

**QUANTIFYING NUTRIENT AND SEDIMENT EXPORT FROM THE  
CHESAPEAKE BAY WATERSHED:  
RETROSPECTIVE ANALYSES AND METHOD IMPROVEMENTS**

by

Qian Zhang

A dissertation submitted to Johns Hopkins University in conformity with the  
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

July 2016

© 2016 Qian Zhang

All Rights Reserved

## ABSTRACT

Toward Chesapeake Bay restoration, management programs have focused for decades on reducing nutrient and sediment loadings from the Chesapeake Bay watershed (CBW). To assess progress and shape future strategies, a critical need is to better understand historical loading changes from different regions of the CBW, using best current methods and data. In this regard, investigators at USGS have developed the “Weighted Regressions on Time, Discharge, and Season (WRTDS)” method for better loading and trend estimation.

Motivated by the above need, this dissertation research focused on applying the WRTDS method to long-term monitoring records for various major tributaries, exploring the method’s uncertainties, and improving the method’s estimation performance. Specific contributions include:

- (1) analysis of long-term seasonal trends of riverine nutrient and sediment loadings from major tributaries to Chesapeake Bay;
- (2) evaluation of decadal-scale changes in sediment and nutrient processing within the Lower Susquehanna River Reservoir System (LSRRS);
- (3) analysis of uncertainty and sensitivity of results to sample availability and storm events when estimating LSRRS input and output loadings;
- (4) investigation of the temporal and spatial patterns of nutrient and sediment export from the Susquehanna River basin and major factors affecting such patterns;
- (5) development of an improved method for making robust interpretations of riverine concentration-discharge relationships and application of this method toward Chesapeake tributaries for a top-down synthesis of export patterns;

- (6) development of improved methods for estimating riverine concentration and loading through incorporation of antecedent discharge conditions into WRTDS and evaluation of the methods' performance under various sampling strategies; and,
- (7) comparison of alternative approaches for quantifying long-range dependence in irregularly sampled water-quality data through Monte-Carlo simulations.

Overall, this research has demonstrated the utility of statistical modeling approaches toward large-scale analysis and synthesis of decadal-scale water-quality data collected in river systems. The applications to the CBW have provided new evidence on the decreasing trapping performance of the LSRRS and new understanding of nutrient and sediment export from various locations in the watershed. In addition, this work has made important methodological advancements with respect to WRTDS estimation performance, interpretation of riverine concentration-discharge relationships, and quantification of long-range dependence in irregularly sampled data.

**Dissertation Readers:**

Dr. William P. Ball, Johns Hopkins University (advisor)

Dr. Ciaran J. Harman, Johns Hopkins University

Dr. Robert M. Hirsch, U.S. Geological Survey

Dr. Peter R. Wilcock, Utah State University

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor Bill Ball for his continuous support of my Ph.D study and research. His guidance has helped me throughout my Ph.D, including but not limited to, research design, results synthesis and presentation, and edits of various manuscripts and this thesis. Bill has often reminded me to think about the big picture and to put my research in perspective. I could not have imagined having a better advisor and mentor for my doctoral research. I always feel very grateful to Bill, ever since his recruiting email came to me six years ago, which has opened the door of an exciting and rewarding journey.

Besides my advisor, I would like to also thank the rest of my dissertation committee, Bob Hirsch, Ciaran Harman, and Peter Wilcock, all of whom have always been very helpful. I clearly remember that about five years ago, when I initially approached Bob with some questions on WRTDS, he promptly replied my email with detailed answers to all my questions and offered to share the software and manual. Ever since, I have benefited tremendously throughout my doctoral research from his continuous support and always insightful advice. Ciaran has constantly pushed me to think hard about how watershed functions and how to properly use mathematical and statistical tools. He has been a very reliable and usually available source of advice. I always think myself fortunate, because Ciaran joined our department when I was still at the early stage of my Ph.D, otherwise I might have missed the opportunity to gain some important perspectives on watershed processes. Peter has given me a lot of encouragements and excellent comments. I learned a lot from him on sediment transport in rivers and the important

perspective of using simple models to approach complex problems. In addition to my dissertation committee, several others have also helped shape and refine the overall direction of my research, including Grace Brush and Seth Guikema, who have served on my Department Qualification Exam, as well as Danial Naiman and Frank Curriero, who have served on my Graduate Board Oral Exam.

Within my home department of Geography and Environmental Engineering, Ed Boucher has been always caring about me and has frequently checked my progress. Rebecca Murphy has given me a lot of guidance on research, literature review, and R coding during our overlapping years. I cannot thank her enough. Dano Wilusz has always been a very responsive resource to chat and share research ideas and findings. I benefited a lot from his critical questions and his edits on my manuscripts. Moreover, I want to thank Hengchen Wei, Di Ha, and Nathan Bickford, who joined our research team for about 1.5-2 years. I enjoyed working with them on closely-related research projects.

My research has also benefited significantly from conversations with many researchers and collaborators outside of JHU. Damian Brady, Dom Di. Toro, Michael Kemp, and Walter Boynton have helped a lot when I came to this field as a rookie. Many USGS individuals, particularly Bob Hirsch, Joel Blomquist, Doug Moyer, Mike Langland, Jeffery Chanat, Jeni Keisman, Scott Ator, and Lori Sprague have shared ideas, monitoring data, GIS maps, R codes, model results, and/or comments on my research approaches and findings. Gary Shenk, Lewis Linker, Guido Yactayo, and Gopal Bhatt at the Chesapeake Bay Program Office have provided watershed source input data, answers to my various questions, and many excellent comments to my research. Kevin McGonigal of the Susquehanna River Basin Commission has provided updated nutrient

and sediment data at several long-term monitoring stations. Ken Staver (University of Maryland) has pushed me to explore further on the uncertainties of our findings. Jim Kirchner (UC Berkeley and ETH Zurich) has provided many insightful comments and suggestions on the LRD analysis. Last but not least, my special thanks go to Damian Brady, Walter Boynton, Doug Moyer, and Bob Hirsch, who have contributed considerable time and efforts to the various manuscripts that we published together.

At different stages, my research has been supported financially by the Water Environment Research Federation (U4R09), National Science Foundation (0854329), Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090), U.S. Geological Survey (G15AC00067), and Maryland Water Resources Research Center (2015MD329B). None of this research would have been possible without the monitoring and modeling data made available by the U.S. Geological Survey, Susquehanna River Basin Commission, Chesapeake Bay Program, and National Center for Water Quality Research at Heidelberg University.

Finally, none of this would have been possible without the support and love of my parents and parents-in-law, my brother and sister, and my wife. My parents and my brother were my first teachers and have always been extremely supportive throughout my endeavors. My son Lucas, who is about 20-month old as I submit my dissertation, has helped me to relax, enjoy life, and stay positive. My parents-in-law helped to take care of little Lucas when I was busy writing this thesis. Most importantly, my wife, Qingyuan Li (Claire), has been a never-ceasing source of support and encouragement and a listener to share frustrations and successes. The completion of this dissertation research is one of my most important journeys. I feel so lucky to have her with me throughout this journey.

