>c</sup>	359	April 27-May 3
	F	May 16, 2014	0.144	138	May 14-17

<sup>a</sup> Multiple peaks; peak values are for largest peak.

<sup>b</sup> 9th largest crest on record (1903-present) [NOAA, 2014b].

<sup>c</sup> Largest crest on record (1946-present) [NOAA, 2014a].

<sup>d</sup> Peak turbidity ignores single spike value on falling turbidity limb (see Figure 5.3).

These six streamflow events include some events for which the streamflow forecasts were largely consistent with the observed discharge, as well as events that were over- or under-forecast. In general, the forecast accuracies of the two streams are similar (e.g., event C was over-forecast for both Raritan River and Brandywine Creek). Because only six streamflow events were considered in this study, the forecast errors inferred from Figure 5.1 and discussed later should only be considered characteristic of these specific streamflow events and may not be representative of the forecast errors during other seasons or over a greater number of streamflow events.

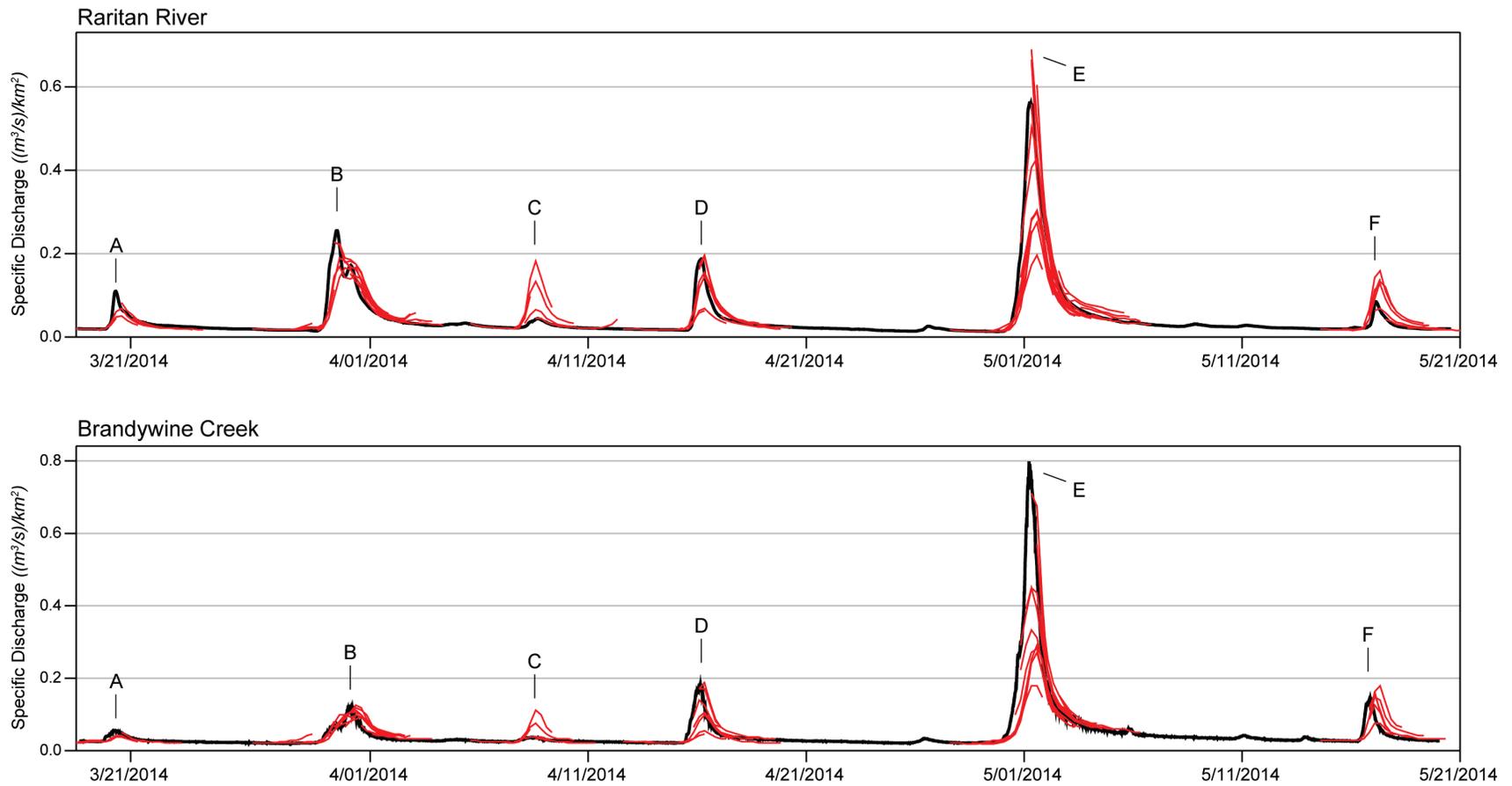


Figure 5.1. NWS streamflow forecasts (red) for events A-F and observed specific discharge (black) from 3/19/2014 – 5/21/2014 for Raritan River (top) and Brandywine Creek (bottom).

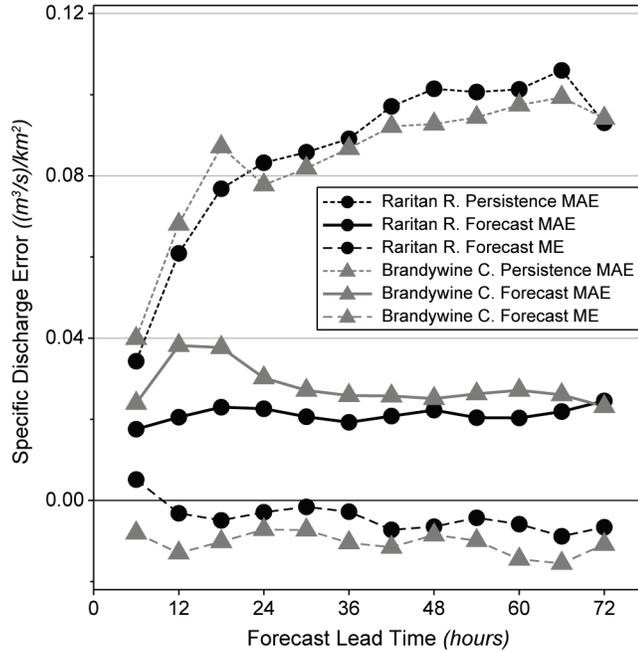
#### 5.4.2 Streamflow Forecast Errors

The streamflow forecast mean error ( $ME$ ) and mean absolute error ( $MAE$ ) for various forecast lead times are shown in Figure 5.2. Forecast lead time refers to the period of time between forecast issue time and the time of the individual points of the forecast. Each data point is calculated from ~63 forecast/observation pairs for Raritan River and ~55 for Brandywine Creek. For both streams  $ME$  is generally negative, indicating that discharge is under-forecast on average (except the shortest lead time for Raritan River). Figure 5.2 also shows that  $ME$  becomes increasingly negative for longer lead times. This is likely related to the time period of the precipitation forecast. Basin-average quantitative precipitation forecasts (QPF) for this region are issued for up to 48 hours into the future (and it is our assumption that the full duration of the QPF is used for streamflow forecasting purposes). As a result, discharge is more likely to be under-forecast at lead times greater than ~48 hours because potential precipitation after this point is not taken into consideration.

There is an oscillation with a period of ~24 hours visible in the  $ME$  for both streams. This likely occurs because most of the study events have their highest discharge periods in the morning hours (local time), which is also the time period that the majority of the streamflow forecasts were issued. As a result, the time periods that are approximate multiples of 24 hours in the future were preferentially higher-flow periods and subject to larger magnitude errors (see later discussion in Section 5.4.5).

Streamflow forecast  $MAE$  for both streams is nearly constant with lead time. The streamflow forecast  $MAE$  for Raritan River is approximately  $0.02 (m^3/s)/km^2$ , while Brandywine Creek  $MAE$  is slightly greater at  $\sim 0.025 (m^3/s)/km^2$  and is highest at 12-

and 18-hour lead times. This “bump” in Brandywine Creek *MAE* is assumed to be an artifact of the streamflow events used in this study, which is supported by the presence of a corresponding bump in the persistence reference for those same lead times.



**Figure 5.2.** Mean error (*ME*) and mean absolute error (*MAE*) in NWS streamflow forecasts (events A-F) for various lead times. Forecast lead times are binned at nearest 6-hour interval (i.e., a point at 12 hours is all forecasts within: 9 hrs < lead time ≤ 15 hrs). Persistence reference represents the observed specific discharge at the forecast issue time repeated for all lead times in the forecast period.

The persistence reference *MAE* is larger than the forecast *MAE* for both streams at all lead times, indicating that the streamflow forecasts are useful compared to this reference. Even at the shortest lead times, the forecast discharge *MAE* is considerably less than the *MAE* of the persistence reference. This is important because times in the near future are the most challenging for a forecast to outperform the persistence reference.

### **5.4.3 Turbidity Forecasts**

Turbidity forecasts corresponding to each issued streamflow forecast were made using the model parameter regression equations shown in Table 5.2 and the event turbidity model (equations (5.1) and (5.2)). Figure 5.3 shows that, in general, the turbidity forecasts are similar to the streamflow forecasts in terms of tending to over- or under-forecast a particular event, with the notable exception of Raritan River “event E”. It should be noted that event E was a historic streamflow event, particularly for Brandywine Creek, that far exceeded the size of the streamflow events used to calibrate the turbidity model. Although turbidity forecasts were made for this event, it was expected that the quality of the forecasts might be less than for the other five events. Figure D 5 and Figure D 6 show the turbidity forecasts for each event in more detail.

**Table 5.2. Regression equations to estimate event model parameters.**

Stream	Regression Equation	$R^2$	$R_{PRED}^2$ <sup>a</sup>
Raritan River	$\text{Log } a_0 = -0.783 - 1.401 \text{ Log } Q_{beg} + 0.00756 \text{ TEMP} + 0.514 \text{ Log } Q_p - 0.166 \text{ Log } (1 + \text{PRCP})$	84.5%	80.1%
$Q_0 = 0.0186 \text{ (m}^3/\text{s)/km}^2$	$b = 2.501 + 0.968 \text{ Log } Q_{beg}/Q_p - 0.201 \text{ Log } (1 + \text{AP5})$	45.7%	37.9%
	$c = 5.886 \times 10^6 + 8.558 \times 10^6 \text{ Log } MQ5 + 8.336 \times 10^6 \text{ Log } DSQ1 + 5.191 \times 10^6 \text{ Log } (1 + \text{AP1})$	30.1%	16.0%
	$r = -5.657 \times 10^{-6} + 7.700 \times 10^{-5} \text{ Log } DSQ2$	29.1%	21.2%
Brandywine Creek	$\text{Log } a_0 = 1.268 + 0.0106 \text{ TEMP} - 0.231 \text{ Log } Q_{beg}/Q_p + 0.110 \text{ Log } \text{DSP50}$	56.2%	44.7%
$Q_0 = 0.0636 \text{ (m}^3/\text{s)/km}^2$	$b = 1.693 + 0.781 \text{ Log } Q_{beg}/Q_p + 0.215 \text{ Log } \text{DSP50}$	63.3%	56.0%
	$c = -1.621 \times 10^7 - 9.116 \times 10^6 \text{ Log } \text{AQ30} - 3.426 \times 10^6 \text{ Log } \text{DSQ2}$	31.5%	22.6%
	$r = 1.575 \times 10^{-4} + 9.954 \times 10^{-5} \text{ Log } \text{AQ2} + 1.930 \times 10^{-5} \text{ Log } \text{DSQ4} - 9.488 \times 10^{-5} \text{ Log } \text{AQI30}$	52.2%	40.2%

Candidate Event Characteristics for Regression Equations

*Antecedent Moisture Conditions based on Precipitation*

- APn* Sum of the Daily Precipitation for the  $n$  days prior to event start ( $n = 1, 2, 5$  and  $10$  days), (mm)  
*DSP* Days since Measureable Precipitation, (days)  
*DSPn* Days since Daily Precipitation of at least  $n$  ( $n = 10, 20$  and  $50$  mm), (days)  
*APIn* Antecedent Precipitation Index ( $n = 5$  and  $10$  day half-life), (mm) (see Chapter 4)

*Antecedent Moisture Conditions based on Discharge*

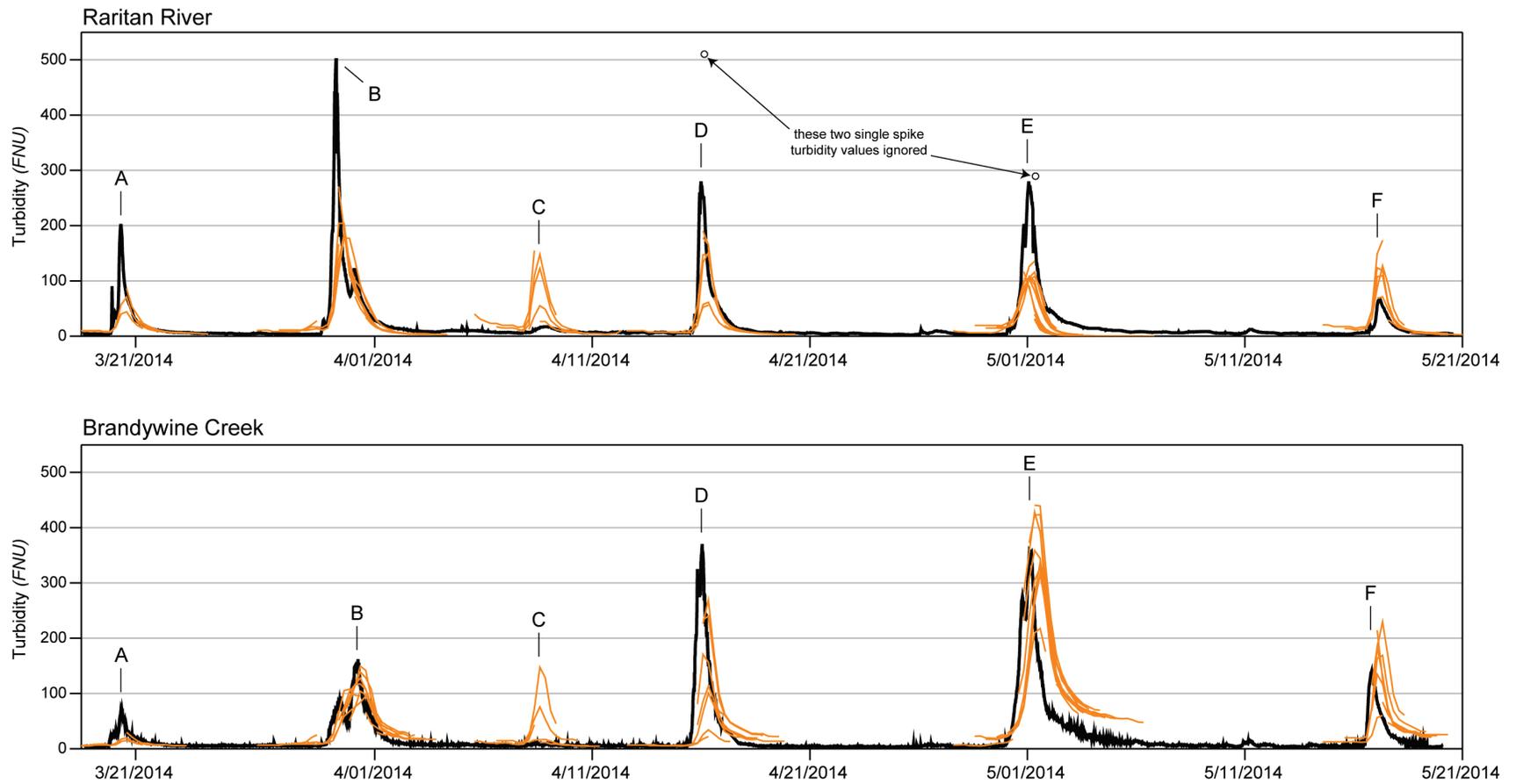
- $Q_{beg}$  Specific Discharge at the Beginning of the Event, ( $\text{m}^3/\text{s)/km}^2$ )  
*AQn* Average Specific Daily Discharge for the  $n$  days prior to event start ( $n = 1, 2, 5, 10$  and  $30$  days), ( $\text{m}^3/\text{s)/km}^2$ )  
*MQn* Maximum Specific Daily Discharge for the  $n$  days prior to event start ( $n = 5, 10$  and  $30$  days), ( $\text{m}^3/\text{s)/km}^2$ )  
*DSQn* Days since Daily Discharge of at least  $n$  times the Mean Annual Daily Discharge ( $n = 1, 2, 3, 4$  and  $5$ ), (days)  
*AQIn* Antecedent Specific Discharge Index ( $n = 5, 10$  and  $30$  day half-life), ( $\text{m}^3/\text{s)/km}^2$ ) (see Chapter 4)

*Other Characteristics*

- $Q_p$  Peak Specific Discharge, ( $\text{m}^3/\text{s)/km}^2$ )  
 $Q_{beg}/Q_p$  Ratio of Discharge at the Beginning of the Event to Peak Discharge (approximation of event baseflow index), (–)  
*PRCP* Event Precipitation, (mm)  
*TEMP* Average Air Temperature for the day of Peak Discharge, (°C)

Note: When the regression equations are used in an operational setting to forecast model parameter values, the event characteristics  $Q_p$ ,  $Q_{beg}$ , *PRCP* and *TEMP* may be determined from either the respective forecasts or observed data depending if the turbidity forecast corresponds to the time prior to, during or after the peak of the streamflow event.