## TABLE OF CONTENTS

<i>ABSTRACT</i> .....	<i>ii</i>
<i>ACKNOWLEDGEMENTS</i> .....	<i>iv</i>
<i>TABLE OF CONTENTS</i> .....	<i>vii</i>
<i>LIST OF TABLES</i> .....	<i>xiv</i>
<i>LIST OF FIGURES</i> .....	<i>xv</i>
<i>Chapter 1. Introduction</i> .....	<i>1</i>
1.1. Chesapeake Bay Hypoxia and Riverine Inputs .....	1
1.2. River Water-Quality Monitoring Program .....	2
1.3. Riverine Flux Estimation and Trend Analysis with the WRTDS Method .....	4
1.4. Source, Fate, and Transport of Nutrients and Sediment in Watersheds .....	8
1.5. Sediment and Nutrient Retention in Conowingo Reservoir .....	10
1.6. 1/f Scaling Signature in River Water-Quality Data and Implications .....	12
1.7. Objectives and Outline .....	14
1.8. Literature Cited .....	17
<i>Chapter 2. Long-term Seasonal Trends of Nitrogen, Phosphorus, and Suspended Sediment Load from the Non-tidal Susquehanna River Basin to Chesapeake Bay</i> .....	<i>27</i>
Abstract.....	27
2.1. Introduction.....	28
2.2. Methods.....	36
2.2.1. Study Sites .....	36
2.2.2. Statistical Methods .....	38
2.2.3. Data Compilation and Analyses .....	40
2.3. Results and Discussion.....	44
2.3.1. History of NO <sub>x</sub> and TN Load at the Conowingo Station (1945-2011).....	44
2.3.2. History of SS, P, and N Load from the Marietta and Conestoga Stations (1986-2011) .....	48

2.3.3. History of SS, P, and N Load at the Conowingo Station (1978-2011) .....	53
<b>2.4. Conclusions.....</b>	<b>65</b>
<b>2.5. Supporting Information .....</b>	<b>66</b>
<b>2.6. Acknowledgements .....</b>	<b>66</b>
<b>2.7. Literature Cited .....</b>	<b>67</b>
<b><i>Chapter 3. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation .....</i></b>	<b><i>75</i></b>
<b>Abstract.....</b>	<b>75</b>
<b>3.1. Introduction.....</b>	<b>76</b>
<b>3.2. Data and Methods.....</b>	<b>80</b>
3.2.1. Study Sites and Data.....	80
3.2.2. Analysis with Standard WRTDS Models .....	81
3.2.3. Analysis with Stationary WRTDS Models .....	84
<b>3.3. Results and Discussion.....</b>	<b>85</b>
3.3.1. Temporal Changes in Concentration-Discharge Relationships .....	85
3.3.2. Changes in Net Deposition: Analysis of Loads from Standard WRTDS Models.....	87
3.3.3. Changes in Net Deposition: Analysis of Loads from Stationary WRTDS Models .....	95
<b>3.4. Management Implications.....</b>	<b>102</b>
<b>3.5. Supporting Information .....</b>	<b>103</b>
<b>3.6. Acknowledgements .....</b>	<b>103</b>
<b>3.7. Literature Cited .....</b>	<b>103</b>
<b><i>Chapter 4. Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns.....</i></b>	<b><i>111</i></b>
<b>Abstract.....</b>	<b>111</b>
<b>4.1. Introduction.....</b>	<b>112</b>
<b>4.2. Methods.....</b>	<b>115</b>
4.2.1. Study Area and Data.....	115
4.2.2. Statistical Method for Loading Estimation: WRTDS .....	116

4.2.3. Trend and Mass-Balance Analyses.....	120
<b>4.3. Results .....</b>	<b>122</b>
4.3.1. Temporal Trends in Flow-normalized Riverine Loadings.....	122
4.3.2. Comparison of Changes in Riverine Yield and Source Input.....	124
4.3.3. Mass Balances of Sub-basins and Effects of Streamflow on Export.....	129
4.3.4. Relative Contributions by Sub-basins and Effects of Land Use on Export.....	133
<b>4.4. Discussion .....</b>	<b>138</b>
4.4.1. Temporal Trends in Flow-normalized Riverine Loadings.....	138
4.4.2. Comparison of Changes in Riverine Yield and Source Input.....	140
4.4.3. Mass Balances of Sub-basins and Effects of Streamflow on Export.....	144
4.4.4. Relative Contributions by Sub-basins and Effects of Land Use on Export.....	146
<b>4.5. Conclusions.....</b>	<b>147</b>
<b>4.6. Supporting Information .....</b>	<b>149</b>
<b>4.7. Acknowledgements .....</b>	<b>150</b>
<b>4.8. Literature Cited .....</b>	<b>150</b>
<b><i>Chapter 5. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season.....</i></b>	<b><i>163</i></b>
<b>Abstract.....</b>	<b>163</b>
<b>5.1. Introduction.....</b>	<b>164</b>
5.1.1. Chesapeake Bay Hypoxia and Watershed Inputs Control.....	164
5.1.2. Statistical Methods for Riverine Loading Estimation.....	165
5.1.3. Insights Gained from Prior Studies on Conowingo Reservoir.....	166
5.1.4. Aims and Contributions of This Work .....	167
<b>5.2. Data and Methods.....</b>	<b>168</b>
5.2.1. Study Area: the NTCBW.....	168
5.2.2. Statistical Methods: WRTDS .....	171
5.2.3. Data Compilation and Analyses .....	174
<b>5.3. Seasonal Trends in the NTCBW.....</b>	<b>178</b>
5.3.1. Difficulty in Evaluating Trends with True-condition Estimates.....	178
5.3.2. Flow-normalized Trends in the NTCBW .....	179
5.3.3. Dominance by Susquehanna River.....	182
<b>5.4. Comparison of Loads among Seasons.....</b>	<b>186</b>