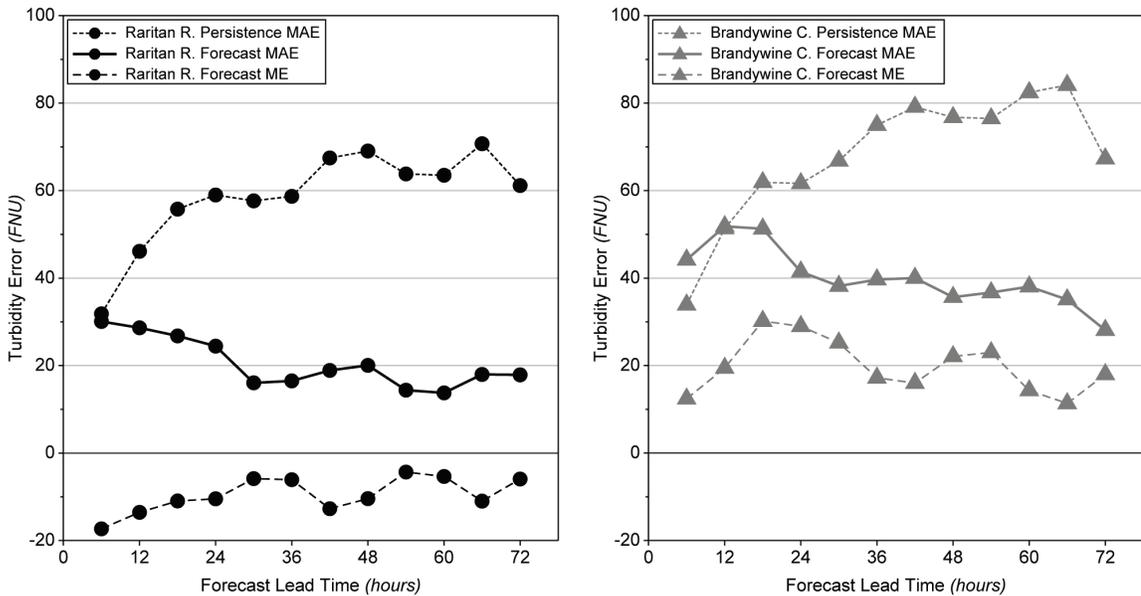
<sup>a</sup>  $R_{PRED}^2$  is calculated by systematically removing each observation and recalculating the regression model (with the same event characteristics present) to determine how well the model predicts each removed observation (Minitab 16 Statistical Software).



**Figure 5.3. Turbidity forecasts (orange) for events A-F and observed turbidity (black) from 3/19/2014 – 5/21/2014 for Raritan River (top) and Brandywine Creek (bottom).**

#### 5.4.4 Turbidity Forecast Errors

The turbidity forecast *ME* and *MAE* for various lead times are shown in Figure 5.4. For both streams, *ME* does not display a clear trend with lead time; however, as in Figure 5.2, a ~24-hour oscillation is present due to the correspondence of forecast issue times and high turbidity time periods. The *ME* for Raritan River is negative for all lead times, whereas Brandywine Creek *ME* is positive. This confirms what is qualitatively visible in Figure 5.3 (i.e., Raritan River turbidity forecasts are more likely to differ from the observed towards under-forecasting with these 6 streamflow events, while Brandywine Creek turbidity forecasts tend towards over-forecasting.)



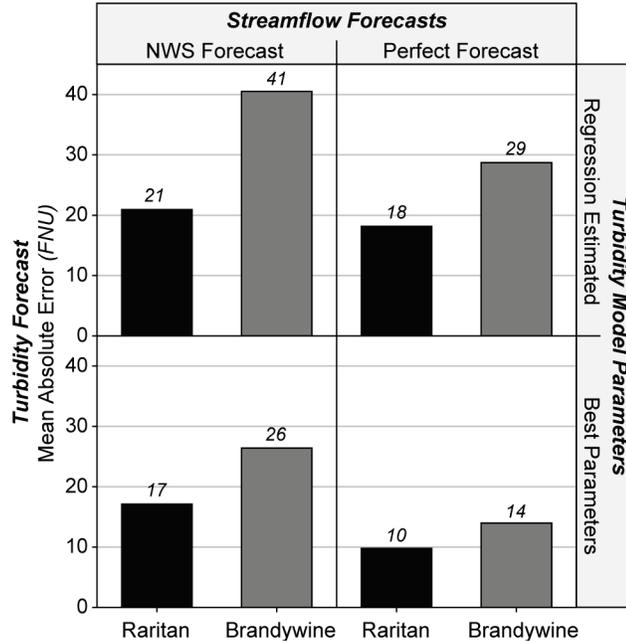
**Figure 5.4. Mean error (*ME*) and mean absolute error (*MAE*) in Raritan River (left) and Brandywine Creek (right) turbidity forecasts for various lead times. Forecast lead times are binned at nearest 6-hour interval (i.e., a point at 12 hours is all forecasts with: 9 hrs < lead time ≤ 15 hrs). Persistence reference represents the observed turbidity at the forecast issue time repeated for all lead times in the forecast period.**

The turbidity forecast *MAE* shows a slight decrease with increasing lead time for both streams (Figure 5.4). This is not what was expected based on the assumption that time periods further into the future are less well known. However, Figure 5.2 shows that streamflow forecast *MAE* also did not increase with lead time. Finally, Brandywine Creek has a *MAE* about twice that of Raritan River for most lead times.

Figure 5.4 shows that the persistence reference *MAE* is clearly greater than the turbidity forecast *MAE* for Raritan River at all lead times beyond 6 hours (the shortest lead time). This indicates that the turbidity forecasts made for this stream were useful, particularly for lead times greater than ~24 hours and result in a *MAE* of about one-third of the persistence reference. Brandywine Creek turbidity forecast *MAE* did not compare as favorably to the persistence reference, and the turbidity forecasts made for this stream had a higher *MAE* than the persistence for 6- and 12-hour lead times. This indicates that for these shortest lead times, less absolute error would be obtained from assuming the future is the same as the present. However, similar to Raritan River, the turbidity forecasts for Brandywine Creek at lead times greater than ~24 hours were useful and have about one-half the *MAE* of the persistence reference.

The general sources of error in the turbidity forecasts were examined by altering the turbidity forecast inputs and model parameter values. To assess the impact of NWS streamflow forecast error, we compared the turbidity forecast *MAE* that results from using a “perfect streamflow forecast” in which each original forecast discharge was replaced by the observed discharge. The error resulting from the estimation (i.e., forecasting) of turbidity model parameters using regression equations (Table 5.2) was evaluated by comparing turbidity forecasts made using “best parameters”, or the optimal

parameters determined by fitting the turbidity model to the observed turbidity for each event. The resulting turbidity forecast *MAE* for each of these scenarios is shown in Figure 5.5.



**Figure 5.5.** Turbidity forecast mean absolute error (*MAE*) for several input data scenarios. “Perfect Forecast” represents a streamflow forecast that perfectly matches the observations at each original forecast time point. “Best Parameters” represents optimal values of turbidity model parameters  $a_0$ ,  $b$ ,  $c$  and  $r$  for each event.

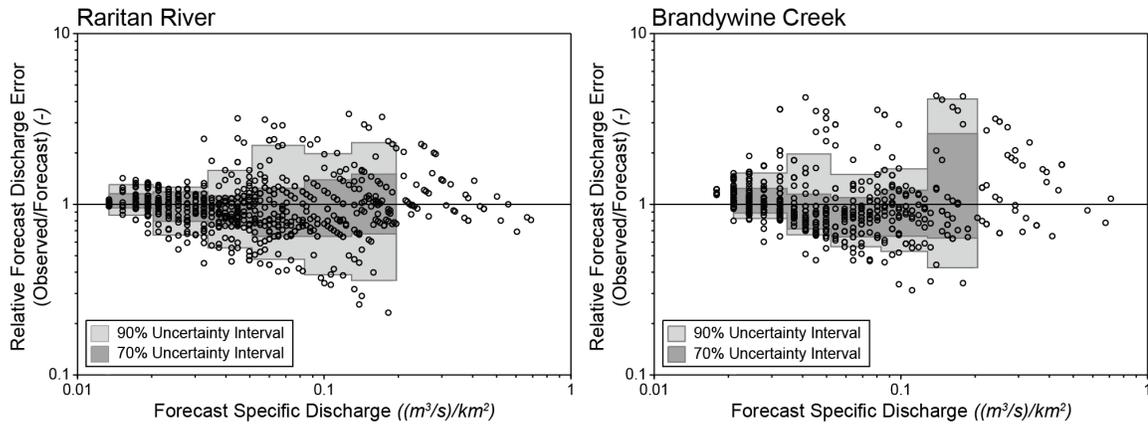
Raritan River turbidity forecasts have a 3-4 *FNU* mean absolute error improvement if either the streamflow forecast is perfect or the best values of the model parameters are used. However, when both of these things occur (perfect streamflow forecast and best parameters), the turbidity forecast is substantially better with a *MAE* of only 10 *FNU*. Brandywine Creek similarly shows the greatest improvement in turbidity forecast *MAE* with the use of both a perfect streamflow forecast and best parameter values, but the turbidity forecasts for this stream are also considerably improved by either

of those two input modifications individually. This suggests that improvements in turbidity forecasts for Brandywine Creek can be achieved separate from any improvements in the NWS streamflow forecast through the refining of the turbidity model parameter estimation methods. However, Raritan River turbidity forecast improvements in absence of improvements in NWS streamflow forecasts may need to come primarily through post-processing or data assimilation strategies.

Also considered but not shown in Figure 5.5 was the influence of streamflow forecast time-resolution on the turbidity forecast *MAE*. A perfect streamflow forecast at the original 6-hr time-resolution was compared to a perfect streamflow forecast at a 1-hr resolution for the same duration into the future. We found no significant difference in turbidity forecast *MAE* for the shorter time-resolution for either model parameter scenario for both streams.

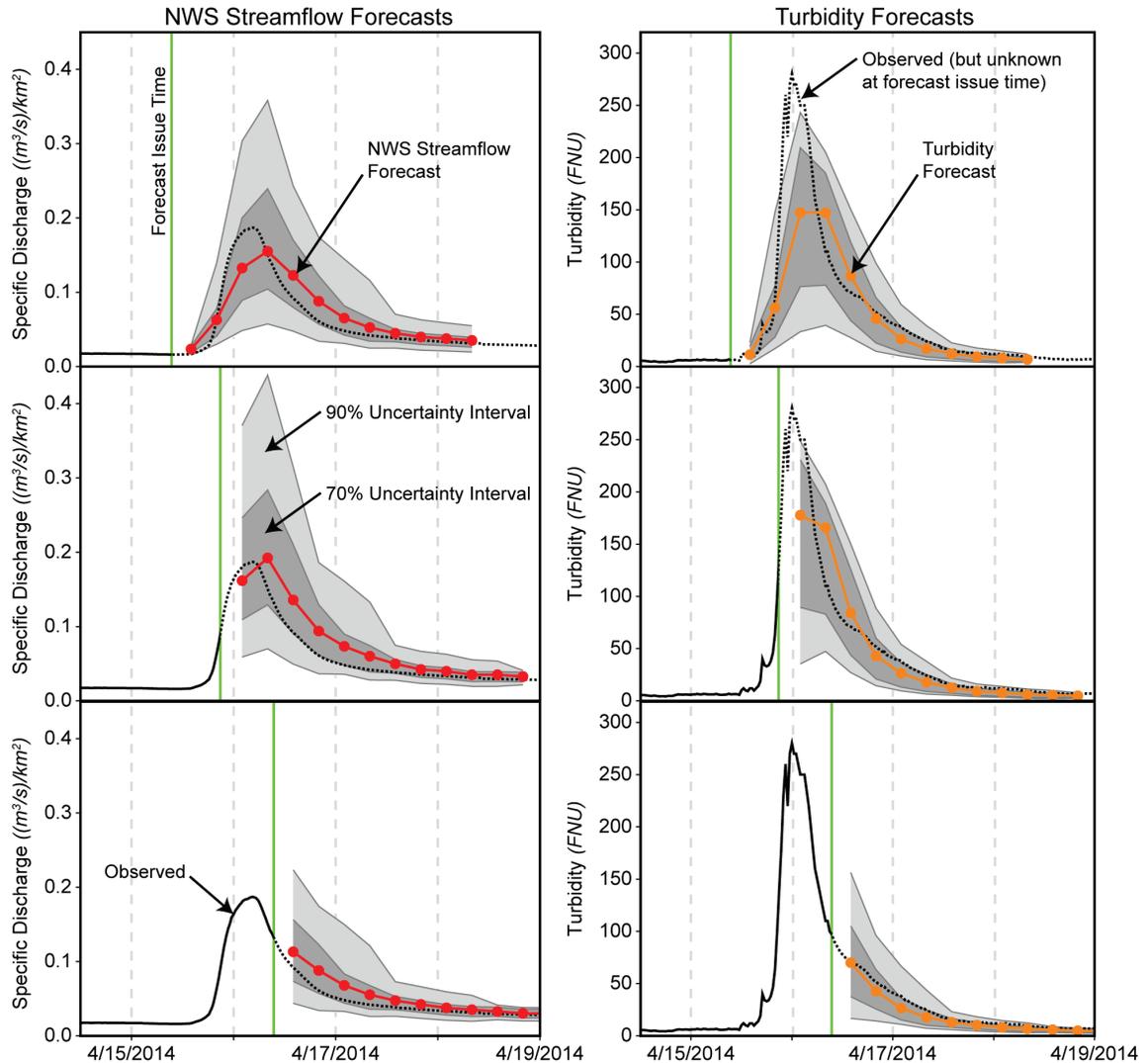
#### **5.4.5 Effect of Streamflow Forecast Uncertainty on Turbidity Forecasts**

To understand the effect of streamflow forecast uncertainty on turbidity forecast uncertainty, the empirical distributions of the relative forecast discharge errors for the study events were randomly sampled and propagated into the turbidity forecast space. Figure 5.6 shows the relative forecast discharge error for forecast points associated with the study events. The x-axis (forecast specific discharge) was divided into regular, logarithmic increments of  $\sim 1.6X$  (2 dB) and the empirical uncertainty intervals for each forecast specific discharge bin are shown. Figure 5.6 shows the resulting uncertainty intervals generally become wider (larger) as the forecast discharge increases.



**Figure 5.6. Relative forecast discharge error for Raritan River (left) and Brandywine Creek (right) as a function of forecast specific discharge for individual forecast data points over events A-F. Empirical uncertainty intervals (70% dark gray, 90% light grey) shaded for  $\sim 1.6X$  (2 dB) x-axis bins. Uncertainty intervals for forecast specific discharge greater than  $\sim 0.2 (m^3/s)/km^2$  were not calculated because data points beyond this threshold were sparse and nearly exclusively from a single event (event E).**

Figure 5.7 shows an example of what the empirical uncertainty intervals look like when applied to streamflow forecasts and propagated into the turbidity forecasts. The uncertainty intervals for specific discharge are larger above the deterministic NWS streamflow forecast than below; however, turbidity uncertainty intervals are narrower in the upper portion than the lower. This is related to the impact higher event peak discharge ( $Q_p$ ) has on the turbidity model parameters determined using the regression equations in Table 5.2. Increasing the forecast discharge for Raritan River, for example, causes a corresponding increase in model parameter  $a_0$  but a decrease in model parameter  $b$ , with the net effect of tempering the proportional turbidity response as streamflow event size increases.



**Figure 5.7. Three example forecasts from Raritan River event D. Left panels show NWS streamflow forecast and empirical uncertainty intervals. Right panels show turbidity forecasts and uncertainty intervals produced from streamflow forecast uncertainty. Observed (solid black), observed but not yet known at forecast issue time (dotted black), streamflow forecast (red), turbidity forecast (orange), 90% uncertainty interval (light gray), 70% uncertainty interval (dark grey) and forecast issue time (green).**

## 5.5 Summary and Conclusions

There are several notable differences between the turbidity forecasts made in this study and the NWS streamflow forecasts on which they are based. First, the turbidity forecasts are a product of input data and corresponding model results alone, whereas the operational NWS streamflow forecasts are hybrid products of both model results and forecaster intervention based on professional experience [*Shamir et al.*, 2006]. Additionally, the turbidity forecasts were not subject to any post-processing (e.g., to correct systematic bias between the latest observation and first forecast value to make the forecast consistent with the latest observation). Thirdly, the turbidity forecasts were not preconditioned using observed turbidity data leading up to the forecast issue time (which is generally the case for streamflow forecasts). Even without those improvements, the turbidity forecasts had lower error than the persistence reference for nearly all lead times, suggesting that they are useful in their current implementation. Furthermore, the turbidity forecasts in this study were based solely on currently-available observed data and forecast inputs, demonstrating that useful water quality forecasts can be made with data readily available to the forecaster.

The two streams used as examples in the current work are representative of a much larger group of gage locations across the U.S. for which all necessary inputs (both observed and forecast) for stream turbidity forecasting currently exist. This suggests that the forecasting performed here could readily be extended to many other streams across the country. Additionally, future applications could utilize turbidity to forecast other water quality constituents of interest. Turbidity is often a useful surrogate to estimate concentrations of suspended sediment [e.g., *Gray and Glysson*, 2003], pathogens [e.g.,

*Christensen et al., 2000*] and agricultural chemicals [e.g., *Hickman, 2004*]. In the Raritan River watershed, for example, fecal coliform levels are elevated and of interest to water supply providers [*New Jersey Water Supply Authority, 2002*]. Turbidity has been shown to correlate with fecal coliform and *E. coli* levels in the watershed [*Money et al., 2009*], suggesting the forecasting of these other water quality constituents may also be possible. In the face of growing demand on water resources, water quality forecasting may become an increasingly important component of efforts to provide safe drinking water.

# CHAPTER 6

---

## Summary and Conclusions

While streamflow forecasting is quite common, stream water quality forecasting remains relatively uncommon, despite the clear benefits that it could provide. The research presented here demonstrates the utility of short-term (i.e., 3-day) water quality forecasting, using turbidity as an example. In pursuing that end goal, this research produced a number of tangible products, including: (1) an improved event turbidity model, (2) a method for power-law model parameter decorrelation, and (3) extensive analyses of the factors influencing turbidity event model parameters. Hopefully, all of these will initiate future examination of the hydrologic and catchment controls on turbidity behavior during hydrologic events, and refinement of the turbidity forecasting methods demonstrated here. The following sections summarize several outcomes of the research presented in this dissertation.

*The improved power-law based model proposed in this research, containing both a discharge rate of change term and a supply term, was able to fit a wider range of both simple and complex turbidity events than prior models. As such, it is an important tool for characterizing water quality during hydrologic events.* Historically, simple power law models (i.e., rating curves) were commonly used to relate water quality constituent “concentrations” to discharge. However, as a monotonic relationship, the power law alone is unable to represent the hysteretic behavior frequently observed for turbidity and other water quality constituents. Previous researchers have proposed models that included a “discharge rate of change” term or a “supply” term, and the research presented here showed that, taken together, a single model containing versions of these terms captures a wide range of simple and complex event turbidity behaviors resulting from diverse energy-driven transport and supply conditions. The resulting model contains four fitting parameters—two are associated with a power-law term ( $a$  and  $b$ ), plus one each for the discharge rate of change ( $c$ ) and the supply term ( $r$ ). For a small subset of the events modeled in this research, specifically events having different numbers of discharge and turbidity peaks (e.g., two turbidity peaks associated with a single discharge peak), even the improved model was not always successful in fully capturing the event turbidity shape. In many cases, multiple turbidity peaks were associated with multiple precipitation periods during a single hydrologic event. Future research should examine the role that event precipitation plays in the occurrence of multi-peaked event turbidity.

*The relationships between event model power-law coefficients and catchment/event characteristics were clarified through the use of a decorrelation and unit simplification procedure. This method involved the transformation of “ $a$ ” into  $a_0$*

using a discharge-dimensioned, gage-specific scaling factor,  $Q_0$ , and resolved two long-standing issues with the power law. First, it has previously been observed by many researchers in diverse fields that the power-law coefficient,  $a$ , and the power-law exponent,  $b$ , are highly correlated, which has implications for determining the underlying factors influencing each parameter. The procedure used here strategically selects a value of  $Q_0$  such that the transformed power-law coefficient,  $a_0$ , is not correlated with  $b$ . Second, this method results in the simplification of the units associated with the power-law coefficient such that they are no longer dependent on the value of the exponent,  $b$ . Furthermore, the gage-specific values of  $Q_0$  were found to be correlated to the characteristics of the catchment. Initial exploration suggests that  $Q_0$  is predictable from catchment characteristics, possibly facilitating applications of this methodology to “ungaged” locations in future work.

*Turbidity event model parameters, particularly those associated with the power-law term ( $a_0$  and  $b$ ), were shown here to be correlated with catchment land cover, baseflow index and antecedent moisture conditions.* The values of event model parameters  $a_0$ ,  $b$ ,  $c$  and  $r$  vary on an event-by-event basis, and two separate analyses of these parameter values were considered in this dissertation. To understand how the characteristics of a catchment relate to general turbidity behavior, the gage median parameter values for 110 gages with a combined total of ~6000 individual events were explored using correlation analysis and spatial mapping. The results indicated that parameter  $a_0$  was well correlated with land cover characteristics, particularly the percentage forest. Parameter  $b$  was found to be highly correlated with many catchment characteristics, but was best correlated with baseflow index. The remaining parameters,  $c$

and  $r$ , showed only weak correlations with the catchment-wide characteristics considered in this research. Some studies, such as *Hunsaker and Levine* [1995] and *Buck et al.* [2004], have found water quality constituents to be correlated with localized catchment characteristics. Future work will, hopefully, explore if these model parameters can be better understood in terms of the characteristics of the catchment local to the gage and/or near the stream (i.e., along the stream corridor).

*Turbidity model parameter event-to-event variability was found to be largely related to the antecedent moisture conditions preceding the hydrologic event.* In order to both establish a basis from which to predict model parameter values for forecasting purposes, and explore the potential physical controls on event parameter values, this research used several methods to predict parameters from the event characteristics. One method involved first using cluster analysis to group training events based on  $a_0$ ,  $b$ ,  $c$  and  $r$  similarity (i.e., similarity of the parameter “set”), then utilizing a classification tree based on event characteristics to predict parameter sets for new events. In a more traditional approach, multiple linear regression was used to predict values of  $a_0$ ,  $b$ ,  $c$  and  $r$  individually from the event characteristics by means of equations developed using training events. The results for both prediction methods show that indicators of antecedent moisture conditions, particularly those derived from discharge, were the most useful predictors of model parameter values. While the cluster/classification approach produced a slightly better overall prediction success rate, the regression approach was chosen to produce turbidity forecasts in this research. This is because the regression approach produced a more stable updating of the turbidity forecast as successive forecasts were issued throughout a single event.

*Useful turbidity forecasts (i.e., better than a persistence reference) based solely on currently-available observed data and forecast inputs can be obtained using the event model, parameter decorrelation and model parameter prediction methods developed in this dissertation.* Within the U.S., an extensive gage/station network of observed and forecast discharge, precipitation and temperature is available through the efforts of the USGS and the NWS. The research presented here utilized these inputs as the basis for forecasting of stream turbidity for 3 days into the future with 6-hr time-resolution (to match typical streamflow forecasts). The observed and forecast inputs were used in two ways—both directly as an independent variable in the turbidity model (i.e., discharge) and to estimate the most appropriate values for the turbidity model parameters using regression equations. Using archived input forecasts, turbidity forecasts were made for events from two streams selected from the larger data set examined here. The forecasts were found to be “useful” when compared to a persistence reference. The turbidity forecast uncertainty resulting from the forecast discharge uncertainty was also considered, and it was found that the uncertainty interval is much wider towards values less than the deterministic forecast.

Future improvements in water quality forecasting are likely to come by implementing some of the strategies currently used in streamflow forecasting, such as data assimilation (manual or automatic). Additionally, there remain many opportunities for refinement of the methods used to forecast event model parameters. Future work involving longer data records than what is currently available may permit greater representation of extreme events and could support additional analyses of the stationarity (or perhaps non-stationarity) of model parameter values for individual streams over time.

Turbidity is both an important primary water quality characteristic and a valuable surrogate; therefore, this research provides an important advancement for operational water management, in and of itself, and in the foundation it lays for the forecasting of other water quality constituents of interest. While the projected impact of climate change on water quantity and quality varies by location, the overall implications are largely negative, and current water management practices may be insufficient to deal with these changes [*Bates et al.*, 2008]. Water quality forecasts may become an important tool facilitating day-to-day operational decisions that optimize water management in order to cope with the water challenges we will likely face in the future.

## REFERENCES

---

- Ali, G., and A. Roy (2010), A case study on the use of appropriate surrogates for antecedent moisture conditions (AMCs), *Hydrol. Earth Syst. Sci.*, 14(10), 1843–1861, doi:10.5194/hess-14-1843-2010.
- Anderson, C. (2005), *Turbidity (ver. 2.1): U.S. Geological Survey Techniques of Water-Resources Investigations, Book 9, Chapter A6, Section 6.7.*
- Arsenault, R., and F. P. Brissette (2014), Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches, *Water Resour. Res.*, 50(7), 6135–6153, doi:10.1002/2013WR014898.
- Asselman, N. E. M. (1999), Suspended sediment dynamics in a large drainage basin: the River Rhine, *Hydrol. Process.*, 13, 1437–1450, doi:10.1002/(SICI)1099-1085(199907)13:10<1437::AID-HYP821>3.0.CO;2-J.
- Asselman, N. E. M. (2000), Fitting and interpretation of sediment rating curves, *J. Hydrol.*, 234(3–4), 228–248, doi:10.1016/S0022-1694(00)00253-5.
- Bača, P. (2008), Hysteresis effect in suspended sediment concentration in the Rybárik basin, Slovakia, *Hydrol. Sci. J.*, 53(1), 224–235, doi:10.1623/hysj.53.1.224.
- Bates, B. C., Z. W. Kundzewicz, S. Wu, and J. P. Palutikof (2008), *Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva.*
- Bergner, F., and G. Zouhar (2000), A new approach to the correlation between the coefficient and the exponent in the power law equation of fatigue crack growth, *Int. J. Fatigue*, 22, 229–239, doi:10.1016/S0142-1123(99)00123-1.
- Bowes, M. J., W. A. House, R. A. Hodgkinson, and D. V. Leach (2005), Phosphorus-discharge hysteresis during storm events along a river catchment: the River Swale, UK., *Water Res.*, 39(5), 751–762, doi:10.1016/j.watres.2004.11.027.
- Brasington, J., and K. Richards (2000), Turbidity and suspended sediment dynamics in small catchments in the Nepal Middle Hills, *Hydrol. Process.*, 14, 2559–2574, doi:10.1002/1099-1085(20001015)14:14<2559::AID-HYP114>3.0.CO;2-E.
- Buck, O., D. K. Niyogi, and C. R. Townsend (2004), Scale-dependence of land use effects on water quality of streams in agricultural catchments., *Environ. Pollut.*, 130(2), 287–299, doi:10.1016/j.envpol.2003.10.018.

- Campbell, F. B., and H. A. Bauder (1940), A rating-curve method for determining silt-discharge of streams, *Trans. AGU*, 21, 603–607.
- Chang, H., and T. N. Carlson (2005), Water quality during winter storm events in Spring Creek, Pennsylvania USA, *Hydrobiologia*, 544(1), 321–332, doi:10.1007/s10750-005-1894-6.
- Christensen, V., X. Jian, and A. C. Ziegler (2000), *Regression Analysis and Real-Time Water-Quality Monitoring to Estimate Constituent Concentrations, Loads, and Yields in the Little Arkansas River, South-Central Kansas, 1995-99: U.S. Geological Survey Water-Resources Investigations Report 00-4126*.
- Cortie, M. (1991), The irrepressible relationship between the Paris law parameters, *Eng. Fract. Mech.*, 40(3), 681–682, doi:10.1002/1099-1085(20001015)14:14<2559::AID-HYP114>3.0.CO;2-E.
- Crockett Consulting (2010), City of Wilmington Source Water Protection Plan, Available from: <http://www.ci.wilmington.de.us/government/sourcewater> (Accessed 1 February 2015)
- Davies-Colley, R. J., and D. G. Smith (2001), Turbidity, suspended sediment, and water clarity: A review, *J. Am. Water Resour. Assoc.*, 37(5), 1085–1101, doi:10.1111/j.1752-1688.2001.tb03624.x.
- De'ath, G., and K. Fabricius (2000), Classification and regression trees: a powerful yet simple technique for ecological data analysis, *Ecology*, 81(11), 3178–3192, doi:10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2.
- Dearmont, D., B. McCarl, and D. Tolman (1998), Costs of water treatment due to diminished water quality: a case study in Texas, *Water Resour. Res.*, 34(4), 849–854, doi:10.1029/98WR00213.
- Doomen, A. M. C., E. Wijma, J. J. G. Zwolsman, and H. Middelkoop (2008), Predicting suspended sediment concentrations in the Meuse River using a supply-based rating curve, *Hydrol. Process.*, 22(12), 1846–1856, doi:10.1002/hyp.6767.
- Duncan, S., and G. Grant (2003), *Clarifying muddy water: Probing the linkages to municipal water quality.*, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.
- Duvert, C., N. Gratiot, O. Evrard, O. Navratil, J. Némery, C. Prat, and M. Esteves (2010), Drivers of erosion and suspended sediment transport in three headwater catchments of the Mexican Central Highlands, *Geomorphology*, 123(3-4), 243–256, doi:10.1016/j.geomorph.2010.07.016.

- Duvert, C. et al. (2012), Towards prediction of suspended sediment yield from peak discharge in small erodible mountainous catchments (0.45–22 km<sup>2</sup>) of France, Mexico and Spain, *J. Hydrol.*, 454–455, 42–55, doi:10.1016/j.jhydrol.2012.05.048.
- Eder, A., P. Strauss, T. Krueger, and J. N. Quinton (2010), Comparative calculation of suspended sediment loads with respect to hysteresis effects (in the Petzenkirchen catchment, Austria), *J. Hydrol.*, 389(1-2), 168–176, doi:10.1016/j.jhydrol.2010.05.043.
- Eder, A., M. Exner-Kittridge, P. Strauss, and G. Blöschl (2014), Re-suspension of bed sediment in a small stream – results from two flushing experiments, *Hydrol. Earth Syst. Sci.*, 18(3), 1043–1052, doi:10.5194/hess-18-1043-2014.
- Falcone, J. A., D. M. Carlisle, D. M. Wolock, and M. R. Meador (2010), GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States, *Ecology*, 91(2), 621, doi:10.1890/09-0889.1.
- Fenneman, N. M., and D. W. Johnson (1946), *Physiographic divisions of the conterminous U.S.*, U.S. Geological Survey, Reston, VA.
- Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham (2011), Completion of the 2006 national land cover database for the conterminous United States, *Photogramm. Eng. Remote Sens.*, 77(9), 858–864.
- Giménez, R., J. Casali, I. Grande, J. Díez, M. A. Campo, J. Álvarez-Mozos, and M. Goñi (2012), Factors controlling sediment export in a small agricultural watershed in Navarre (Spain), *Agric. Water Manag.*, 110, 1–8, doi:10.1016/j.agwat.2012.03.007.
- Gray, J., and G. Glysson (2003), *Proceedings of the Federal Interagency Workshop on Turbidity and other Sediment Surrogates, April 30-May 2, 2002, Reno, Nevada: U.S. Geological Survey Circular 1250.*
- Gupta, H. V., C. Perrin, G. Blöschl, A. Montanari, R. Kumar, M. Clark, and V. Andréassian (2014), Large-sample hydrology: a need to balance depth with breadth, *Hydrol. Earth Syst. Sci.*, 18(2), 463–477, doi:10.5194/hess-18-463-2014.
- Guy, H. (1964), *An analysis of some storm-period variables affecting stream sediment transport: U.S. Geological Survey Professional Paper 462-E.*
- Helsel, D., and R. Hirsch (2002), *Statistical methods in water resources, Techniques of Water-Resource Investigations of the U.S. Geological Survey: Book 4, Hydrologic Analysis and Interpretation.*
- Heng, S., and T. Suetsugi (2014), Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent

- instantaneous sediment yield prediction, *J. Hydrol.*, 512, 240–253, doi:10.1016/j.jhydrol.2014.03.003.
- Henley, W. F., M. A. Patterson, R. J. Neves, and D. A. Lemly (2000), Effects of sedimentation and turbidity on lotic food webs: a concise review for natural resource managers, *Rev. Fish. Sci.*, 8(2), 125–139, doi:10.1080/10641260091129198.
- Hickman, E. (2004), *Pesticide Compounds in Streamwater in the Delaware River Basin, December 1998 – August 2001: U.S. Geological Survey Scientific Investigations Report 2004-5105*.
- Holder, R. (1985), *Multiple regression in hydrology*, Institute of Hydrology, Wallingford, Oxfordshire.
- House, W., and M. Warwick (1998), Hysteresis of the solute concentration/discharge relationship in rivers during storms, *Water Res.*, 32(8), 2279–2290, doi:10.1080/10641260091129198.
- Hunsaker, C. T., and D. a. Levine (1995), Hierarchical Approaches to the Study of Water Quality in Rivers, *Bioscience*, 45(3), 193–203, doi:10.2307/1312558.
- Hussain, M., S. Kar, and R. Puniyani (1999a), Relationship between power law coefficients and major blood constituents affecting the whole blood viscosity, *J. Biosci.*, 24(3), 329–337, doi:10.1007/BF02941247.
- Hussain, M. A., S. Kar, R. R. Puniyani, and R. V. Gupte (1999b), Investigation on the blood viscosity autoregulation factor in some diseases using power law model, in *Proceedings of the 26th National Conference on Fluid Mechanics and Fluid Power: December 15 - 17, 1999*, pp. 289–293.
- IPCC (2014), Summary for Policymakers, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by C. B. Field et al., pp. 1–32, Cambridge University Press, Cambridge, United Kingdom, and New York, NY, USA.
- James, A. L., and N. T. Roulet (2009), Antecedent moisture conditions and catchment morphology as controls on spatial patterns of runoff generation in small forest catchments, *J. Hydrol.*, 377(3–4), 351–366, doi:10.1016/j.jhydrol.2009.08.039.
- Jansson, M. B. (2002), Determining sediment source areas in a tropical river basin, Costa Rica, *Catena*, 47(1), 63–84, doi:10.1016/S0341-8162(01)00173-4.
- Jones, A. S., D. K. Stevens, J. S. Horsburgh, and N. O. Mesner (2011), Surrogate Measures for Providing High Frequency Estimates of Total Suspended Solids and

- Total Phosphorus Concentrations, *J. Am. Water Resour. Assoc.*, 47(2), 239–253, doi:10.1111/j.1752-1688.2010.00505.x.
- Juracic, M., L. M. Vitturi, S. Rabitti, and G. Rampazzo (1986), Suspended matter properties and its role in pollutant transfer from the river to the sea. Case study: Adige river-Adriatic Sea, *Sci. Total Environ.*, 55, 243–249, doi:10.1016/0048-9697(86)90183-X.
- Kirchner, J. W., X. Feng, C. Neal, and A. J. Robson (2004), The fine structure of water-quality dynamics: The (high-frequency) wave of the future, *Hydrol. Process.*, 18(7), 1353–1359, doi:10.1002/hyp.5537.
- Kohler, M. A., and R. K. Linsley (1951), *Predicting the runoff from storm rainfall: U.S. Weather Bureau Research Paper No. 34.*
- Krueger, T., J. N. Quinton, J. Freer, C. J. A. Macleod, G. S. Bilotta, R. E. Brazier, P. Butler, and P. M. Haygarth (2009), Uncertainties in data and models to describe event dynamics of agricultural sediment and phosphorus transfer., *J. Environ. Qual.*, 38(3), 1137–1148, doi:10.2134/jeq2008.0179.
- Krueger, T., J. N. Quinton, J. Freer, C. J. A. Macleod, G. S. Bilotta, R. E. Brazier, J. M. B. Hawkins, and P. M. Haygarth (2012), Comparing empirical models for sediment and phosphorus transfer from soils to water at field and catchment scale under data uncertainty, *Eur. J. Soil Sci.*, 63, 211–223, doi:10.1111/j.1365-2389.2011.01419.x.
- Laaha, G., and G. Blöschl (2006), A comparison of low flow regionalisation methods—catchment grouping, *J. Hydrol.*, 323(1-4), 193–214, doi:10.1016/j.jhydrol.2005.09.001.
- Lana-Renault, N., and D. Regüés (2009), Seasonal patterns of suspended sediment transport in an abandoned farmland catchment in the Central Spanish Pyrenees, *Earth Surf. Process. Landforms*, 34, 1291–1301, doi:10.1002/esp.1825.
- Lawler, D. M., G. E. Petts, I. D. L. Foster, and S. Harper (2006), Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK, *Sci. Total Environ.*, 360(1–3), 109–126, doi:10.1016/j.scitotenv.2005.08.032.
- Levin, R. B., P. R. Epstein, T. E. Ford, W. Harrington, E. Olson, and E. G. Reichard (2002), U.S. drinking water challenges in the twenty-first century, *Environ. Health Perspect.*, 110(Suppl 1), 43–52.
- Lewis, J., S. R. Mori, E. T. Keppeler, and R. R. Ziemer (2001), Impacts of logging on storm peak flows, flow volumes and suspended sediment loads in Caspar Creek, California, in *Land Use and Watersheds: Human Influence on Hydrology and*

*Geomorphology in Urban and Forest Areas Volume 2*, edited by M. S. Wigmosta and S. J. Burges, pp. 85–125, American Geophysical Union.