5.4.1. Comparison of Load Magnitudes among Seasons .....	186
5.4.2. Comparison of Load Trends among seasons .....	190
<b>5.5. Comparison of Loads among Tributaries.....</b>	<b>191</b>
5.5.1. Particulate (SS, PP, PN) and Particulate-dominated (TP) Species .....	191
5.5.2. Dissolved (DN, DP) and Dissolved-dominated (TN) Species .....	195
5.5.3. Fractional Contributions of the Tributaries.....	197
5.5.4. Changes in N:P Ratios in the Tributaries.....	200
<b>5.6. Next Steps: Explaining Changes and Trends .....</b>	<b>204</b>
<b>5.7. Summary.....</b>	<b>204</b>
<b>5.8. Supporting Information .....</b>	<b>205</b>
<b>5.9. Acknowledgements .....</b>	<b>206</b>
<b>5.10. Literature Cited .....</b>	<b>206</b>
<b><i>Chapter 6. What Can We Learn from Limited Data? Statistical Inferences and Uncertainties of Riverine Fluxes and Trends with Limited Sampling of Extreme-Flow Events .....</i></b>	<b><i>219</i></b>
<b>Abstract.....</b>	<b>219</b>
<b>6.1. Introduction.....</b>	<b>220</b>
<b>6.2. Analyses &amp; Results.....</b>	<b>224</b>
6.2.1. WRTDS Mass-Balance Analysis for the Major Storm Events .....	224
6.2.2. WRTDS Weight Analysis.....	227
6.2.3. WRTDS Regression Surface Analysis.....	228
6.2.4. WRTDS Flow-Normalized Loading Analysis .....	230
<b>6.3. Summary, Implications, and Prospects.....</b>	<b>233</b>
<b>6.4. Supporting Information .....</b>	<b>234</b>
<b>6.5. Acknowledgements .....</b>	<b>234</b>
<b>6.6. Literature Cited .....</b>	<b>234</b>
<b><i>Chapter 7. An Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-Quality Data .....</i></b>	<b><i>239</i></b>
<b>Abstract.....</b>	<b>239</b>

<b>7.1. Introduction and Background .....</b>	<b>240</b>
7.1.1. Complexities with $C-Q$ Interpretation .....	241
7.1.2. Solution: The WRTDS Method .....	244
<b>7.2. Methods: WRTDS <math>\beta_2</math> Extraction and Use.....</b>	<b>247</b>
<b>7.3. Results and Discussion.....</b>	<b>249</b>
7.3.1. Visualization of WRTDS $\beta_2$ Coefficients .....	249
7.3.2. Trends of WRTDS $\beta_2$ Coefficients with Respect to Season and Discharge .....	251
7.3.3. Temporal Trends in WRTDS $\beta_2$ Coefficients with Uncertainty Analysis .....	255
7.3.4. Limitations of the Proposed Approach .....	257
<b>7.4. Summary.....</b>	<b>258</b>
<b>7.5. Supporting Information .....</b>	<b>259</b>
<b>7.6. Acknowledgements .....</b>	<b>259</b>
<b>7.7. Literature Cited .....</b>	<b>259</b>
<b><i>Chapter 8. Non-stationary Concentration-Discharge Relationships: A Synthesis of Nutrient and Sediment Patterns in the Major Tributaries to Chesapeake Bay.....</i></b>	<b>267</b>
<b>Abstract.....</b>	<b>267</b>
<b>8.1. Introduction.....</b>	<b>268</b>
<b>8.2. Methods.....</b>	<b>271</b>
8.2.1. Study Sites .....	271
8.2.2. Monitoring Data .....	275
8.2.3. WRTDS Method.....	277
8.2.4. Data Analysis.....	278
<b>8.3. Results.....</b>	<b>280</b>
8.3.1. Changes in Coefficients with Discharge.....	280
8.3.2. Temporal Changes in Coefficients for Selected Discharge Conditions .....	285
<b>8.4. Discussion .....</b>	<b>292</b>
8.4.1. Changes in Coefficients with Discharge.....	292
8.4.2. Temporal Changes in Coefficients for Selected Discharge Conditions .....	297
<b>8.5. Conclusions.....</b>	<b>300</b>
<b>8.6. Supporting Information .....</b>	<b>301</b>

8.7. Acknowledgements .....	302
8.8. Literature Cited .....	302
<i>Chapter 9. Improving Riverine Constituent Concentration and Flux Estimation by Accounting for Antecedent Discharge Conditions .....</i>	<i>313</i>
Abstract.....	313
9.1. Introduction.....	314
9.2. Statistical Models: Original and Modified WRTDS .....	318
9.3. Data and Analysis .....	323
9.3.1. Testing Sites and Data .....	323
9.3.2. Monte Carlo Sub-sampling.....	325
9.3.3. Model Evaluation.....	327
9.4. Results and Discussion .....	328
9.4.1. Selected Cases with Major Improvement in Performance.....	329
9.4.2. Performance Comparison for All Simulations: Effects of Model Choices.....	335
9.4.3. Performance Comparison for All Simulations: Effects of Sampling Strategies .....	341
9.5. Summary and Final Remarks.....	347
9.6. Supporting Information .....	349
9.7. Acknowledgements .....	349
9.8. Literature Cited .....	350
<i>Chapter 10. Evaluation of Methods for Estimating Long-Range Dependence in Water Quality Time Series with Irregular Sampling.....</i>	<i>361</i>
Abstract.....	361
10.1. Introduction.....	362
10.1.1. Autocorrelations in Time Series .....	362
10.1.2. Overview of Approaches for LRD Quantification.....	363
10.1.3. Motivation and Objective of this Work .....	367
10.2. Quantification of Sampling Irregularity in Typical River Water-Quality Data .....	370
10.2.1. Modeling of Sampling Irregularity .....	370
10.2.2. Examination of Sampling Irregularity in Real River Water-Quality Data.....	373
10.2.3. Simulation of Time Series with Irregular Sampling .....	376

<b>10.3. Evaluation of Various LRD Estimation Methods for Irregular Time Series .....</b>	<b>378</b>
10.3.1. Summary of Estimation Methods .....	378
10.3.2. Evaluation of Method Performance .....	381
10.3.3. Quantification of LRD in Real Water-Quality Data .....	390
<b>10.4. Summary and Prospects .....</b>	<b>394</b>
<b>10.5. Supporting Information .....</b>	<b>395</b>
<b>10.6. Acknowledgements .....</b>	<b>395</b>
<b>10.7. Literature Cited .....</b>	<b>396</b>
<b><i>Chapter 11. Summary, Application, and Future Work .....</i></b>	<b><i>403</i></b>
11.1. Summary and Implications .....	403
11.2. Application and Relevancy .....	408
11.3. Recommendations for Future Work .....	412
11.4. Literature Cited .....	417
<b><i>Appendix A. Supporting Information to Chapter 2 .....</i></b>	<b><i>421</i></b>
<b><i>Appendix B. Supporting Information to Chapter 3 .....</i></b>	<b><i>435</i></b>
<b><i>Appendix C. Supporting Information to Chapter 4 .....</i></b>	<b><i>487</i></b>
<b><i>Appendix D. Supporting Information to Chapter 5 .....</i></b>	<b><i>499</i></b>
<b><i>Appendix E. Supporting Information to Chapter 6 .....</i></b>	<b><i>535</i></b>
<b><i>Appendix F. Supporting Information to Chapter 8 .....</i></b>	<b><i>549</i></b>
<b><i>Appendix G. Supporting Information to Chapter 9 .....</i></b>	<b><i>553</i></b>
<b><i>Appendix H. Supporting Information to Chapter 10 .....</i></b>	<b><i>559</i></b>
<b><i>CURRICULUM VITAE .....</i></b>	<b><i>575</i></b>