- Linsley, R. K. (1982), Rainfall-runoff models - An overview, in *Rainfall-runoff relationship*, edited by V. P. Singh, pp. 3–22, Water Resources Publications, Highlands Ranch, Colorado, USA.
- López-Tarazón, J. A., R. J. Batalla, D. Vericat, and T. Francke (2009), Suspended sediment transport in a highly erodible catchment: The River Isábena (Southern Pyrenees), *Geomorphology*, 109(3–4), 210–221.
- Mather, A. L., and R. L. Johnson (2014), Quantitative characterization of stream turbidity-discharge behavior using event loop shape modeling and power law parameter decorrelation, *Water Resour. Res.*, 50(10), 7766–7779, doi:10.1002/2014WR015417.
- Megnounif, A., A. Terfous, and S. Ouillon (2013), A graphical method to study suspended sediment dynamics during flood events in the Wadi Sebdu, NW Algeria (1973–2004), *J. Hydrol.*, 497, 24–36, doi:10.1016/j.jhydrol.2013.05.029.
- Merz, R., J. Parajka, and G. Blöschl (2009), Scale effects in conceptual hydrological modeling, *Water Resour. Res.*, 45(9), W09405, doi:10.1029/2009WR007872.
- Miller, J. R. (1997), The role of fluvial geomorphic processes in the dispersal of heavy metals from mine sites, *J. Geochemical Explor.*, 58(2–3), 101–118, doi:10.1016/S0375-6742(96)00073-8.
- Money, E. S., G. P. Carter, and M. L. Serre (2009), Modern space/time geostatistics using river distances: Data integration of turbidity and E. coli measurements to assess fecal contamination along the Raritan River in New Jersey, *Environ. Sci. Technol.*, 43(10), 3736–3742, doi:10.1021/es803236j.
- Monk, W., P. Wood, D. Hannah, and D. Wilson (2007), Selection of river flow indices for the assessment of hydroecological change, *River Res. Appl.*, 23, 113–122, doi:10.1002/rra.964.
- Mukundan, R., D. C. Pierson, L. Wang, A. H. Matonse, N. R. Samal, M. S. Zion, and E. M. Schneiderman (2013a), Effect of projected changes in winter streamflow on stream turbidity, Esopus Creek watershed in New York, USA, *Hydrol. Process.*, 27(21), 3014–3023, doi:10.1002/hyp.9824.
- Mukundan, R., D. C. Pierson, E. M. Schneiderman, D. M. O'Donnell, S. M. Pradhanang, M. S. Zion, and A. H. Matonse (2013b), Factors affecting storm event turbidity in a New York City water supply stream, *Catena*, 107, 80–88, doi:10.1016/j.catena.2013.02.002.

- Nash, J. E., and J. V Sutcliffe (1970), River flow forecasting through conceptual models part I — A discussion of principles, *J. Hydrol.*, 10(3), 282–290, doi:10.1016/0022-1694(70)90255-6.
- Neff, R., H. Chang, C. G. Knight, R. G. Najjar, B. Yarnal, and H. A. Walker (2000), Impact of climate variation and change on Mid-Atlantic Region hydrology and water resources, *Clim. Res.*, 14, 207–218.
- New Jersey Water Supply Authority (2002), Raritan Basin Portrait of a Watershed, Available from:  
[http://www.raritanbasin.org/alliance/documents/summary\\_report.pdf](http://www.raritanbasin.org/alliance/documents/summary_report.pdf) (Accessed 1 February 2015)
- Nijssen, B. et al. (2003), Simulation of high latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e): 2: Comparison of model results with observations, *Glob. Planet. Change*, 38(1-2), 31–53, doi:10.1016/S0921-8181(03)00004-3.
- Nistor, C. J., and M. Church (2005), Suspended sediment transport regime in a debris-flow gully on Vancouver Island, British Columbia, *Hydrol. Process.*, 19(4), 861–885, doi:10.1002/hyp.5549.
- NOAA (2014a), Top Ten Highest Historical Crests: Brandywine Creek at Wilmington, Available from:  
[http://www.erh.noaa.gov/marfc/Rivers/FloodClimo/Top\\_Flood\\_Crests/Sch/Wilmington-SEPA-Top10-Table.pdf](http://www.erh.noaa.gov/marfc/Rivers/FloodClimo/Top_Flood_Crests/Sch/Wilmington-SEPA-Top10-Table.pdf) (Accessed 1 February 2015)
- NOAA (2014b), Top Ten Highest Historical Crests: Raritan River at Manville, Available from:  
[http://www.erh.noaa.gov/marfc/Rivers/FloodClimo/Top\\_Flood\\_Crests/Rtn/Manville-RTN-Top10-Table.pdf](http://www.erh.noaa.gov/marfc/Rivers/FloodClimo/Top_Flood_Crests/Rtn/Manville-RTN-Top10-Table.pdf) (Accessed 1 February 2015)
- NOAA National Weather Service (n.d.), River Forecasts, Available from:  
<http://water.weather.gov/ahps/forecasts.php> (Accessed 23 April 2015)
- Oeurng, C., S. Sauvage, and J.-M. Sánchez-Pérez (2010), Dynamics of suspended sediment transport and yield in a large agricultural catchment, southwest France, *Earth Surf. Process. Landforms*, 35(11), 1289–1301, doi:10.1002/esp.1971.
- Olden, J. D., J. J. Lawler, and N. L. Poff (2008), Machine Learning Methods Without Tears: A Primer for Ecologists, *Q. Rev. Biol.*, 83(2), 171–193, doi:10.1086/587826.
- Olden, J. D., M. J. Kennard, and B. J. Pusey (2012), A framework for hydrologic classification with a review of methodologies and applications in ecohydrology, *Ecohydrology*, 5(4), 503–518, doi:10.1002/eco.251.

- Phillips, S. (2002), *The U.S. Geological Survey and the Chesapeake Bay — The Role of Science in Environmental Restoration: U.S. Geographical Survey Circular 1220*.
- Polsky, C., J. Allard, N. Currit, R. Crane, and B. Yarnal (2000), The Mid-Atlantic Region and its climate: past, present, and future, *Clim. Res.*, 14, 161–173.
- Portland Water Bureau (2011), Discover your drinking water, Available from: <http://www.portlandoregon.gov/water/article/225481> (Accessed 1 February 2015)
- Pryor, S. C., and J. T. Schoof (2008), Changes in the seasonality of precipitation over the contiguous USA, *J. Geophys. Res.*, 113(D21), D21108, doi:10.1029/2008JD010251.
- Rankl, J. (2004), *Relations between total-sediment load and peak discharge for rainstorm runoff on five ephemeral streams in Wyoming: U.S. Geological Survey Water-Resources Investigations Report 02-4150*.
- Reidy Liermann, C. A., J. D. Olden, T. J. Beechie, M. J. Kennard, P. B. Skidmore, C. P. Konrad, and H. Imaki (2012), Hydrogeomorphic classification of Washington State rivers to support emerging environmental flow management strategies, *River Res. Appl.*, 28, 1340–1358, doi:10.1002/rra.1541.
- Rinaldo, A., A. Marani, and R. Rigon (1991), Geomorphological dispersion, *Water Resour. Res.*, 27(4), 513–525, doi:10.1029/90WR02501.
- Rodríguez-Blanco, M. L., M. M. Taboada-Castro, L. Palleiro, and M. T. Taboada-Castro (2010), Temporal changes in suspended sediment transport in an Atlantic catchment, NW Spain, *Geomorphology*, 123(1–2), 181–188, doi:10.1016/j.geomorph.2010.07.015.
- Rose, S. (2003), Comparative solute–discharge hysteresis analysis for an urbanized and a “control basin” in the Georgia (USA) Piedmont, *J. Hydrol.*, 284(1–4), 45–56, doi:10.1029/90WR02501.
- Saco, P. M., and P. Kumar (2002), Kinematic dispersion in stream networks 2. Scale issues and self-similar network organization, *Water Resour. Res.*, 38(11), 27–1–27–15, doi:10.1029/2001WR000694.
- Samal, N. R., A. H. Matonse, R. Mukundan, M. S. Zion, D. C. Pierson, R. K. Gelda, and E. M. Schneiderman (2013), Modelling potential effects of climate change on winter turbidity loading in the Ashokan Reservoir, NY, *Hydrol. Process.*, 27(21), 3061–3074, doi:10.1002/hyp.9910.
- Savitzky, A., and M. Golay (1964), Smoothing and differentiation of data by simplified least squares procedures., *Anal. Chem.*, 36(8), 1627–1639, doi:10.1021/ac60214a047.

- Sawicz, K. (2009), *Catchment Classification: An Empirical Analysis of Hydrologic Similarity*, Pennsylvania State University.
- Sawicz, K., C. Kelleher, T. Wagener, P. Troch, M. Sivapalan, and G. Carrillo (2014), Characterizing hydrologic change through catchment classification, *Hydrol. Earth Syst. Sci.*, 18(1), 273–285, doi:10.5194/hess-18-273-2014.
- Seaber, P. R., F. P. Kapinos, and G. L. Knapp (1987), *Hydrologic unit maps: U.S. Geological Survey Water-Supply Paper 2294*.
- Seeger, M., M.-P. Errea, S. Beguería, J. Arnáez, C. Martí, and J. M. García-Ruiz (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, *J. Hydrol.*, 288(3-4), 299–311, doi:10.1016/j.jhydrol.2003.10.012.
- Shamir, E., T. M. Carpenter, P. Fickenscher, and K. P. Georgakakos (2006), Evaluation of the National Weather Service Operational Hydrologic Model and Forecasts for the American River Basin, *J. Hydrol. Eng.*, doi:10.1061/(ASCE)1084-0699(2006)11:5(392).
- Siwek, J., J. P. Siwek, and M. Żelazny (2013), Environmental and land use factors affecting phosphate hysteresis patterns of stream water during flood events (Carpathian Foothills, Poland), *Hydrol. Process.*, 27(25), 3674–3684, doi:10.1002/hyp.9484.
- Sloto, R. A., and M. Y. Crouse (1996), *HYSEP: A computer program for streamflow hydrograph separation and analysis: U.S. Geological Survey Water-Resources Investigations Report 96-4040*.
- Smith, H. G., and D. Dragovich (2009), Interpreting sediment delivery processes using suspended sediment-discharge hysteresis patterns from nested upland catchments, south-eastern Australia, *Hydrol. Process.*, 23, 2415–2426, doi:10.1002/hyp.7357.
- Steege, A., G. Govers, J. Nachtergaele, I. Takken, L. Beuselinck, and J. Poesen (2000), Sediment export by water from an agricultural catchment in the Loam Belt of central Belgium, *Geomorphology*, 33(1–2), 25–36, doi:10.1016/S0169-555X(99)00108-7.
- Syvitski, J. P., M. D. Morehead, D. B. Bahr, and T. Mulder (2000), Estimating fluvial sediment transport: The rating parameters, *Water Resour. Res.*, 36(9), 2747–2760, doi:10.1029/2000WR900133.
- Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res. Atmos.*, 106(D7), 7183–7192, doi:10.1029/2000JD900719.

- Terajima, T., T. Sakamoto, Y. Nakai, and K. Kitamura (1997), Suspended sediment discharge in subsurface flow from the head hollow of a small forested watershed, northern Japan, *Earth Surf. Process. Landforms*, 22(11), 987–1000, doi:10.1002/(SICI)1096-9837(199711)22:11<987::AID-ESP790>3.0.CO;2-X.
- Thomas, R. (1988), Monitoring baseline suspended sediment in forested basins: the effects of sampling on suspended sediment rating curves, *Hydrol. Sci. J.*, 33(5), 499–514, doi:10.1080/02626668809491277.
- Towler, E., B. Rajagopalan, R. S. Summers, and D. Yates (2010), An approach for probabilistic forecasting of seasonal turbidity threshold exceedance, *Water Resour. Res.*, 46(6), W06511, doi:10.1029/2009WR007834.
- U.S. EPA (1999), *Guidance Manual for Compliance with the Interim Enhanced Surface Water Treatment Rule: Turbidity Provisions: U.S. Environmental Protection Agency, EPA 815-R-99-010*.
- VanSickle, J., and R. L. Beschta (1983), Supply-based models of suspended sediment transport in streams, *Water Resour. Res.*, 19(3), 768–778, doi:10.1029/WR019i003p00768.
- Vongvixay, A., C. Grimaldi, C. Gascuel-Oudou, P. Laguionie, M. Faucheux, N. Gilliet, and M. Mayet (2010), Analysis of suspended sediment concentration and discharge relations to identify particle origins in small agricultural watersheds, in *Sediment Dynamics for a Changing Future (Proceedings of the ICCE symposium, SSGGW, Poland 14-18 June 2010)*, IAHR Publ. 337, pp. 76–83.
- Walling, D. E. (2005), Tracing suspended sediment sources in catchments and river systems, *Sci. Total Environ.*, 344, 159–184, doi:10.1016/j.scitotenv.2005.02.011.
- Welles, E., S. Sorooshian, G. Carter, and B. Olsen (2007), Hydrologic verification: A call for action and collaboration, *Bull. Am. Meteorol. Soc.*, 88, 503–511, doi:10.1175/BAMS-88-4-503.
- Wickham, J. D., T. G. Wade, and K. H. Riitters (2011), An environmental assessment of United States drinking water watersheds, *Landsc. Ecol.*, 26(5), 605–616, doi:10.1007/s10980-011-9591-5.
- Williams, G. (1989), Sediment concentration versus water discharge during single hydrologic events in rivers, *J. Hydrol.*, 111, 89–106, doi:10.1016/0022-1694(89)90254-0.
- Wilson, C. G., A. N. T. Papanicolaou, and K. D. Denn (2012), Partitioning fine sediment loads in a headwater system with intensive agriculture, *J. Soils Sediments*, 12(6), 966–981, doi:10.1007/s11368-012-0504-2.

- WWAP (World Water Assessment Programme) (2012), *The United Nations World Water Development Report 4: Managing Water under Uncertainty and Risk*, Paris.
- Yang, T.-M., S.-K. Fan, C. Fan, and N.-S. Hsu (2014), Establishment of turbidity forecasting model and early-warning system for source water turbidity management using back-propagation artificial neural network algorithm and probability analysis., *Environ. Monit. Assess.*, 186(8), 4925–34, doi:10.1007/s10661-014-3748-z.
- Zhang, W., and W. Summer (2002), Sediment transport analysed by energy derived concepts, in *Modelling erosion, sediment transport and sediment yield*, edited by W. Summer and D. Walling, pp. 137–146, UNESCO, Paris.
- Zilberstein, V. (1992), On correlations between the power-law parameters, *Int. J. Fract.*, 58, R57–R59, doi:10.1007/BF00015624.

# APPENDIX A

---

Supporting Information for Chapter 2

**Quantitative Characterization of Stream Turbidity-Discharge Behavior  
using Event Loop Shape Modeling and Power Law Parameter  
Decorrelation**

*Amanda L. Mather and Richard L. Johnson*

Note: Table A1 is shown in parts over the next several pages.

**Table A 1. Study gages and catchment characteristics.**

Gage	USGS ID	Stream	Starting Date of Data Used <sup>a</sup>
G1	01646305	Dead Run	2008-10-01
G2	01645762	S F Little Difficult Run	2007-10-01
G3	01656903	Flatlick Br.	2007-10-01
G4	01645704	Difficult Run	2007-10-01
G5	01649190	Paint Br.	2007-10-03
G6	01480300	W Br. Brandywine C.	2007-10-04
G7	01473169	Valley C.	2007-10-01
G8	01480617	W Br. Brandywine C.	2007-10-01
G9	01472157	French C.	2007-10-01
G10	01480700	E Br. Brandywine C.	2010-11-11
G11	01649500	NE Br. Anacostia R.	2007-10-01
G12	01480870	E Br. Brandywine C.	2007-10-01
G13	01632900	Smith C.	2010-04-16
G14	01481000	Brandywine C.	2007-10-01
G15	01481500	Brandywine C.	2007-10-01
G16	01400500	Raritan R.	2008-04-15
G17	01673000	Pamunkey R.	2007-10-01
G18	01608500	S Br. Potomac R.	2010-10-27
G19	02035000	James R.	2007-10-02
G20	01463500	Delaware R.	2007-10-02

<sup>a</sup> End date of data interval used was 2013-04-22 for all gages.

**Table A 1 (continued). Study Gages and Catchment Characteristics.**

Gage	Drainage Area <sup>b</sup>	Drainage Density <sup>c</sup>	Average Slope <sup>d</sup>	Average Elevation <sup>d</sup>
	(km <sup>2</sup> )	(km/km <sup>2</sup> )	(deg)	(m)
	DA	DD	SLP	ELEV
G1	5.31	0.73	2.79	85
G2	7.02	1.15	3.32	120
G3	10.9	1.33	1.96	111
G4	14.2	1.25	3.15	124
G5	33.9	1.26	2.61	122
G6	48.4	0.67	2.29	221
G7	53.8	0.98	4.02	108
G8	143	0.82	3.28	193
G9	153	0.94	4.21	161
G10	157	0.99	3.26	170
G11	188	1.25	2.18	67
G12	233	1.05	3.48	157
G13	242	1.67	5.53	422
G14	743	1.11	2.79	149
G15	813	1.19	3.4	144
G16	1269	1.84	2.69	135
G17	2799	1.65	1.7	101
G18	3847	1.56	9.12	398
G19	16199	1.72	5.02	290
G20	17553	1.17	2.91	385

<sup>b</sup> USGS NWIS<sup>c</sup> DD = (total stream length)/(DA); Total stream length from USGS NHD streamlines.<sup>d</sup> Global Mapper; USGS NED DEM.

**Table A 1 (continued). Study Gages and Catchment Characteristics.**

Gage	Mean Annual	Streamflow	Streamflow	Baseflow Index <sup>h</sup>
	Runoff <sup>e</sup>	Variability <sup>f</sup>	Skew <sup>g</sup>	
	(m)	(-)	(-)	
	MAR	SV	SKEW	BFI
G1	0.39	3.45	2.14	0.433
G2	0.35	3.25	1.763	0.626
G3	0.57	3.35	2.41	0.465
G4	0.55	3.43	2.346	0.449
G5	0.33	1.8	1.741	0.574
G6	0.58	1.96	1.744	0.576
G7	0.62	1.56	1.428	0.723
G8	0.53	1.37	1.432	0.695
G9	0.57	1.52	1.559	0.669
G10	0.57	1.12	1.344	0.724
G11	0.42	2.49	2.443	0.435
G12	0.62	1.23	1.405	0.718
G13	0.23	1.64	1.803	0.708
G14	0.57	1.25	1.421	0.734
G15	0.57	1.23	1.395	0.73
G16	0.67	1.72	1.807	0.625
G17	0.24	1.6	2.155	0.619
G18	0.31	1.65	2.037	0.633
G19	0.33	1.27	1.811	0.682
G20	0.74	1	1.386	0.741

<sup>e</sup> USGS NWIS daily streamflow<sup>f</sup> SV = st.dev./mean; USGS NWIS daily streamflow<sup>g</sup> SKEW = mean/median; USGS NWIS daily streamflow<sup>h</sup> *Arnold et al. [1995]* as in *Sawicz et al. [2011]*; USGS NWIS daily streamflow

**Table A 1 (continued). Study Gages and Catchment Characteristics.**

Gage	Forest <sup>i</sup>	Urban <sup>i</sup>	Water & Wetlands <sup>i</sup>	Crops <sup>i</sup>	Pasture <sup>i</sup>
	(%)	(%)	(%)	(%)	(%)
	FOREST	URBAN	WATWET	CROPS	PASTURE
G1	26.4	70.4	0.9	0	2.2
G2	55.7	38	3.6	0.3	1.2
G3	13.4	82.9	0.7	0.2	1.6
G4	35	60.4	1.8	0.9	1.3
G5	32.7	56.6	2.2	2.4	4.6
G6	21.3	10.4	2.6	32.2	27.8
G7	31.2	51.8	1.2	3.7	5.1
G8	33.6	19.8	2	17.6	20.2
G9	49.8	9	2.3	11.7	17.7
G10	37.1	15.2	4.2	15.7	19.5
G11	24.7	62.7	5.1	1.4	4.2
G12	36.3	22.6	3.3	13.6	16.7
G13	45.9	7.9	0.1	3	43.2
G14	31.9	23.2	2.1	13.5	22
G15	32.9	23.4	2.2	12.7	21.5
G16	40.4	21.5	5.1	13.9	13.9
G17	57.9	5	6.8	5.3	14.8
G18	80.6	3.9	0.5	0.4	14.5
G19	76.6	6.4	1.1	0.4	13.9
G20	67.5	10.1	5.7	6.3	8.8

<sup>i</sup>NLCD 2006. (Categories shown are not inclusive and may not total 100%).

**Table A 1 (continued). Study Gages and Catchment Characteristics.**

Gage	HSG A Soils <sup>†</sup>	HSG B Soils <sup>†</sup>	HSG C Soils <sup>†</sup>	HSG D Soils <sup>†</sup>
	(%)	(%)	(%)	(%)
	SOILA	SOILB	SOILC	SOILD
G1	1.7	91.7	4.3	2.3
G2	5.4	70.5	20.4	3.8
G3	0	9.8	80.2	9.9
G4	3.9	52.8	13.4	29.9
G5	5.4	70.7	18.7	5.2
G6	5.2	75.6	9.5	9.7
G7	1.8	79.6	12.4	6.1
G8	2.5	68.2	16.8	12.5
G9	5.2	64.6	19.8	10.4
G10	5.8	74.5	9.2	10.6
G11	2.7	42.2	33.1	22
G12	5.2	75.1	9.7	10.1
G13	0.1	52.2	41.3	6.4
G14	4.5	72.8	11.8	10.9
G15	4.4	72.3	12.1	11.2
G16	0.5	49.1	39.2	10.1
G17	0.6	50.9	36.6	10
G18	14.3	21.2	54.9	9.7
G19	0.3	54.8	34.5	10.3
G20	4.1	15.8	67.3	11.7

<sup>†</sup> www.soilinfo.psu.edu; Hydrologic Soil Groups (HSG) from NRCS STATSGO database (Categories shown by not total 100% as areas of open water are omitted).

## References (APPENDIX A Only)

- Arnold, J. G., Allen, P. M., Muttiah, R., & Bernhardt, G. (1995). Automated Base Flow Separation and Recession Analysis Techniques. *Ground Water*, 33(6), 1010–1018. doi:10.1111/j.1745-6584.1995.tb00046.x
- Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. a., & Carrillo, G. (2011). Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences*, 15(9), 2895–2911. doi:10.5194/hess-15-2895-2011

# APPENDIX B

---

Supporting Information for Chapter 3

## **Characterization and Exploratory Analysis of Stream Turbidity during 5928 Events from 110 U.S. Gages**

*Amanda L. Mather and Richard L. Johnson*

### **Section B.1: Gage Data**

Table B 1 lists the 110 U.S. Geological Survey (USGS) gages used in this study, and Figure 3.1 in Chapter 3 shows the gage locations and catchment outlines. The “NCDC Station ID” corresponds to the NOAA National Climatic Data Center (NCDC) station used in the calculation of several meteorological catchment characteristics as described in Section B.2. The “Expanded Event Window” is discussed in Section 3.3.1 of Chapter 3.

**Table B 1. Gage data.**

USGS Gage ID	Stream Name	Drainage Area (km <sup>2</sup> )	Starting date of data used	Ending date of data used	NCDC Station ID	Expanded Event Window?
01104455	STONY BROOK	1.24	2007-10-01	2013-05-18	USW00014702	
01400500	RARITAN RIVER	1269.10	2008-04-15	2013-04-22	USC00283029	
01408029	MANASQUAN RIVER	163.95	2007-10-01	2013-11-08	USC00283181	
01408500	TOMS RIVER	318.57	2011-09-21	2013-11-08	USC00283181	3x
01463500	DELAWARE RIVER	17560.20	2007-10-02	2013-04-22	USC00304731	
01472157	FRENCH CREEK	153.07	2007-10-01	2013-04-22	USC00369464	
01473169	VALLEY CREEK	53.87	2007-10-01	2013-04-22	USC00369464	
01473500	SCHUYLKILL RIVER	4558.40	2012-02-23	2013-11-08	USC00363632	
01480300	WEST BRANCH BRANDYWINE CREEK	48.43	2007-10-04	2013-04-22	USC00369464	
01480617	WEST BRANCH BRANDYWINE CREEK	143.23	2007-10-01	2013-04-22	USC00369464	
01480700	EAST BRANCH BRANDYWINE CREEK	156.95	2010-11-11	2013-04-22	USC00369464	
01480870	EAST BRANCH BRANDYWINE CREEK	232.84	2007-10-01	2013-04-22	USC00369464	
01481000	BRANDYWINE CREEK	743.33	2007-10-01	2013-04-22	USC00369464	
01481500	BRANDYWINE CREEK	813.26	2007-10-01	2013-04-22	USC00369464	
01493112	CHESTERVILLE BRANCH	15.85	2012-06-09	2013-11-08	USW00013701	
01548303	STRAIGHT RUN	16.78	2012-07-18	2013-11-08	USC00369408	
01608500	SOUTH BRANCH POTOMAC RIVER	3848.74	2010-10-27	2013-04-23	USC00467730	
01632900	SMITH CREEK	242.42	2010-04-16	2013-04-23	USC00442208	
01645704	DIFFICULT RUN	14.22	2007-10-01	2013-04-23	USC00448737	
01645762	S F LITTLE DIFFICULT RUN	7.02	2007-10-01	2013-04-23	USC00448737	
01646000	DIFFICULT RUN	149.70	2012-10-10	2013-11-21	USC00448737	
01646305	DEAD RUN	5.31	2008-01-10	2013-04-23	USC00182325	
01648010	ROCK CREEK	164.98	2012-10-01	2013-11-08	USC00186350	
01649190	PAINT BRANCH	33.93	2007-10-03	2013-04-23	USC00180700	
01649500	NORTH EAST BRANCH ANACOSTIA RIVER	188.55	2007-10-01	2013-04-25	USC00180700	
01656903	FLATLICK BRANCH	10.88	2007-10-01	2013-04-23	USC00448737	
01673000	PAMUNKEY RIVER	2799.79	2007-10-04	2013-04-23	USC00445050	2x
02035000	JAMES RIVER	16205.63	2007-10-02	2013-04-23	USC00448600	
02203863	SHOAL CREEK	22.53	2012-03-19	2013-10-01	USC00095666	
02203900	SOUTH RIVER	256.41	2012-04-04	2013-10-01	USC00095666	
02204130	HONEY CREEK	67.34	2011-05-03	2012-08-16	USC00095666	
02205522	PEW CREEK	18.13	2007-10-01	2013-10-01	USC00098950	
02207335	YELLOW RIVER	673.40	2010-10-01	2012-07-16	USC00095666	
02208130	SHOAL CREEK	10.10	2007-10-01	2013-10-01	USC00098950	
02336152	SOUTH FORK PEACHTREE	15.28	2012-03-17	2013-10-01	USC00095666	
02337410	DOG RIVER	172.24	2007-10-01	2013-10-01	USC00091640	
03067510	SHAVERS FORK	155.92	2011-12-13	2013-10-31	USC00468308	
03238140	TAYLOR CREEK	10.98	2007-10-03	2013-10-23	USW00093812	
03238745	TWELVEMILE CREEK	101.01	2007-10-01	2013-10-24	USW00093812	
03238772	FOURMILE CREEK	8.03	2007-10-04	2012-07-16	USW00093812	
03254550	BANKLICK CREEK	77.70	2007-10-06	2013-10-23	USW00093814	
03260015	PLEASANT RUN CREEK	16.55	2007-10-06	2013-10-31	USW00093814	
03262001	WOOLPER CREEK	62.68	2007-10-16	2013-10-31	USW00093814	
03277075	GUNPOWDER CREEK	94.79	2007-10-09	2013-10-23	USW00093814	
03353200	EAGLE CREEK	266.77	2010-09-30	2013-10-31	USC00129557	
03374100	WHITE RIVER	29279.95	2010-12-03	2013-09-08	USC00120784	3x
03408500	NEW RIVER	989.38	2007-10-01	2013-10-31	USW00053868	
03432100	HARPETH RIVER	172.49	2008-10-01	2013-05-06	USC00401720	
03601630	LOCKE BRANCH	2.10	2007-10-01	2012-10-17	USC00402989	
04027000	BAD RIVER	1546.23	2011-03-02	2013-11-07	USC00475286	
04040000	ONTONAGON RIVER	3470.60	2011-07-08	2013-10-31	USC00208706	
04087030	MENOMONEE RIVER	89.87	2008-12-29	2013-10-03	USC00473058	2x
04087050	LITTLE MENOMONEE RIVER	20.72	2008-11-08	2013-10-03	USC00473058	
04087088	UNDERWOOD CREEK	46.88	2010-02-18	2013-10-03	USC00475474	
04087119	HONEY CREEK	26.68	2008-12-06	2013-09-20	USC00475474	
04087120	MENOMONEE RIVER	318.57	2008-11-05	2013-10-03	USC00475474	

**Table B 1 (continued). Gage data.**

USGS Gage ID	Stream Name	Drainage Area (km <sup>2</sup> )	Starting date of data used	Ending date of data used	NCDC Station ID	Expanded Event Window?
04142000	RIFLE RIVER	828.80	2011-06-08	2013-10-03	USC00208800	
04165500	CLINTON RIVER	1901.06	2011-07-15	2013-11-07	USC00206982	
04166500	RIVER ROUGE	484.33	2011-04-01	2013-10-03	USC00202691	
04195500	PORTAGE RIVER	1108.52	2011-04-23	2013-09-19	USC00330862	2x
04199500	VERMILION RIVER	678.58	2011-06-23	2013-10-03	USW00004849	
04200500	BLACK RIVER	1025.64	2011-06-23	2013-10-03	USW00004849	
04206425	CUYAHOGA RIVER	1437.45	2012-05-24	2013-10-03	USW00014813	
04208000	CUYAHOGA RIVER	1831.13	2011-02-17	2013-10-03	USW00014813	
05054000	RED RIVER OF THE NORTH	17612.00	2007-10-01	2013-11-13	USC00218947	2x
05517000	YELLOW RIVER	1126.65	2012-05-10	2013-11-13	USC00126989	2x
06795500	SHELL CREEK	699.30	2008-03-29	2012-10-21	USC00251825	
06800500	ELKHORN RIVER	15203.30	2008-03-27	2013-10-16	USW00014941	
06893300	INDIAN CREEK	68.89	2011-05-18	2012-08-15	USW00003967	
06893350	TOMAHAWK CREEK	61.90	2011-07-08	2012-08-15	USW00003967	
06893390	INDIAN CREEK	166.20	2007-10-02	2013-11-08	USW00003967	
06893820	LITTLE BLUE RIVER	254.34	2009-10-08	2013-11-08	USC00234154	
06893830	ADAIR CREEK	13.31	2008-10-11	2013-11-08	USC00234154	
06893890	EAST FORK LITTLE BLUE RIVER	89.10	2009-12-01	2013-11-08	USC00234154	
06893970	SPRING BRANCH CREEK	21.76	2007-10-05	2013-11-08	USC00234154	
06894000	LITTLE BLUE RIVER	476.56	2007-10-03	2013-11-08	USW00053879	
06914950	BIG BULL CREEK	74.33	2010-10-23	2012-08-15	USW00093909	
06914990	LITTLE BULL CREEK	20.36	2010-09-01	2012-08-15	USW00093909	
07017610	BIG RIVER	1059.31	2011-10-14	2013-11-13	USC00236826	
07018500	BIG RIVER	2375.03	2011-10-14	2013-11-13	USC00236826	
07061270	EAST FORK BLACK RIVER	135.20	2007-10-01	2013-11-08	USC00236826	
07075270	SOUTH FORK OF LITTLE RED RIVER	193.47	2010-07-24	2013-11-08	USC00032794	
07143672	LITTLE ARKANSAS RIVER	1965.81	2007-10-01	2013-11-08	USC00145152	2x
07144100	LITTLE ARKANSAS RIVER	3209.01	2007-10-01	2013-11-08	USC00145152	
07191222	BEATY CREEK	153.12	2007-10-01	2013-11-08	USC00032930	
07263296	MAUMELLE RIVER	189.07	2007-10-09	2013-11-08	USC00030130	
08048000	W FK TRINITY RIVER	6773.29	2010-08-21	2013-11-08	USC00419532	
08068000	W FK SAN JACINTO RIVER	2144.36	2009-08-18	2013-11-08	USC00414382	
08068500	SPRING CREEK	1059.26	2007-10-01	2013-11-08	USW00053910	
08070200	E FK SAN JACINTO RIVER	1004.92	2007-10-01	2013-11-08	USW00053902	2x
08188060	SAN ANTONIO RIVER	9246.30	2011-07-27	2013-11-08	USW00012921	
11274790	TUOLUMNE RIVER	779.59	2007-10-01	2013-11-02	USW00053150	
11462500	RUSSIAN RIVER	937.58	2008-07-19	2013-11-02	USW00023275	
11467000	RUSSIAN RIVER	3470.60	2008-06-07	2013-11-02	USW00023275	
11501000	SPRAGUE RIVER	4053.35	2007-11-17	2013-11-02	USC00353232	2x
11502500	WILLIAMSON RIVER	7770.00	2007-10-17	2013-11-02	USC00353232	2x
12100490	WHITE RIVER	1230.25	2010-05-27	2013-09-05	USC00456295	
12101500	PUYALLUP RIVER	2455.32	2012-05-03	2013-09-05	USC00456295	
12186000	SAUK RIVER	393.68	2011-09-23	2013-09-05	USC00451992	
12214500	SUMAS RIVER	80.29	2011-04-21	2013-09-05	USC00451484	
14182500	LITTLE NORTH SANTIAM RIVER	290.08	2007-10-01	2013-09-05	USC00352292	
14206950	FANNO CREEK	81.59	2007-10-03	2013-09-05	USC00355945	
14210000	CLACKAMAS RIVER	1737.89	2007-10-01	2013-09-05	USC00352693	
14211010	CLACKAMAS RIVER	2434.60	2007-10-01	2013-09-05	USC00352693	
14240525	NF TOUTLE RIVER	453.25	2010-05-01	2013-09-05	USC00451760	
14242580	TOUTLE RIVER	1284.64	2010-04-01	2013-09-05	USC00451760	
14316500	N UMPQUA RIVER	1230.25	2010-06-15	2012-10-14	USC00354835	
14338000	ELK CREEK	344.47	2008-01-26	2011-06-05	USC00355055	
14361500	ROGUE RIVER	6345.50	2012-09-06	2013-03-14	USC00355055	
071912213	SPAVINAW CREEK	420.95	2007-10-01	2013-11-08	USC00032930	

## **Section B.2: Catchment Data**

A total of 29 catchment characteristics were determined for each of the 110 study gages. A summary of these results is shown in Table 3.1 of Chapter 3. The details of the data sources and methods used to determine each catchment characteristics are discussed in the following sections.

### **Data Sources**

The GIS software Global Mapper was used to delineate each catchment based on gage location and digital elevation data from the USGS National Elevation Dataset (NED) (10m resolution), which was used to calculate several topographic and morphologic characteristics (SF, DD, SLP, ELEV, RELIEF and RELIEF RATIO).

Land cover data was obtained from the USGS 2006 National Land Cover Database (NLCD) (30m resolution) and was used to calculate the fraction of land cover classes present (FOREST, AG, URBAN and WETWAT).

Soil data from the Center for Environmental Informatics ([www.soilinfo.psu.edu](http://www.soilinfo.psu.edu)); Hydrologic Soil Groups (HSG) from Natural Resources Conservation Service (NRCS) State Soil Geographic (STATSGO) database) was used to calculate the fraction of HSG classes present (SOILA, SOILB, SOILC and SOILD).

Daily meteorological data was obtained from the NOAA National Climatic Data Center (NCDC) from the available station closest to the centroid of the catchment (Table B 1) for the 6-year period of 2007-10-01 to 2013-09-30 and used to calculate catchment climate characteristics (PSI and AI).

Climate 30-year (1981-2010) normals from the PRISM Climate Group (<http://prism.oregonstate.edu>) were used to calculate several catchment climate characteristics (MAT, MAP and RR). These characteristics could have been calculated from NCDC station data; however, spatial datasets were used throughout this study, when available, with an eye towards future un-gaged applications.

Surface water features (e.g., streamlines, water bodies) were obtained from the National Hydrography Dataset (NHD) and used to calculate several characteristics related to the drainage network and impoundments present (DD, ID, IA and IMA).

Mean daily discharge data was obtained from the USGS for the 6-year period of 2007-10-01 to 2013-09-30 and used to compute the flow characteristics (MAR, RR, SV, SKEW and BFI).

### **Catchment Characteristics**

**LAT** – Gage latitude, (degrees); as published by USGS.

**LONG** – Gage longitude, (degrees); as published by USGS.

**DA** – Drainage area, (km<sup>2</sup>); as published by USGS.

**SF** – Shape factor, (-); defined as the catchment width divided by the catchment length; calculated as the catchment drainage area (DA) divided by the catchment length squared. The catchment length was the distance from the gage to the most distant catchment boundary, along a line approximately following the main stem of the stream.

**DD** – Drainage density, (km/km<sup>2</sup>); calculated as the total length of USGS NHD streamlines within the catchment bounds divided by the catchment drainage area (DA).

**SLP** – Mean catchment slope, (degrees); determined using Global Mapper and the USGS NED.

**ELEV** – Mean catchment elevation, (m); determined using Global Mapper and the USGS NED.

**RELIEF** – Relief, (m); calculated as the maximum elevation minus the minimum elevation. This was determined using Global Mapper and the USGS NED.