## LIST OF TABLES

### Chapter 2

Table 2.1. Details of the study sites. <sup>a</sup> .....	37
---	----

### Chapter 4

Table 4.1. Details of the long-term monitoring sites and sub-basins in the Susquehanna River basin. <sup>a</sup> .....	118
--	-----

### Chapter 5

Table 5.1. Details of the monitoring sites. ....	170
Table 5.2. Summary of long-term median values of annual loads and yields in the nontidal Chesapeake Bay watershed (NTCBW) and nine tributaries. <sup>a</sup> .....	177

### Chapter 8

Table 8.1. Details of the 15 long-term monitoring sites in the Chesapeake Bay watershed. <sup>a</sup> .....	274
Table 8.2. Temporal coverage of observed water-quality data at the 15 Chesapeake sites. ....	276
Table 8.3. Period-of-record changes ( $\Delta$ ) in estimated WRTDS $\beta_2$ coefficients at the 15 Chesapeake sites under three different discharge conditions. ....	290

### Chapter 9

Table 9.1. Details of the nine river monitoring sites.....	324
Table 9.2. Temporal coverages of water-quality data at the nine study sites. ....	325
Table 9.3. Summary of cases with major improvements in estimation performance. Each case shows results that are based on one selected model run. ....	332

### Chapter 10

Table 10.1. Quantification of sampling irregularity for selected water-quality constituents at nine sites of the Chesapeake Bay River Input Monitoring program and six sites of the Lake Erie and Ohio tributary monitoring program.....	375
--	-----

## LIST OF FIGURES

### Chapter 1

- Figure 1.1. The Chesapeake Bay watershed and monitoring sites at the fall-line of the nine major tributaries – Susquehanna River (1), Potomac River (2), James River (3), Rappahannock River (4), Appomattox River (5), Pamunkey River (6), Mattaponi River (7), Patuxent River (8), and Choptank River (9). Inset shows the major physiographic provinces. This figure was reproduced from Moyer *et al.* (2012) with permission. .... 3
- Figure 1.2. Source input, speciation, major hydrological pathway, and output signals of nitrogen, phosphorus, and suspended sediment. .... 8
- Figure 1.3. Map of the Lower Susquehanna River Reservoir System consisting of Lake Clarke, Lake Aldred, and Conowingo Reservoir. Yellow triangles indicate the three monitoring sites: Conowingo, Marietta, and Conestoga. .... 11

### Chapter 2

- Figure 2.1. Map of the Susquehanna River basin and long-term monitoring sites. Sites in this study include two main-stem stations, Marietta (No. 7) and Conowingo (No. 9), and one tributary station, Conestoga (No. 8). The non-tidal Susquehanna River basin, shaded in yellow, covers portions of New York, Pennsylvania, and Maryland. This figure was reproduced from Figure 6 in Sprague *et al.* (2000) with permission, with the original numbering of the monitoring sites. .... 30
- Figure 2.2. Map of Lower Susquehanna River Reservoir System and the study sites. The reservoir system consists of Lake Clarke, Lake Aldred, and Conowingo Reservoir. The Marietta station (No. 7) is just above the reservoirs and the Conowingo station (No. 9) is just below the reservoirs. The Conestoga station (No. 8) monitors streamflow from the Conestoga River, a major tributary to the Susquehanna River. This figure was adapted from Figure 1 in Langland (2009) with permission. See Figure 2.1 for locations of the three sites in the non-tidal Susquehanna River basin. .... 33
- Figure 2.3. Observed data of NO<sub>x</sub> concentration and seasonal streamflow discharge in (a) Jan-Mar, (b) Apr-Jun, (c) Jul-Sep, and (d) Oct-Dec, in the Susquehanna River at the Conowingo station for the period from 1945 to 2011. .... 43

Figure 2.4. Seasonal averages of NO <sub>x</sub> (a) true-condition concentration, (b) flow-normalized concentration, (c) true-condition load, and (d) flow-normalized load in the Susquehanna River at the Conowingo station. All estimates have been normalized by the median of respective long-term annual averages at the Conowingo station (locating at y =1.0 in each panel). .....	45
Figure 2.5. Estimates of “true-condition” Jan-May TN load in the Susquehanna River at the Conowingo station for the period 1945 to 2011. Plot (a) shows TN loads obtained in this study. Plot (b) shows TN loads reported by Murphy <i>et al.</i> (2011) for the period 1949 to 2009. For that study, TN loads prior to 1980 were obtained using regression equations between TN and NO <sub>x</sub> load developed by Hagy <i>et al.</i> (2004), and TN loads from 1981 to 2009 were directly obtained from the USGS RIM Program website (US Geological Survey, 2012a). Plot (c) shows differences in estimates between (a) and (b). .....	46
Figure 2.6. Seasonal averages of flow-normalized SS load from the Marietta and Conestoga stations. All loads have been normalized by the median of long-term annual SS loads at the Conowingo station (locating at y =1.0). .....	48
Figure 2.7. Seasonal averages of flow-normalized load of (a) TP, (b) PP, (c) DP, (d) DOP (dissolved orthophosphate), and (e) DHP (dissolved hydrolysable P) from the Marietta and Conestoga stations. All loads have been normalized by the median of respective long-term annual loads at the Conowingo station at the reservoir outlet (locating at y =1.0 in each panel). .....	50
Figure 2.8. Seasonal averages of flow-normalized load of (a) TN, (b) PN, (c) DN, (d) DNO <sub>x</sub> , and (e) DKN from the Marietta and Conestoga stations. All loads have been normalized by the median of respective long-term annual loads at the Conowingo station at the reservoir outlet (locating at y =1.0 in each panel). .....	51
Figure 2.9. Seasonal averages of flow-normalized SS load in the Susquehanna River at the Conowingo station. All loads have been normalized by the median of long-term annual SS loads at Conowingo (locating at y =1.0). The TMDL of 2,510,000 kg/day set for Susquehanna River (US Environmental Protection Agency, 2010) is inserted for comparison.....	54
Figure 2.10. Seasonal averages of flow-normalized load of (a) TP, (b) PP, (c) DP, (d) DOP (dissolved orthophosphate), and (e) DHP (dissolved hydrolysable P) in the Susquehanna River at the Conowingo station. All loads have been normalized by the median of respective long-term annual loads at Conowingo (locating at y =1.0 in each	

panel). The TMDL of 3,700 kg P/day set for Susquehanna River (US Environmental Protection Agency, 2010) is inserted in (a) for comparison. ....	55
Figure 2.11. Seasonal averages of flow-normalized load of (a) TN, (b) PN, (c) DN, (d) DNO <sub>x</sub> , and (e) DKN in the Susquehanna River at the Conowingo station. All loads have been normalized by the median of respective long-term annual loads at Conowingo (locating at $y = 1.0$ in each panel). The TMDL of 101,000 kg N/day set for Susquehanna River (US Environmental Protection Agency, 2010) is inserted in (a) for comparison. ....	56
Figure 2.12. Rates of storage change in (a) SS, (b) PP, and (c) PN within Conowingo Reservoir based on flow-normalized load. All rates of change have been calculated as the differences between the loads at Conowingo (system output) and the combined loads from Marietta and Conestoga (system input), and then normalized by the median of respective long-term annual loads at the Conowingo station. ....	60
Figure 2.13. The evolving patterns of SS concentration vs. river discharge in each season at (a-d) the Marietta and (e-h) the Conowingo stations. A 5-year interval was selected to show the evolution. Patterns in 2002 (dashed line) were also added in (e) to (h) to aid comparison. ....	63

Chapter 3

Figure 3.1. Map of the Lower Susquehanna River Reservoir System consisting of Lake Clarke, Lake Aldred, and Conowingo Reservoir. Yellow triangles indicate the three monitoring sites: Conowingo, Marietta, and Conestoga. (See Table B1 for site details.) This figure was modified after Figure 1 in Langland (2015) with simplifications. ....	77
Figure 3.2. Observed concentration-discharge ( $C-Q$ ) relationships for suspended sediment (SS) and total phosphorus (TP) in the Susquehanna River at Conowingo, MD, for three separate periods between 1987 and 2013. Solid lines are fitted LOWESS (locally weighted scatterplot smoothing) curves. Vertical black dashed lines in (a)-(b) correspond to 3000 m <sup>3</sup> /s. Vertical purple dashed lines in (a)-(b) indicate the literature scour threshold of 11300 m <sup>3</sup> /s (400000 ft <sup>3</sup> /s). Data points with vertical solid lines in (b) indicate left-censored concentration samples. (Detection limit varied with samples.).....	86
Figure 3.3. Estimated output/input ratios (O/I) based on 35-day moving averages of WRTDS-estimated output and input loadings. Plots (a)-(c) are annual boxplots for (a)	

suspended sediment (SS), (b) total phosphorus (TP), and (c) total nitrogen (TN). (Note that each boxplot represents 365 daily data points and that  $O/I < 1.0$  reflects net deposition.) Plots (d)-(f) show the results of uncertainty analysis based on 100 synthetic data sets – see text. Plots show the averages of annual median values (blue dots) and the 95% confidence intervals (black error bars). ..... 90

Figure 3.4. Fractional contribution of total streamflow volume ( $V_w$ ) and nutrient and sediment mass discharges in the Susquehanna River at Conowingo by each of the five flow classes for reservoir input (a) and output (b) in 1987-2013. (Ranges of the five flow classes –  $Q_1$ : 25~396  $m^3/s$ ;  $Q_2$ : 399~787  $m^3/s$ ;  $Q_3$ : 790~1,464  $m^3/s$ ;  $Q_4$ : 1,467~7,646  $m^3/s$ ;  $Q_5$ : 7,674~20,077  $m^3/s$ .) ..... 94

Figure 3.5. Frequency plots of ranked loadings for reservoir output, input, and net increase in storage for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) based on the three historical stationary WRTDS models. For output (panels a, d, g) and input (panels b, e, h), dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. For net increase in storage (panels c, f, i), dashed lines indicate the 95% confidence intervals based on 100 sets of net storage estimates obtained from the resulting 10x10 input x output matrix. See Figures B23–B31 in Appendix B5-I for the full range of x-axis (*i.e.*, 0–365 days/year). ..... 97

Figure 3.6. Modeled cumulative reservoir storage over the course of three selected wet, average, and dry calendar years (*i.e.*, 2003, 2007, and 2001) with respect to suspended sediment (SS) loads based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. See Figure B41 in Appendix B5-III for results on SS output and input loadings. .... 101

#### Chapter 4

Figure 4.1. Map of the Susquehanna River basin (SRB), showing the seven long-term monitoring sites (No. 1-7) and the seven sub-basins (SB1-SB7). Conowingo (#7) is the non-tidal SRB’s outlet. Inset shows the SRB’s location within the Chesapeake Bay watershed. The diagram of “Simplified River Network” shows four sites on the river mainstem

	and three sites on the tributaries to Susquehanna River. See Table 4.1 for details of the sites and sub-basins. ....	117
Figure 4.2.	Reconstructed time series of (a) annual discharge and annual <i>flow-normalized (FN)</i> loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN at the seven Susquehanna sites. To aid comparison, all <i>y-axis</i> values have been scaled by respective long-term annual medians (see Table C2).....	123
Figure 4.3.	Reconstructed time series of WRTDS-estimated TP riverine yield (flow-normalized and true-condition estimates) and yields from major source inputs (fertilizer, manure, point source, and sum of all sources) for the seven Susquehanna sites: (a) Towanda, (b) Danville, (c) Lewisburg, (d) Newport, (e) Marietta, (f) Conestoga, and (g) Conowingo. Note that the <i>y-axis</i> scale varies with plot. ....	126
Figure 4.4.	Reconstructed time series of WRTDS-estimated TN riverine yield (flow-normalized and true-condition estimates) and yields from major source inputs (atmospheric deposition, fertilizer, manure, point source, and sum of all sources) for the seven Susquehanna sites: (a) Towanda, (b) Danville, (c) Lewisburg, (d) Newport, (e) Marietta, (f) Conestoga, and (g) Conowingo. Note that the <i>y-axis</i> scale varies with plot.....	127
Figure 4.5.	Comparison between declines in total source input and declines in flow-normalized riverine yield for the seven Susquehanna sites for the period of 1987-2011. See Table C3 for details. The blue diagonal line indicates the 1:1 reference line. (Note that the Conestoga site is clearly from a different population than the other sites. These non-Conestoga sites show no relation for TP and a low-slope relation for TN.) .....	129
Figure 4.6.	Reconstructed time series of (a) annual discharge and annual <i>true-condition</i> net contributions of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN by the seven Susquehanna sub-basins. To aid comparison, <i>all y-axis</i> values have been scaled by respective long-term annual maxima.....	130
Figure 4.7.	Fitted linear slopes for relations between annual flow-weighted concentration ( $C_{\text{Annual-FW}}$ ) and area-normalized annual discharge ( $[Q/A]_{\text{Annual}}$ ) on log-log scale at the seven Susquehanna sites for each water-quality constituent. Detailed data and slope fits are presented in Figure C2. ....	132
Figure 4.8.	Relations between area-normalized annual discharge ( $[Q/A]_{\text{Annual}}$ ) and annual ratios of (a) DP/TP, (b) DN/TN, (c) PP/SS, and (d) TN/TP at the seven Susquehanna sites. Plot (e) shows relations between area-normalized daily discharge ( $[Q/A]_{\text{Daily}}$ ) and daily TN/TP ratio for each site. Plot (f) summarizes the daily TN/TP ratio at each site with	