**RELIEF RATIO** – Relief ratio, (-); calculated as the relief (RELIEF) divided by the catchment length. The catchment length was the distance from the gage to the most distant catchment boundary, along a line approximately following the main stem of the stream.

**SOILA** – Fraction of type A SCS hydrologic soil group, (%); spatial data obtained from the Center for Environmental Informatics, proportions within catchment bounds determined using Global Mapper.

**SOILB** – Fraction of type B SCS hydrologic soil group, (%); spatial data obtained from the Center for Environmental Informatics, proportions within catchment bounds determined using Global Mapper.

**SOILC** – Fraction of type C SCS hydrologic soil group, (%); spatial data obtained from the Center for Environmental Informatics, proportions within catchment bounds determined using Global Mapper.

**SOILD** – Fraction of type D SCS hydrologic soil group, (%); spatial data obtained from the Center for Environmental Informatics, proportions within catchment bounds

determined using Global Mapper. Mixed classes (i.e., A/D, B/D and C/D) were considered class D.

**FOREST** – Fraction of forest land cover, (%); spatial data obtained from USGS 2006 National Land Cover Database, includes: Deciduous Forest, Evergreen Forest and Mixed Forest classifications.

**AG** – Fraction of agriculture land cover, (%); spatial data obtained from USGS 2006 National Land Cover Database, includes: Cultivated Crops and Pasture/Hay classifications.

**URBAN** – Fraction of urban land cover, (%); spatial data obtained from USGS 2006 National Land Cover Database, includes: Developed Open Space, Developed Low Intensity, Developed Medium Intensity and Developed High Intensity classifications.

**WETWAT** – Fraction of wetland or water land cover, (%); spatial data obtained from USGS 2006 National Land Cover Database, includes: Open Water, Woody Wetlands and Emergent Herbaceous Wetlands classifications.

**MAT** – Mean annual temperature, (°C); calculated from PRISM spatial dataset based on catchment boundaries using Global Mapper.

**MAP** – Mean annual precipitation, (m); calculated from PRISM spatial dataset based on catchment boundaries using Global Mapper.

**PSI** – Precipitation seasonality index, (-); calculated from the NCDC dataset, where MAP is the mean annual precipitation and  $x_n$  is the mean monthly precipitation [*Walsh and Lawler, 1981; Sawicz et al., 2011*]:

$$PSI = \frac{1}{MAP} \sum_{n=1}^{12} \left| x_n - \frac{MAP}{12} \right|$$

**AI** – Aridity index, (-); calculated as the mean annual potential evapotranspiration (PET) divided by the mean annual precipitation (MAP). The PET was calculated from the NCDC dataset daily average temperature using the Hargreaves method [*Hargreaves and Samani, 1985*]. Extraterrestrial radiation was estimated from gage latitude and Julian day [*Allen et al., 1998*].

**ID** – Impoundment density, (number/km<sup>2</sup>); calculated as the number of USGS NHD waterbodies classified as “lake/pond” or “reservoir” within (partly or fully) the catchment divided by the catchment drainage area (DA).

**IA** – Fraction of impoundment area, (%); calculated as the total area of all USGS NHD waterbodies classified as “lake/pond” or “reservoir” within (partly or fully) the catchment divided by the catchment drainage area (DA). (Note: this catchment characteristic is similar to WETWAT above, but it is derived from a different dataset and does not include “wetlands”.)

**IMA** – Fraction of largest impoundment area, (%); calculated as the area of the largest USGS NHD waterbody classified as “lake/pond” or “reservoir” within (partly or fully) the catchment divided by the catchment drainage area (DA).

**MAR** – Mean annual runoff, (m); calculated as the sum of the USGS mean daily discharge dataset divided by the number of years in the data period and divided by the catchment drainage area (DA).

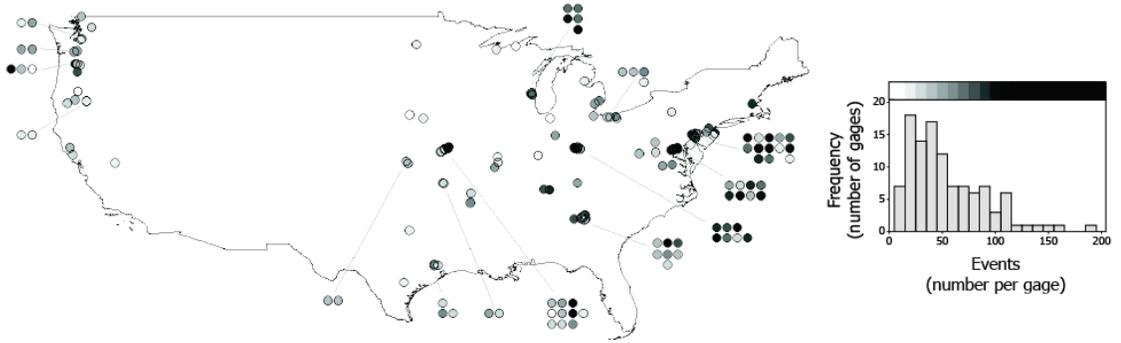
**RR** – Runoff ratio, (-); calculated as the mean annual runoff (MAR) divided by the mean annual precipitation (MAP).

**SV** – Streamflow variability, (-); calculated as the standard deviation of the USGS mean daily discharge dataset divided by the mean of the USGS mean daily discharge dataset.

**SKEW** – Skewness in daily flows, (-); calculated as the mean of the USGS mean daily discharge dataset divided by the median of the USGS mean daily discharge dataset.

**BFI** – Baseflow index, (-); calculated from the USGS mean daily discharge dataset using the methods of *Arnold et al.* [1995] as described in *Sawicz et al.* [2011].

### Section B.3: Study Events



**Figure B 1. Number of events per gage used in this study. Histogram displays distribution of number of events per gage and shows color key for map of gage event totals.**

## References (APPENDIX B Only)

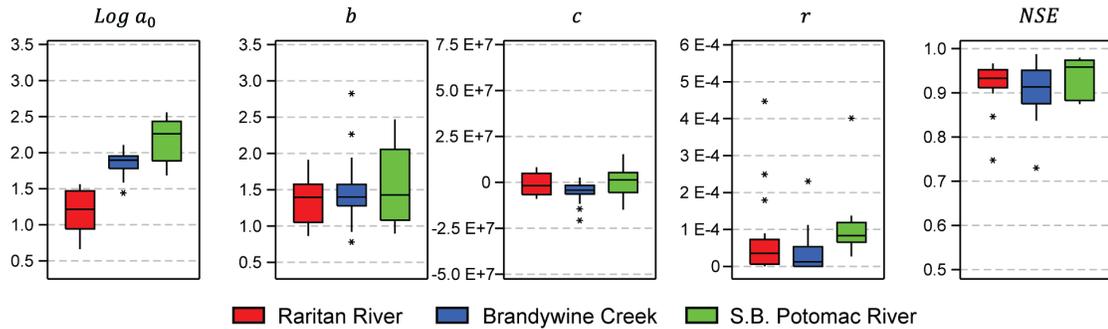
- Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), *Crop evapotranspiration - Guidelines for computing crop water requirements: FAO Irrigation and Drainage Paper 56*, Food and Agriculture Organization of the United Nations, Rome.
- Arnold, J. G., P. M. Allen, R. Muttiah, and G. Bernhardt (1995), Automated Base Flow Separation and Recession Analysis Techniques, *Ground Water*, 33(6), 1010–1018, doi:10.1111/j.1745-6584.1995.tb00046.x.
- Hargreaves, G. H., and Z. Samani (1985), Reference Crop Evapotranspiration from Ambient Air Temperature, in *American Society of Agricultural Engineering Meeting, Chicago, 1985*, vol. 24, p. paper 85–2517.
- Sawicz, K., T. Wagener, M. Sivapalan, P. A. Troch, and G. Carrillo (2011), Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, *Hydrol. Earth Syst. Sci.*, 15(9), 2895–2911, doi:10.5194/hess-15-2895-2011.
- Walsh, R. P. D., and D. M. Lawler (1981), Rainfall Seasonality: Description, Spatial Patterns and Change Through Time, *Weather*, 36(7), 201–208, doi:10.1002/j.1477-8696.1981.tb05400.x.

# APPENDIX C

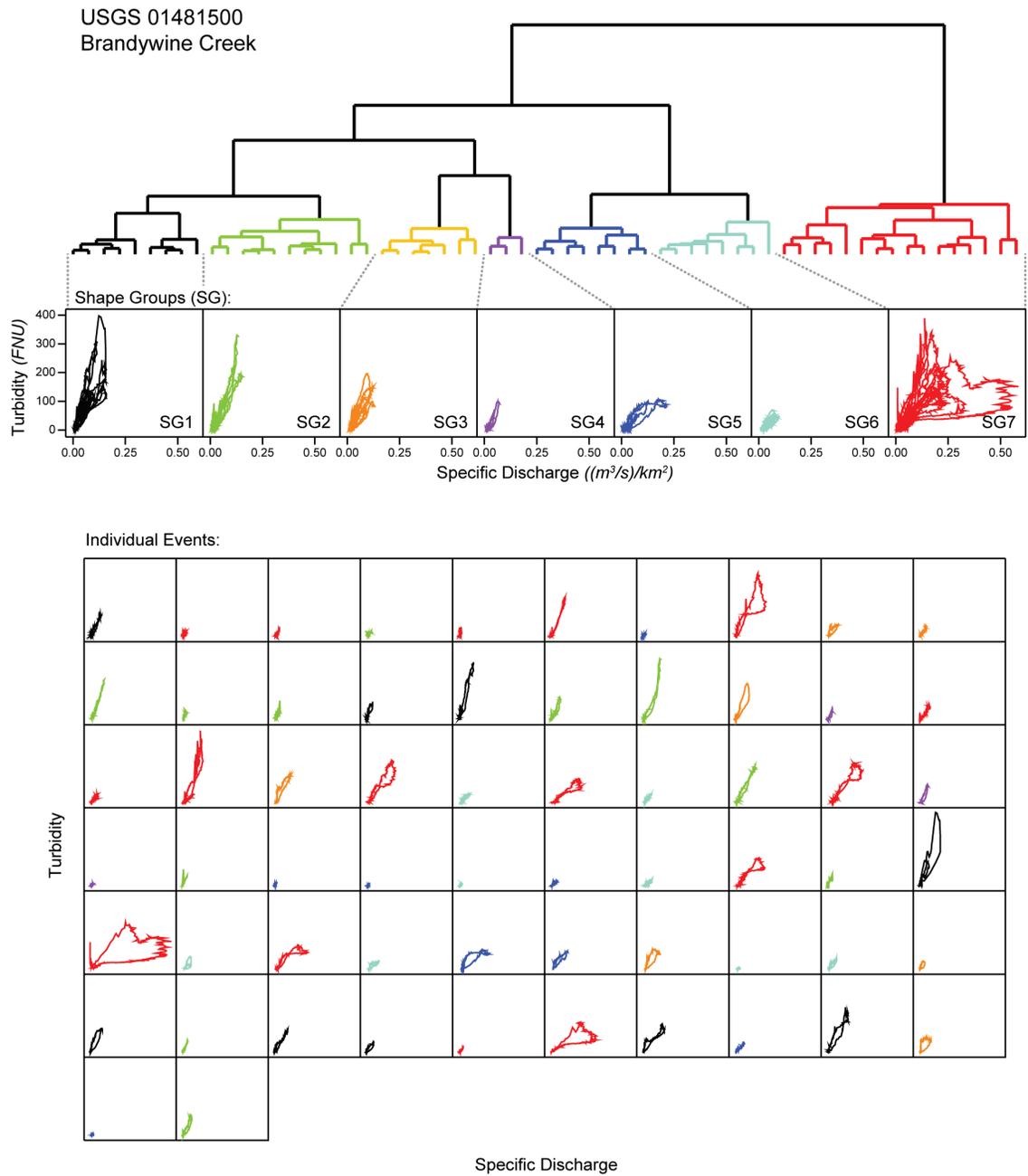
## Supporting Information for Chapter 4

### Event-based Prediction of Stream Turbidity using Regression and Classification Tree Approaches

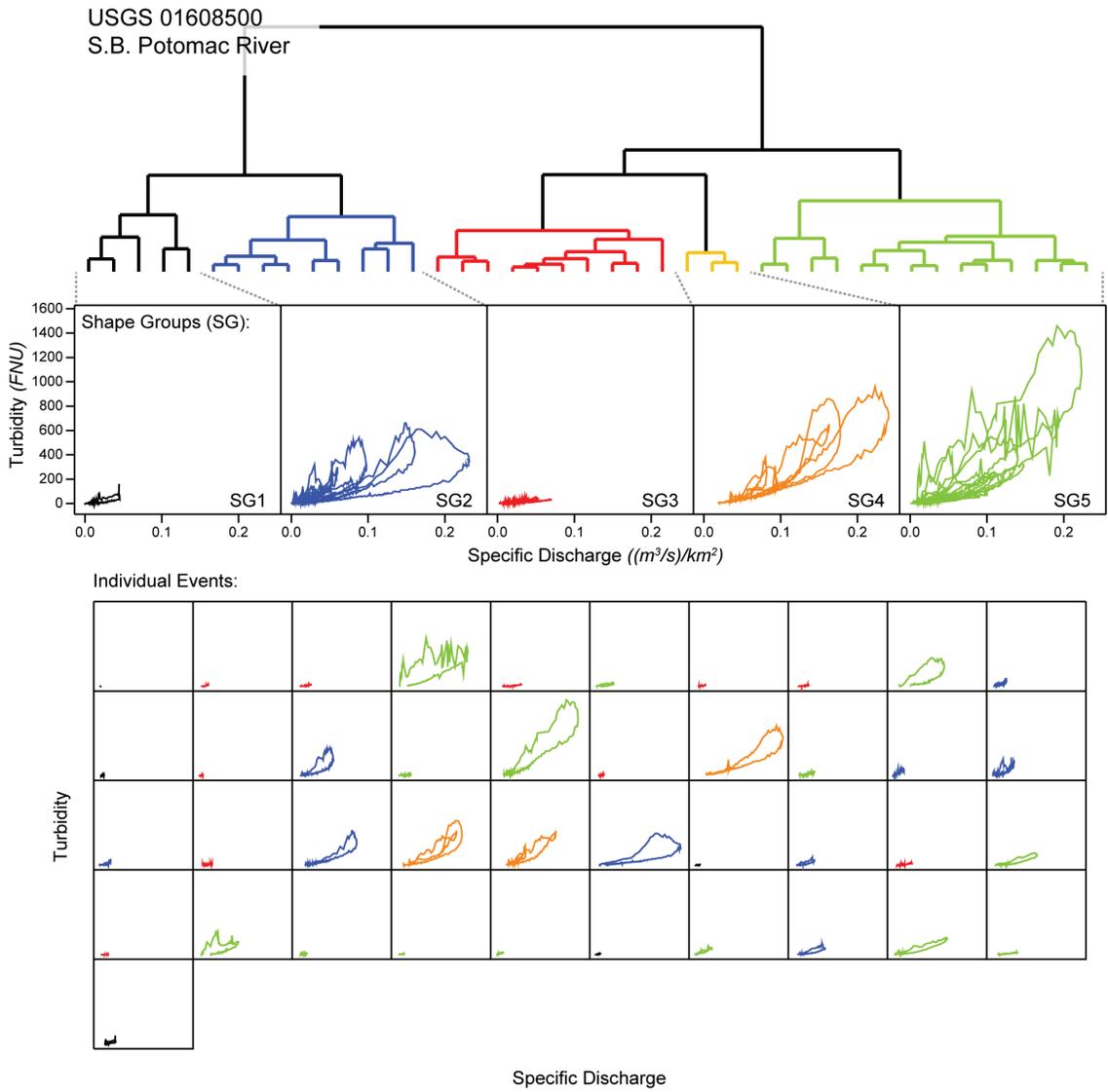
*Amanda L. Mather and Richard L. Johnson*



**Figure C 1. Boxplots of best-fit event model parameter values for the events in the validation datasets:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)/(s/m<sup>3</sup>)),  $r$  (km<sup>2</sup>/m<sup>3</sup>) and  $NSE$  (-). Boxplot boxes contain the middle 50% of the data (with the median shown as a line within the box), whiskers extend to the nearest data point within 1.5 box lengths from box edge and stars denote events which fall outside the upper and lower whiskers (outliers); boxplot colors indicate stream.**



**Figure C 2. Dendrogram showing cluster analysis results and the events associated with each cluster (shape group) for Brandywine Creek. The events are shown as turbidity vs. discharge loops and are colored by shape group. The small plots of individual events all have the same axis scaling.**



**Figure C 3. Dendrogram showing cluster analysis results and the events associated with each cluster (shape group) for S.B. Potomac River. The events are shown as turbidity vs. discharge loops and are colored by shape group. The small plots of individual events all have the same axis scaling.**

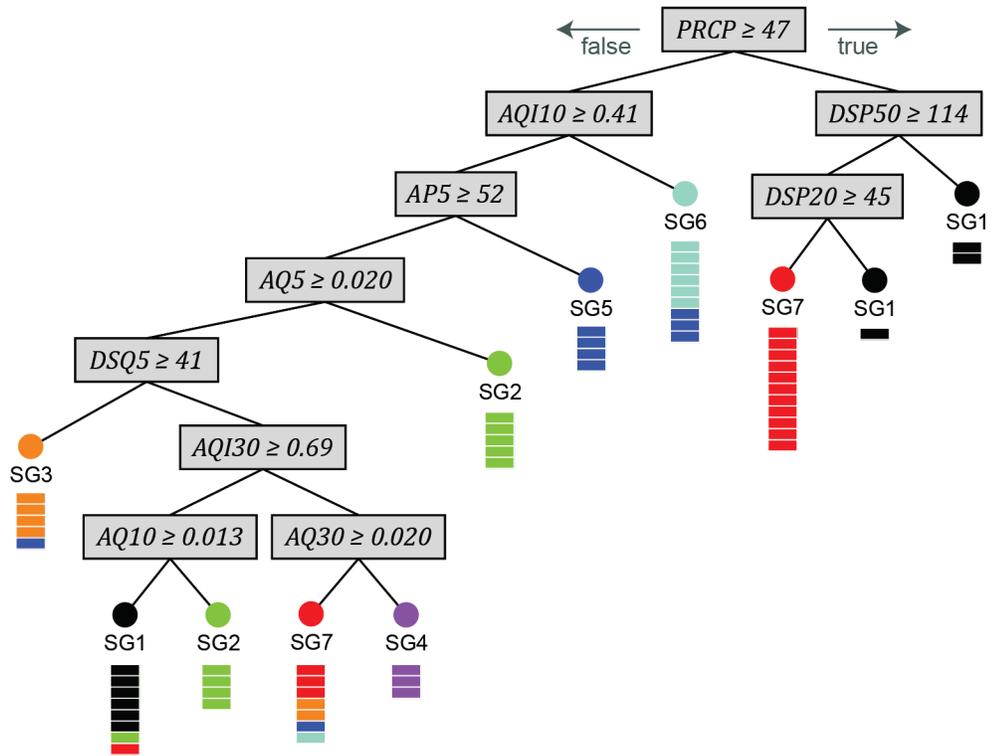


Figure C 4. Classification Tree for Brandywine Creek. Terminal node (circles) class identity colored by shape group, as shown in Figure C 2. Membership for training events shown with rectangles representing individual events under each terminal node. Only 59 of the total 62 events are shown due to three events having missing characteristic data for a branch the event encounters.