boxplots. The region between dashed lines in plots (d)-(f) represent the nominally co-limitation condition by both N and P ( <i>i.e.</i> , $10 \leq \text{TN/TP} \leq 20$ ).....	134
Figure 4.9. Fractional contributions (FC) of each sub-basin to (a) annual discharge and annual <i>true-condition</i> loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN at Conowingo (river fall-line). Note that FCs of all sub-basins sum up to one in each year. ....	135
Figure 4.10. Relations between area fractions of two types of land use ( <i>i.e.</i> , forested and non-forested) and log-transformed period-of-record medians of (a) area-normalized annual discharge ( $[\text{Q/A}]_{\text{Annual Median}}$ ) and annual <i>true-condition</i> loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN in the seven sub-basins. Each point represents one sub-basin. Dashed lines are linear fits. ....	137

## Chapter 5

Figure 5.1. Chesapeake Bay watershed and monitoring sites at the fall-line of the nine major tributaries — Susquehanna River (1), Potomac River (2), James River (3), Rappahannock River (4), Appomattox River (5), Pamunkey River (6), Mattaponi River (7), Patuxent River (8), and Choptank River (9). Inset shows the major physiographic provinces. This figure was reproduced from Moyer <i>et al.</i> (2012) with permission. ....	169
Figure 5.2. True-condition seasonal estimates of (a) total phosphorus (TP) and (b) total nitrogen (TN) loads from the nontidal Chesapeake Bay watershed (NTCBW). Y-axis values are loadings scaled by the respective long-term medians.....	179
Figure 5.3. Flow-normalized seasonal estimates of (a) total phosphorus (TP), (b) particulate phosphorus (PP), (c) dissolved phosphorus (PP), (d) suspended sediment (SS), (e) total nitrogen (TN), (f) particulate nitrogen (PN), and (g) dissolved nitrogen (DN) loads from the nontidal Chesapeake Bay watershed (NTCBW). Y-axis values are loadings scaled by the respective long-term medians (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).....	180
Figure 5.4. Flow-normalized seasonal estimates of (a) total phosphorus (TP), (b) particulate phosphorus (PP), (c) dissolved phosphorus (PP), (d) suspended sediment (SS), (e) total nitrogen (TN), (f) particulate nitrogen (PN), and (g) dissolved nitrogen (DN) loads from all nine rivers but the Susquehanna ( <i>i.e.</i> , “nontidal Chesapeake Bay watershed – Susquehanna” [“NTCBW-SUS”]). Y-axis values are loadings scaled by	

the respective long-term medians (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

.....	184
Figure 5.5. Intra-annual variations of river flow, total nitrogen (TN) load, total phosphorus (TP) load, and suspended sediment (SS) load in the nontidal Chesapeake Bay watershed (NTCBW) and the nine individual tributaries, as based on long-term (1970s-2012) median values. X-axis represents calendar months. Y-axis values have been scaled by the respective values in January. ....	188
Figure 5.6. Seasonal averages of streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the four largest tributaries (SUS, Susquehanna; POT, Potomac; JAM, James; RAP, Rappahannock). Y-axis values are values scaled by the respective long-term medians (see Table 5.2). ....	193
Figure 5.7. Seasonal averages of streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the five smallest tributaries (APP, Appomattox; PAM, Pamunkey; MAT, Mattaponi; PAT, Patuxent; CHO, Choptank). Y-axis values are values scaled by the respective long-term medians (see Table 5.2). ....	194
Figure 5.8. Seasonal fractional contributions of the four largest tributaries to the nontidal Chesapeake Bay watershed (NTCBW) for streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) (SUS, Susquehanna; POT, Potomac; JAM, James; RAP, Rappahannock).....	198
Figure 5.9. Seasonal fractional contributions of the five smallest tributaries to the nontidal Chesapeake Bay watershed (NTCBW) for streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) (APP, Appomattox; PAM, Pamunkey; MAT, Mattaponi; PAT, Patuxent; CHO, Choptank). ....	199
Figure 5.10. Seasonal TN:TP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW) and the nine individual tributaries, as calculated based on WRTDS flow-normalized estimates. On each panel, y-axis indicates the N:P ratios, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P (10 ≤ ratio ≤ 20). ....	202
Figure 5.11. Annual TN:TP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW) and the nine individual tributaries, as calculated based on WRTDS flow-normalized	

estimates. On each panel, y-axis indicates the N:P ratios, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ). ..... 203

Chapter 6

Figure 6.1. Comparison of hydrograph and suspended sediment (SS) chemograph for the Marietta and Conowingo sites during two major storm events, *i.e.*, 2004 Hurricane Ivan and 2011 Tropical Storm Lee, and WRTDS concentration estimates obtained based on three different sampling scenarios. Horizontal dashed lines indicate the scour threshold of 11,300 m<sup>3</sup>/s. Vertical dashed lines indicate the key dates during each event, for which the differences of estimates among the three sampling scenarios are summarized in Table E2..... 226

Figure 6.2. Estimated suspended sediment concentration as a function of discharge for the Conowingo site for four selected dates, *i.e.*, September 1 of 1990 and 2005 (after Hirsch, 2012) and May 15 of 1990 and 2005 (after Zhang *et al.*, 2013). Note that both axes are on logarithmic scale. Dashed lines indicate 95% confidence intervals, which were obtained from uncertainty analyses with 100 bootstrap model runs..... 230

Figure 6.3. Flow-normalized annual load trends of suspended sediment at the Conowingo site, as obtained with different scenarios of sampling data. Solid blue line is based on the original concentration record. Dashed blue lines indicate 95% confidence intervals, as obtained from uncertainty analyses of 100 bootstrap runs. Solid green line is based on a flow-censored concentration record, which eliminates all concentration samples with discharge ( $Q$ ) > scour threshold (*i.e.*, 11,300 m<sup>3</sup>/s). Solid purple line is based on censored records for both concentration and discharge, which eliminates all concentration samples with  $Q > 11,300 \text{ m}^3/\text{s}$  and replaces  $Q$  above 11,300 m<sup>3</sup>/s with a nominal value (*i.e.*, 20 m<sup>3</sup>/s)..... 232

Chapter 7

Figure 7.1. Illustration of common issues on interpretation of concentration-discharge ( $C-Q$ ) relationships using the example of suspended sediment concentration data in Susquehanna River at Conowingo, MD. (a) Direct linear fit (red line) to the entire data sets assumes linearity in the  $C-Q$  relationship. The LOWESS fit (blue line) doesn't invoke such an assumption. (b) LOWESS fits for  $C-Q$  data in three different

temporal periods. (c) LOWESS fits for $C-Q$ data in four different seasons. (d) LOWESS fits for $C-Q$ data in four quartiles of discharge.....	242
Figure 7.2. Contour plot showing estimated WRTDS $\beta_2$ coefficients as a function of time and discharge for suspended sediment in Susquehanna River at Conowingo, MD. Black open circles indicate the $t-Q$ combinations where concentration samples were taken. ....	250
Figure 7.3. Contour plot showing estimated change in $\beta_2$ coefficients between year 1990 and year 2010 for suspended sediment in Susquehanna River at Conowingo, MD. ....	251
Figure 7.4. Boxplots of estimated daily $\beta_2$ coefficients by (a) discharge percentile and (b) calendar month for suspended sediment in Susquehanna River at Conowingo, MD. The dashed black lines indicates the estimated slope (0.83) obtained from linear fit to all data – see Figure 7.1a. The red dashed line indicates no sensitivity to discharge ( <i>i.e.</i> , chemostasis). ....	253
Figure 7.5. Boxplots of estimated $\beta_2$ coefficients by calendar month for a selected discharge, <i>i.e.</i> , 885 m <sup>3</sup> /s (approximately the median value of the daily discharge record), for suspended sediment in Susquehanna River at Conowingo, MD. These coefficients were extracted from the $\beta_2$ surface ( <i>i.e.</i> , Figure 7.2) for this particular discharge throughout the period of record. ....	254
Figure 7.6. Annual averages of estimated $\beta_2$ coefficients for three selected discharges for suspended sediment in Susquehanna River at Conowingo, MD. These coefficients were extracted from the $\beta_2$ surface ( <i>i.e.</i> , Figure 7.2) for these hypothetical cases of constant discharge throughout the period of record. The dashed lines represent the 90% confidence interval for coefficients derived from 50 replicates model runs that were based on resampled (with replacement) concentration data – see Section 7.2. ....	256

## Chapter 8

Figure 8.1. Chesapeake Bay watershed and the 15 monitoring sites that include nine River Input Monitoring (RIM) sites on the fall-line of nine major tributaries and six Susquehanna River Basin Commission (SRBC) sites at upstream locations within the Susquehanna River basin. This figure was modified after Figure 1 in Moyer <i>et al.</i> (2012). ....	273
Figure 8.2. Contour plot showing estimated WRTDS $\beta_2$ coefficients as a function of time and discharge for total phosphorus in Susquehanna River at Conowingo, MD. Black open circles indicate the time-discharge combinations where concentration samples have	

been taken. The  $\beta_2$  coefficients correspond to three broad categories, namely, (1) dilution (*i.e.*,  $\beta_2 < 0$ ); (2) chemostasis ( $\beta_2 \approx 0$ ); and (3) mobilization ( $\beta_2 > 0$ )..... 279

Figure 8.3. Boxplot summary of estimated WRTDS  $\beta_2$  coefficients by discharge decile for suspended sediment (SS) at the 15 Chesapeake sites. (X-axis: 1 = 0<sup>th</sup>~10<sup>th</sup>, 2 = 10<sup>th</sup>~20<sup>th</sup> ..., 9 = 80<sup>th</sup>~90<sup>th</sup>, 10 = 90<sup>th</sup>~100<sup>th</sup>.) ..... 281

Figure 8.4. Boxplot summary of estimated WRTDS  $\beta_2$  coefficients by discharge decile for total phosphorus (TP) at the 15 Chesapeake sites. (X-axis: 1 = 0<sup>th</sup>~10<sup>th</sup>, 2 = 10<sup>th</sup>~20<sup>th</sup> ..., 9 = 80<sup>th</sup>~90<sup>th</sup>, 10 = 90<sup>th</sup>~100<sup>th</sup>.) ..... 282

Figure 8.5. Boxplot summary of estimated WRTDS  $\beta_2$  coefficients by discharge decile for total nitrogen (TN) data at the 15 Chesapeake sites. (X-axis: 1 = 0<sup>th</sup>~10<sup>th</sup>, 2 = 10<sup>th</sup>~20<sup>th</sup> ..., 9 = 80<sup>th</sup>~90<sup>th</sup>, 10 = 90<sup>th</sup>~100<sup>th</sup>.) ..... 283

Figure 8.6. Annual averages of estimated WRTDS  $\beta_2$  coefficients for three selected discharges for suspended sediment (SS) at the 15 Chesapeake sites. Dashed lines represent the 90% confidence interval as derived from 50 bootstrap runs based on resampled data. .... 287

Figure 8.7. Annual averages of estimated WRTDS  $\beta_2$  coefficients for three selected discharges for total phosphorus (TP) at the 15 Chesapeake sites. Dashed lines represent the 90% confidence interval as derived from 50 bootstrap runs based on resampled data. .... 288

Figure 8.8. Annual averages of estimated WRTDS  $\beta_2$  coefficients for three selected discharges for total nitrogen (TN) at the 15 Chesapeake sites. Dashed lines represent the 90% confidence interval as derived from 50 bootstrap runs based on resampled data. .... 289

Figure 8.9. Comparison between period-of-record decline in riverine yield and decline in total source input for total nitrogen (TN) across the 15 sites. Blue dashed line is 1:1 reference line. .... 296

## Chapter 9

Figure 9.1. Time series of (a) daily discharge and (b-m) the derived flow variables for Maumee River at Waterville, OH between 2005 and 2009..... 322

Figure 9.2. Example time series of (a) the original NO<sub>x</sub> concentration data (black open circle) and (b-d) three types of concentration subsets (non-black solid circle) in Maumee River at Waterville, OH between 2005 and 2006. See text for details of the three sampling strategies..... 326

Figure 9.3. Comparison of the original WRTDS model and the *LTF*A (*long-term flow anomaly*) model for NO<sub>x</sub> in Grand River near Painesville, OH. (a) Time series of residuals from the original model. (b) Relationship between the original model residuals and LTF

as modeled with LOWESS curve (green line). Spearman's correlation ( $\gamma$ ) is shown on the plot. (c) Comparison between the LTFA model residuals and the original model residuals. (d) Enlarged view of the first quadrant of plot (c). (e) Time series of annual concentration based on observed data, the original model, and the LTFA model. ... 330

Figure 9.4. Comparison of the original WRTDS model and the *STFA* (*short-term flow anomaly*) model for Cl in Maumee River at Waterville, OH. (a) Time series of residuals from the original model. (b) Relationship between the original model residuals and STFA, as modeled with LOWESS curve (green line). Spearman's correlation ( $\gamma$ ) is shown on the plot. (c) Comparison between the STFA model residuals and the original model residuals. (d) Enlarged view of the first quadrant of plot (c). (e) Time series of annual concentration based on observed data, the original model, and the STFA model..... 333