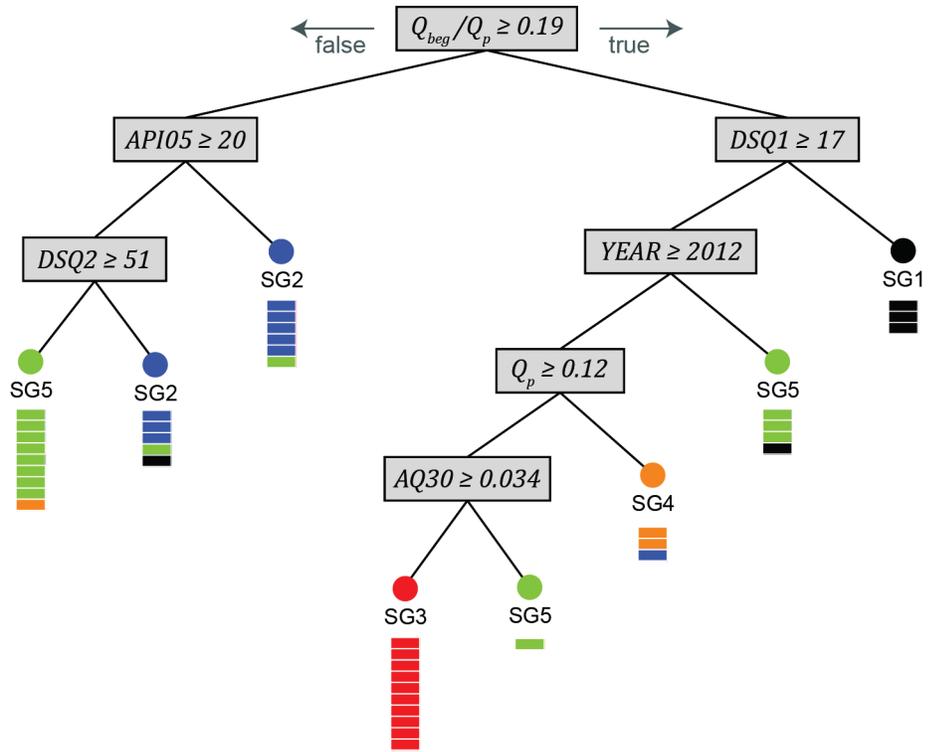


Figure C 5. Classification Tree for S.B. Potomac River. Terminal node (circles) class identity colored by shape group, as shown in Figure C 3. Membership for training events shown with rectangles representing individual events under each terminal node.

### Raritan River

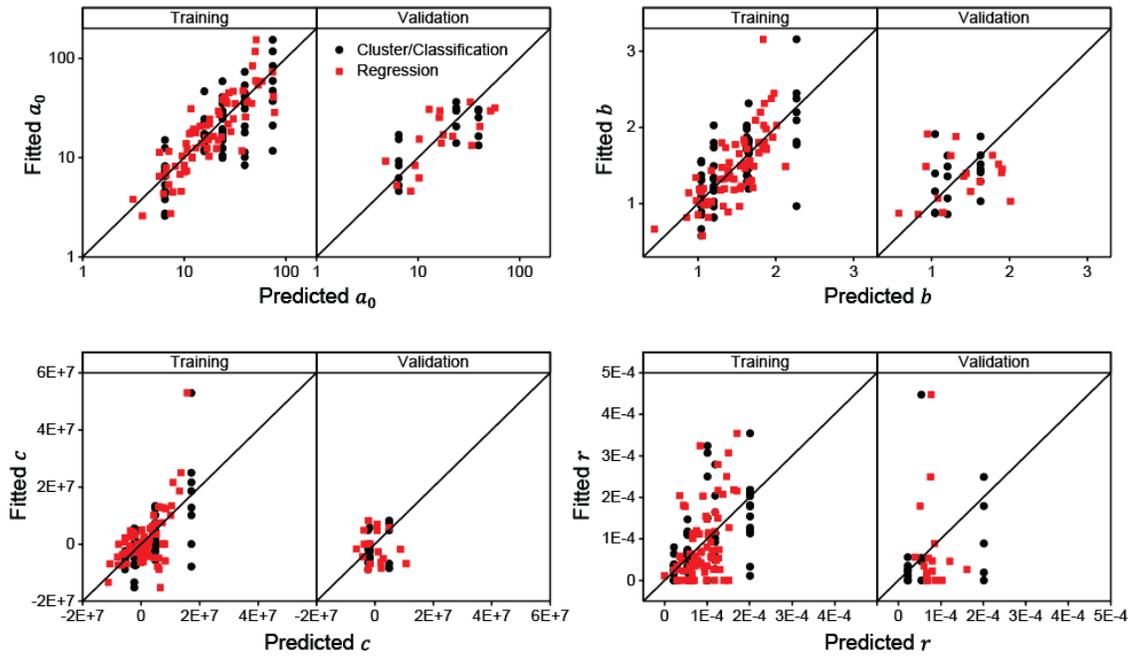


Figure C 6. Comparison of predicted parameter values and fitted parameter values for Raritan River. Units are:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)) and  $r$  (km<sup>2</sup>/m<sup>3</sup>).

### Brandywine Creek

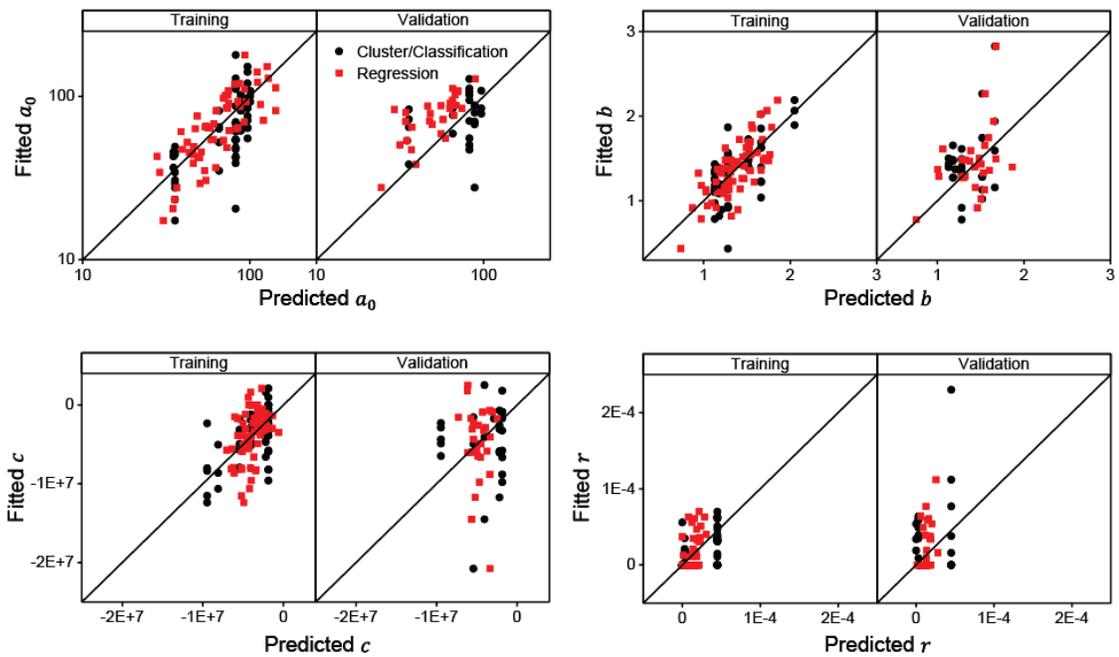


Figure C 7. Comparison of predicted parameter values and fitted parameter values for Brandywine Creek. Units are:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)) and  $r$  (km<sup>2</sup>/m<sup>3</sup>).

### S.B. Potomac River

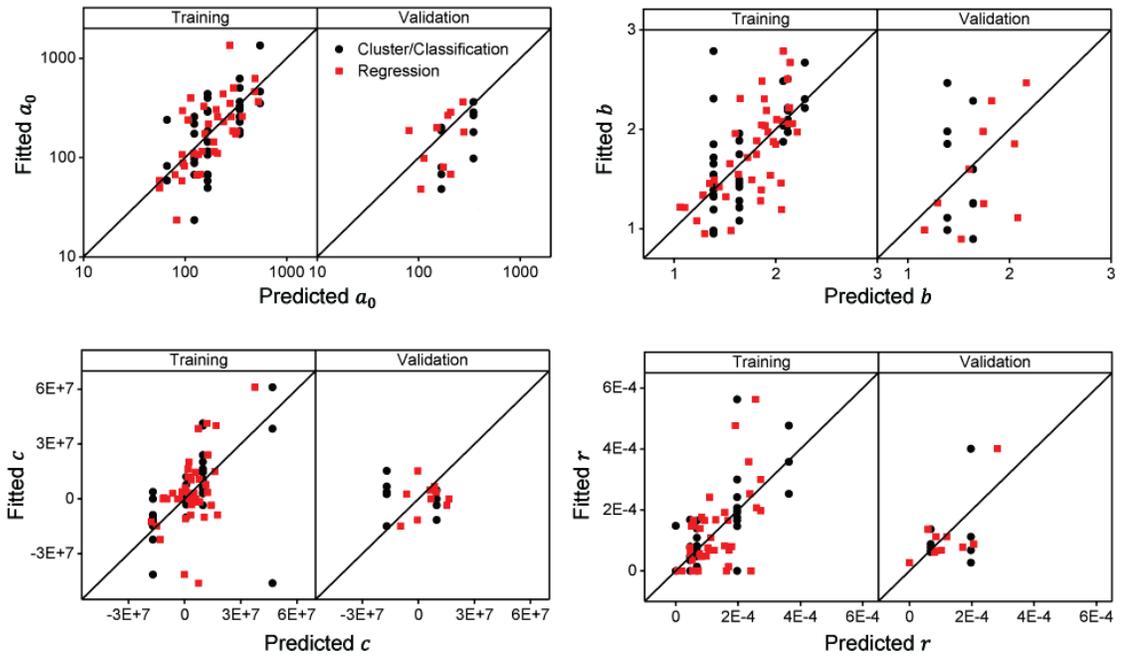


Figure C 8. Comparison of predicted parameter values and fitted parameter values for S.B. Potomac River. Units are:  $a_0$  (FNU),  $b$  (-),  $c$  (FNU(km<sup>2</sup>s)(s/m<sup>3</sup>)) and  $r$  (km<sup>2</sup>/m<sup>3</sup>).

# APPENDIX D

---

Supporting Information for Chapter 5  
**Forecasting Turbidity during Streamflow Events for  
Two Mid-Atlantic U.S. Streams**

*Amanda L. Mather and Richard L. Johnson*

## Section D.1: Flow Diagram of Calibration and Operational Forecasting

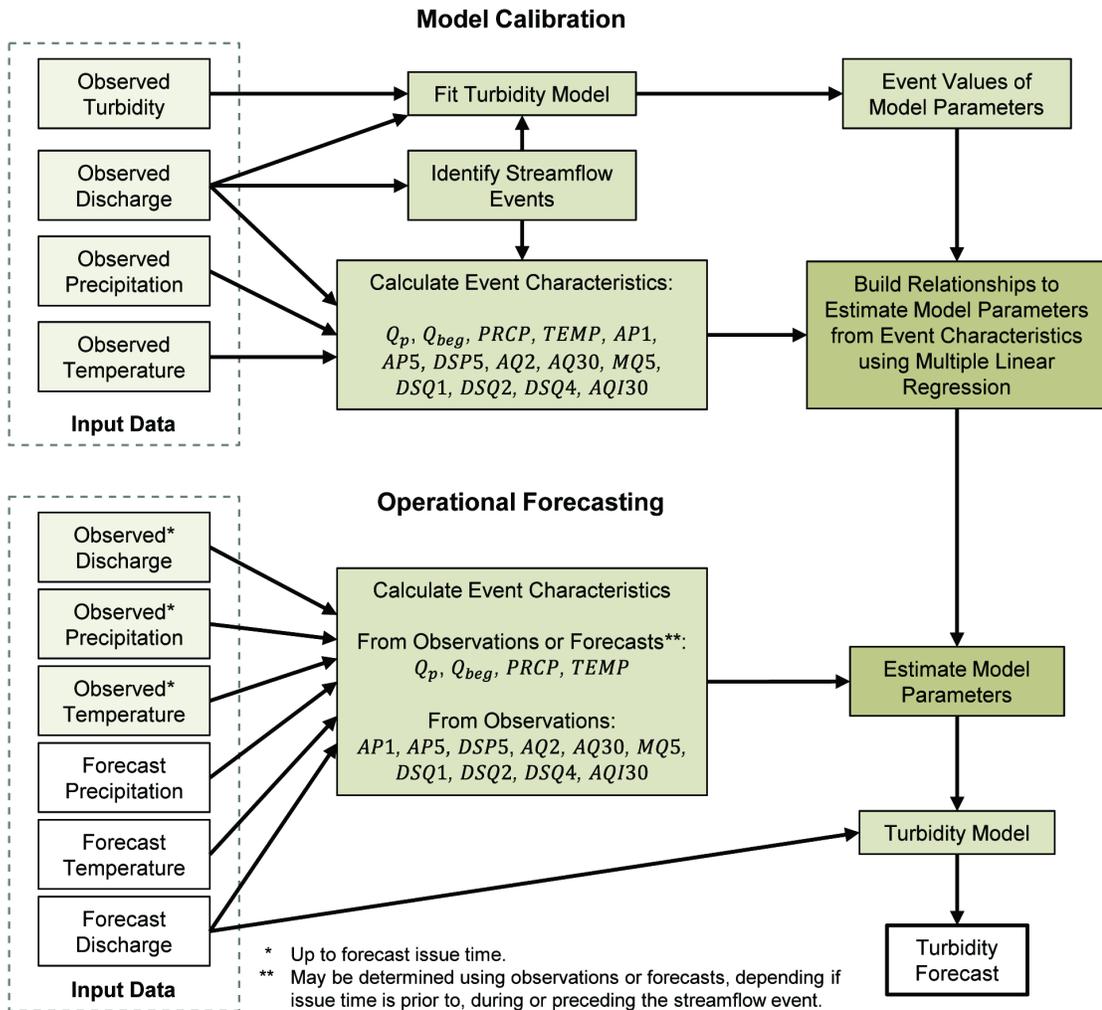


Figure D 1. Flow diagram showing turbidity model calibration and operational forecasting.

## Section D.2: $V(t)$ – Streamflow Event Cumulative Discharge

$$V(t) = \int_0^T Q(t)dt \cong \sum_{i=0}^T Q_i \Delta t$$

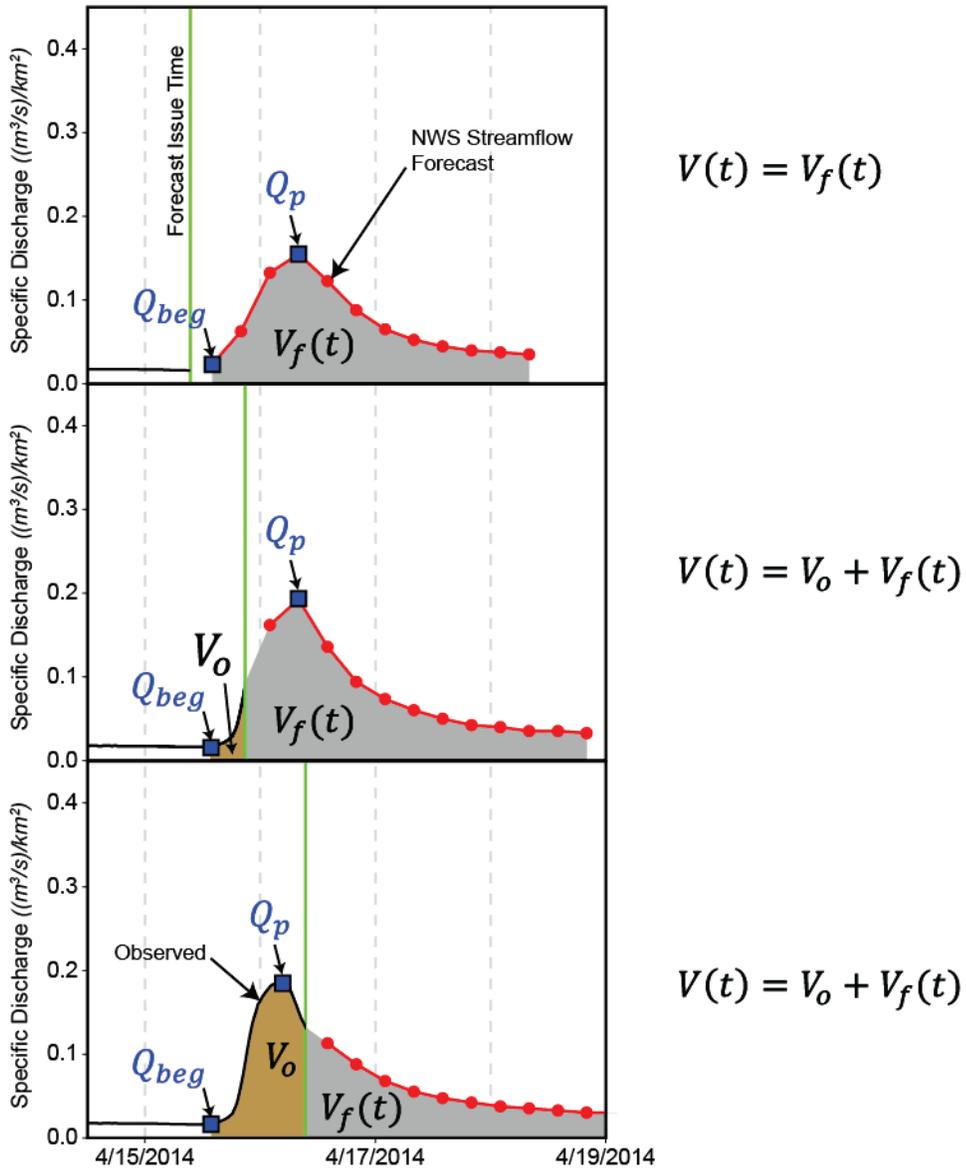


Figure D 2. Diagram showing calculation of  $V(t)$ , which is used in the turbidity model (equation (5.1) in Chapter 5).

## Section D.3: Uncertainty Intervals

### Definitions

“*forecast point*” – a single time in the future (e.g., 3/20/15 2:00 pm) with an associated forecast (e.g., a forecast discharge of  $0.10 \text{ (m}^3\text{/s)/km}^2$ ).

“*streamflow forecast*” or “*turbidity forecast*” – the typically 12 forecast points (3 days x 4/day (6-hr)) issued at a particular time (e.g., 3/20/15 9:37am).

### Pre-processing

1. The observed discharge corresponding to each discharge forecast point was determined.
2. Pairs of forecast and observed discharge data were binned into logarithmic intervals of  $\sim 1.6X$  (or 2 dB) based on forecast discharge. A given observed discharge data point may be part of several forecast/observation data pairs due to the regularity of forecast point times (i.e., 0:00, 6:00, 12:00 and 18:00 UTC) and the frequency of forecast issue (at least one forecast issued per day).
3. For each data pair, relative forecast discharge error (observed divided by forecast “*obs/fcst*”) was calculated.
4. For each forecast discharge bin, data pairs were arranged in ascending order based on values of *obs/fcst*.
5. For each forecast discharge bin, the array of ascending *obs/fcst* points were evenly spaced and linearly interpolated such that the array now has 1001

numbers. The first number is the lowest value of  $obs/fcst$ , while the last number is the highest value of  $obs/fcst$ .

### **Simulation**

6. For each originally issued streamflow forecast, a “new” streamflow forecast was made by performing the following steps for each point in the forecast:
  - a. Generate a random number from 0-1, multiply the number by 1000, and then round to the nearest whole number (resulting in 1001 whole number between 0-1000).
  - b. Determine which of the forecast discharge bins to use (from “step 2”) based on which interval the original forecast discharge for the data point falls into.
  - c. Use the number generated in “step 6a” to select (“look-up”) a value of  $obs/fcst$  from the array created in “step 5”.
  - d. Multiply the selected value of  $obs/fcst$  by the original forecast discharge for the data point to obtain a “new” forecast for the point.
7. Repeat “step 6” 10,000 times for each originally issued streamflow forecast.
8. Make turbidity forecasts for each “new” streamflow forecast. (i.e., 10,000 turbidity forecasts will be made corresponding to the 10,000 streamflow forecasts associated with each single original streamflow forecast).

### **Post-processing**

9. Each original forecast point (both discharge and turbidity) now has 10,000 “new” forecast values associated with it. Repeat the following for each point:
  - a. Arrange the 10,000 values in ascending order.

- b. Select the 500<sup>th</sup>, 1500<sup>th</sup>, 8500<sup>th</sup> and 9500<sup>th</sup> numbers from the array.
- c. The interval between the 500<sup>th</sup> and the 9500<sup>th</sup> numbers is the 90% uncertainty interval, between the 1500<sup>th</sup> and the 8500<sup>th</sup> is the 70% uncertainty interval.

## Section D.4: Streamflow Forecasts for Each Streamflow Event

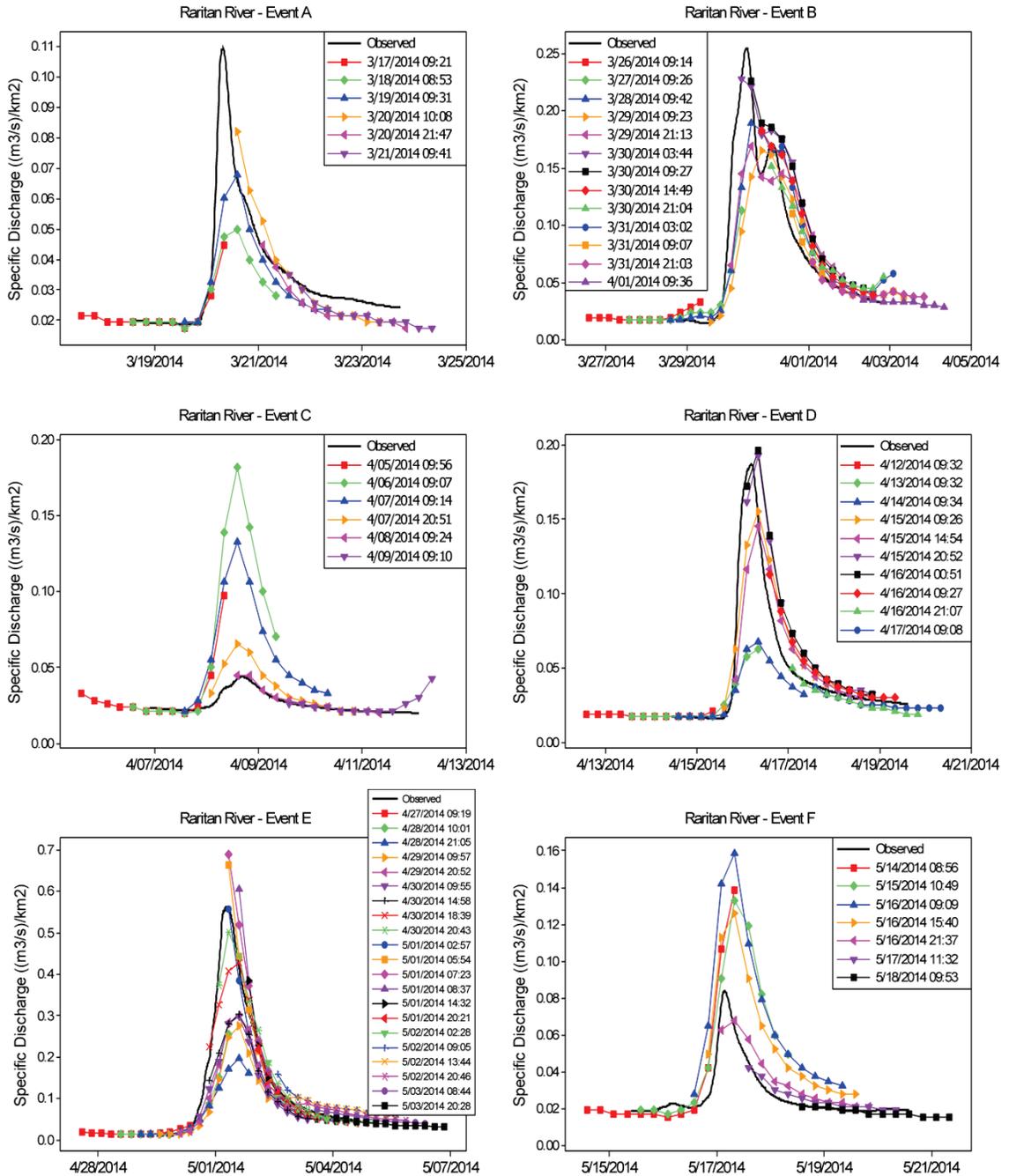


Figure D 3. NWS streamflow forecasts for Raritan River events A-F.

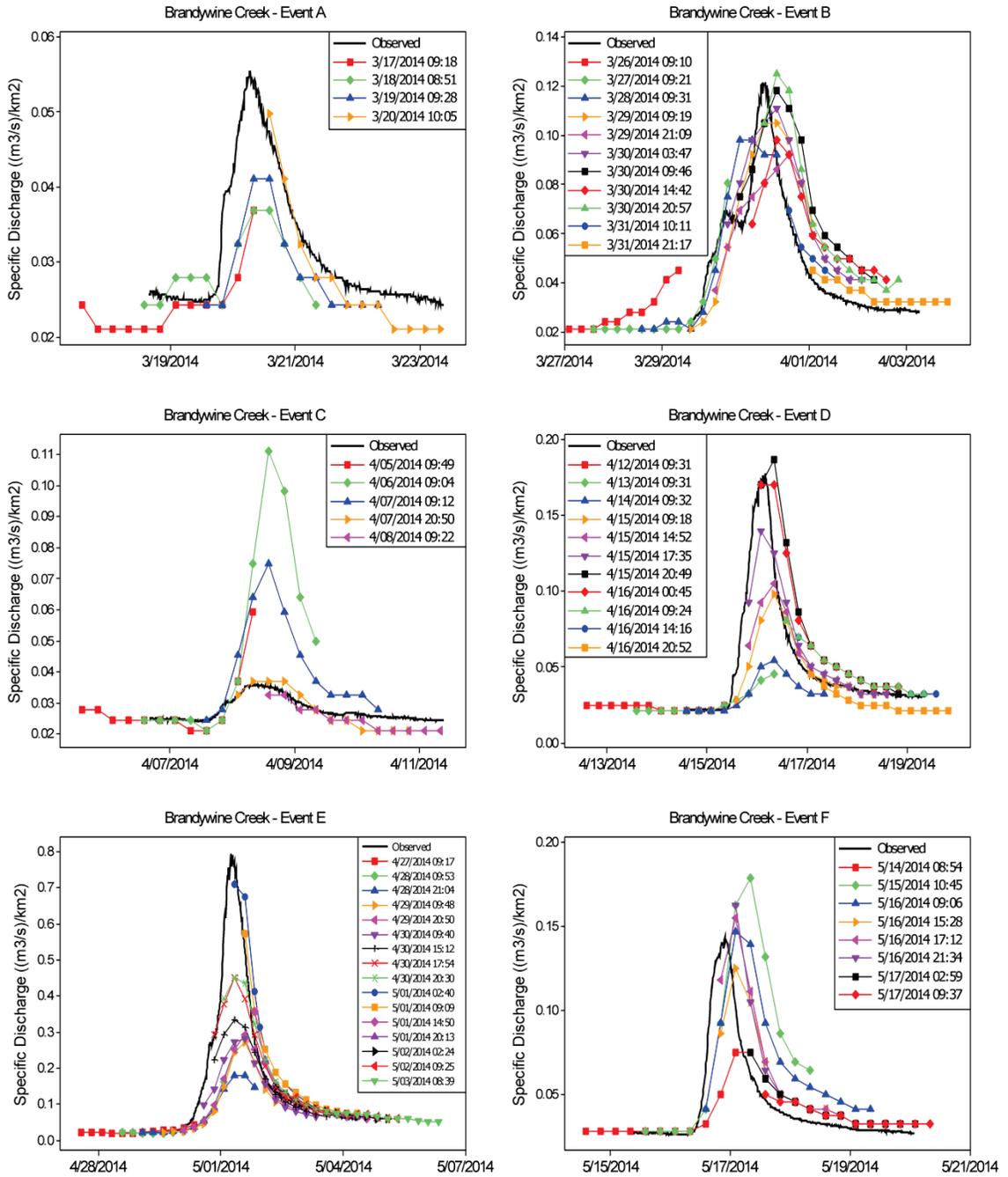


Figure D 4. NWS streamflow forecasts for Brandywine Creek events A-F.

## Section D.5: Turbidity Forecasts for Each Streamflow Event

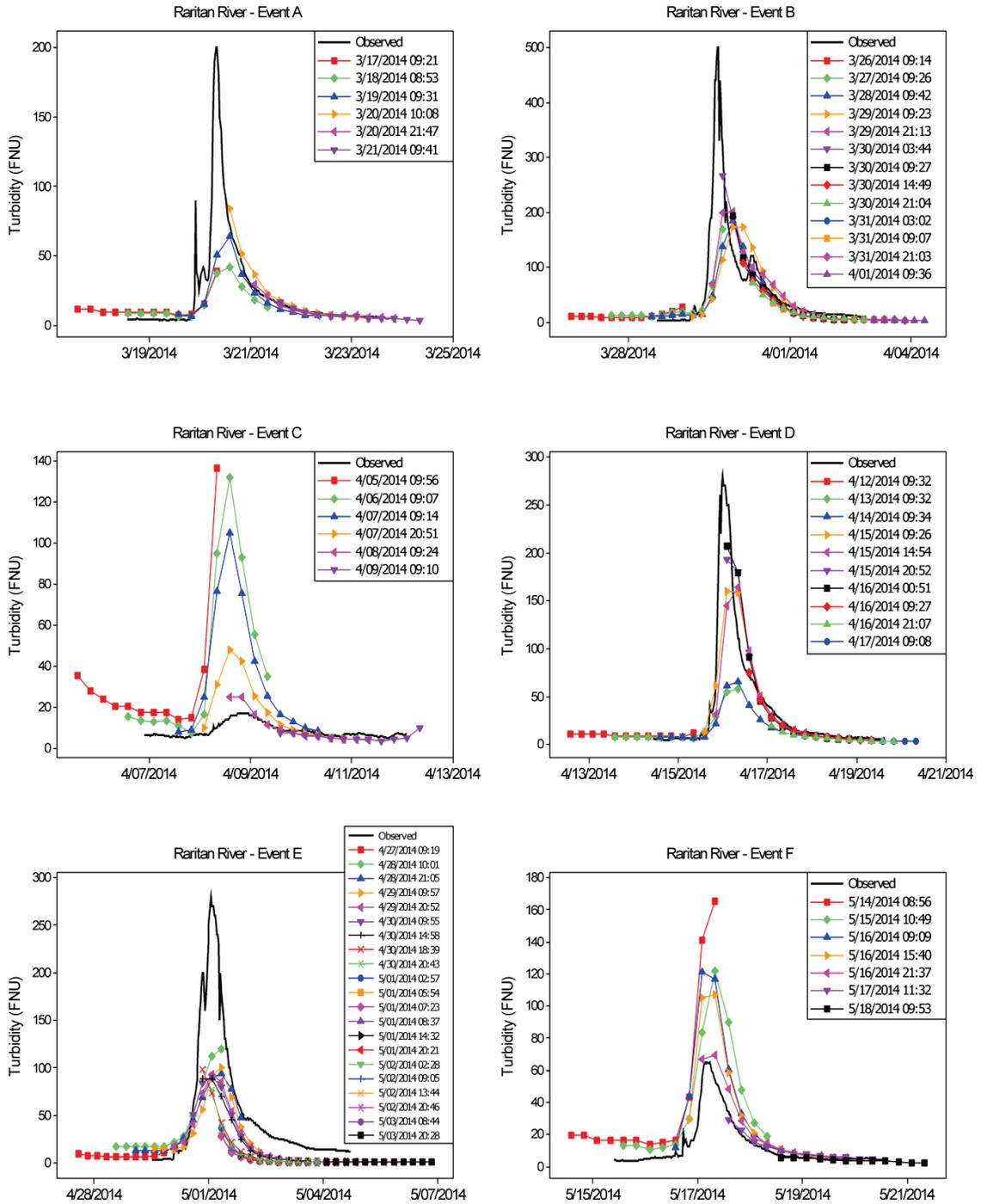


Figure D 5. Turbidity forecasts for Raritan River events A-F.

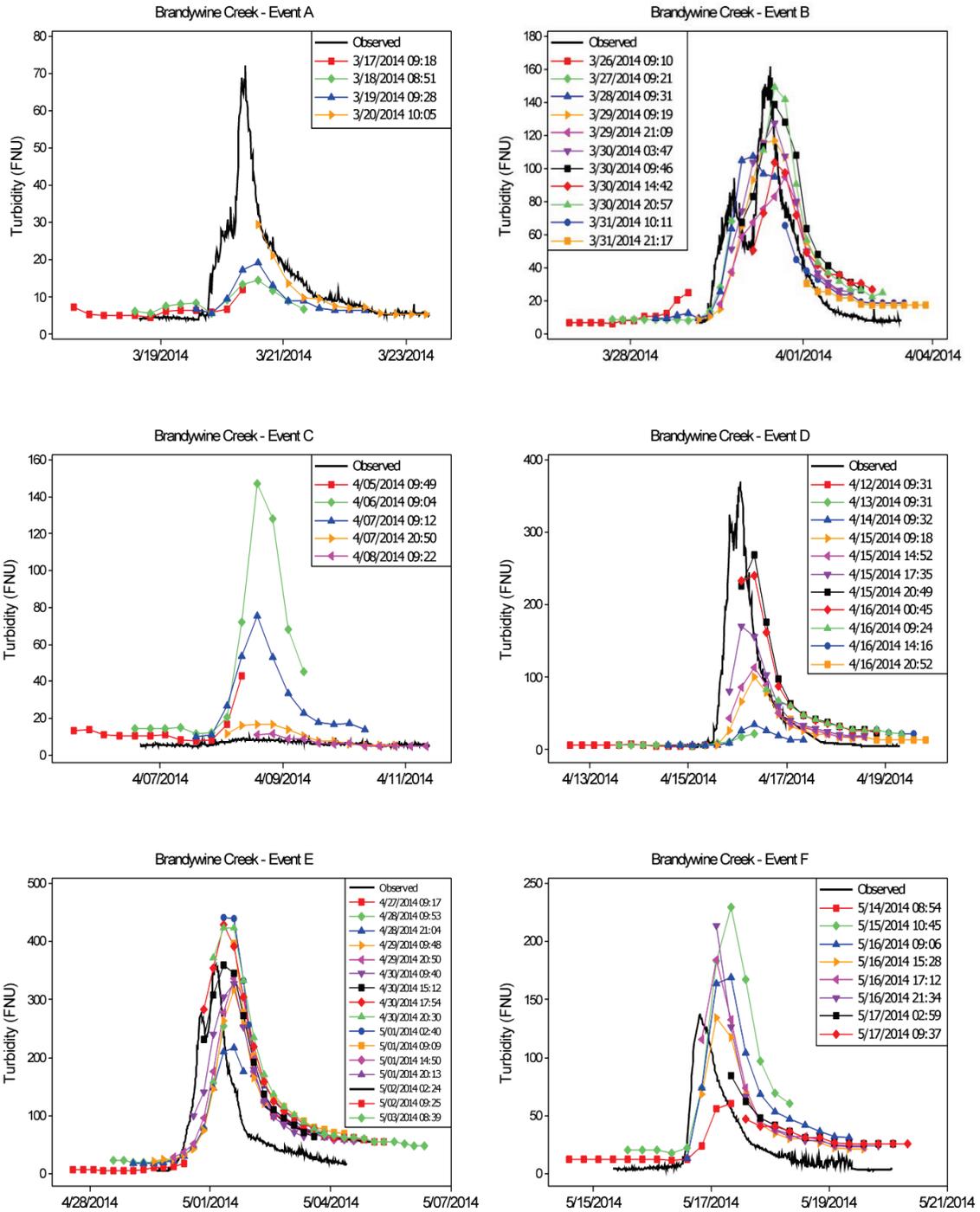


Figure D 6. Turbidity forecasts for Brandywine Creek events A-F.