Figure 9.5. Boxplots of ratio between the modified Nash-Sutcliffe efficiency (E) of each modified model and E-value of the original model for the nine sites for (a) Cl, (b) NO<sub>x</sub>, (c) TKN, (d) SRP, (e) TP, and (f) SS. The ratios are divided into three regions: (1) major improvement (ratio > 1.2; blue bars), (2) minor improvement (1 < ratio < 1.2; grey bars), and (3) inferior performance (ratio < 1.0; red bars). Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy A*..... 336

Figure 9.6. Boxplots of absolute values of Spearman's correlation coefficient ( $|\gamma|$ ) between residuals from the original WRTDS model and each of the twelve proposed flow variables. Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy A*..... 339

Figure 9.7. Performance of the original WRTDS model under each of the three sampling strategies. The performance is quantified with modified Nash-Sutcliffe efficiency (E) for (a) concentration and (b) flux and with absolute percent bias (|PBIAS|) for (c) concentration and (d) flux. Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy..... 342

Figure 9.8. Coefficient of variation in daily discharge (Q) data and constituent concentration data. Each boxplot summarizes nine coefficients, each of which was based on data at one individual monitoring site. .... 343

Figure 9.9. Performance of the original and modified models under the three sampling strategies. Plots (a)-(c) summarizes the performance with medians of modified Nash-Sutcliffe efficiency (E) for *concentration* based on 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy. Plots (d)-(f) compares the three sampling strategies directly using ratios of E median..... 345

Figure 9.10. Performance of the original and modified models under the three sampling strategies. Plots (a)-(c) summarizes the performance with medians of absolute value of percent bias (|PBIAS|) for *flux* based on 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy. Plots (d)-(f) compares the three sampling strategies directly using ratios of |PBIAS| median. .... 346

Chapter 10

Figure 10.1. Synthetic time series with 200 time steps for three representative LRD strengths that correspond to white noise ( $\beta = 0$ ), pink noise ( $\beta = 1$ ), and brown noise ( $\beta = 2$ ). The 1<sup>st</sup> row shows the simulated time series without any gap. The three rows below show the same time series as in the 1<sup>st</sup> row but with data gaps that were simulated using three different negative binomial (NB) distributions -- 2<sup>nd</sup> row: NB( $\lambda = 1, \mu = 1$ ); 3<sup>rd</sup> row: NB( $\lambda = 1, \mu = 14$ ); 4<sup>th</sup> row: NB( $\lambda = 0.01, \mu = 1$ ). .... 366

Figure 10.2. Examples of gap interval simulation using binomial distributions, NB (shape  $\lambda$ , mean  $\mu$ ). Simulation parameters:  $L = 9125$  days,  $\Delta t_{nominal} = 1$  day. The three panels show simulation with fixed (a)  $\mu = 1$ , (b)  $\mu = 14$ , and (c)  $\lambda = 1$ . Note that  $\Delta t_{average} / \Delta t_{nominal} = \mu + 1$ . .... 372

Figure 10.3. Examples of quantified sampling irregularity with negative binomial (NB) distributions: total nitrogen in Choptank River (top) and total phosphorus in Cuyahoga River (bottom). Theoretical CDF and quantiles are based on the fitted NB distributions. See Table 10.1 for estimated mean and shape parameters. .... 374

Figure 10.4. Illustration of the interpolation methods for gap-filling. The gap-free data (A1) was simulated with a series length of 500, with the first 30 data shown. (x: omitted data for gap-filling; +: interpolated data; NOCB: next observation carried backward; LOCF: last observation carried forward; lowess: locally weighted scatterplot smoothing.). 380

Figure 10.5. Comparison of bias in estimated LRD strengths in irregular data that are simulated with  $\beta = 1$  (30 replicates), series length of 9125, and gap intervals simulated with (a) NB ( $\lambda = 0.01, \mu = 1$ ), (b) NB ( $\lambda = 0.1, \mu = 1$ ), (c) NB ( $\lambda = 1, \mu = 1$ ), and (d) NB ( $\lambda = 10, \mu = 1$ ). .... 382

Figure 10.6. Comparison of bias in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 2. .... 386

Figure 10.7. Comparison of root-mean-squared error (RMSE) in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 2. ....	387
Figure 10.8. Comparison of bias in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 15. ....	388
Figure 10.9. Comparison of root-mean-squared error (RMSE) in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 15. ....	389
Figure 10.10. Quantification of LRD strengths in real water-quality data from the two regional monitoring networks, as estimated using the set of examined methods. All estimations were performed on concentration residuals (in natural log concentration units) after accounting for effects of time, discharge, and season. The two dashed lines in each panel indicate white noise ( $\beta = 0$ ) and flicker noise ( $\beta = 1$ ), respectively. See Table 10.1 for site and data details. ....	392

## Chapter 1. Introduction

### 1.1. Chesapeake Bay Hypoxia and Riverine Inputs

Chesapeake Bay has experienced persistent summertime hypoxia in its bottom waters, which has been attributed to a combination of anthropogenic nutrient and sediment inputs from the bay watershed (Malone *et al.*, 1988; Boynton and Kemp, 2000; Kemp *et al.*, 2005) and naturally occurring vertical stratification (Boicourt, 1992; Pritchard and Schubel, 2001; Murphy *et al.*, 2011). Of the two factors, the more attainable means of controlling hypoxia is the reduction of nitrogen (N), phosphorus (P), and suspended sediment (SS) inputs, which have been a principal focus of Chesapeake Bay Watershed (CBW) management for decades. In 2010, the strength of this reduction endeavor was enhanced by newly promulgated regulations on total maximum daily loads (TMDLs) for Chesapeake Bay (USEPA, 2010).

The Chesapeake Bay Modeling System is a computational system that is used currently toward establishing such TMDLs, using projected loadings that are based on many factors, including land use, and under assumed climatic conditions (Linker *et al.*, 2013; Shenk and Linker, 2013). In this context, the Modeling Workgroup of the Chesapeake Bay Program partnership has recently identified the following as the first of its “priority items” for its 2017 Midpoint Assessment of the Chesapeake Bay TMDL: “Revisiting the Watershed Model calibration with the goal of improving local watershed results” (Chesapeake Bay Program Office, 2012). One specific need in this regard is to better understand the nature and causes of historical load changes from different regions of the watershed so that these changes can be compared with watershed model output and

used to improve the model. Such understandings will be essential to help managers shape future strategies of river monitoring and adaptive management. Toward that end, one must first better understand the applicability and uncertainty of the statistical methods for evaluating historical records of nutrient and sediment data.

## **1.2. River Water-Quality Monitoring Program**

Chesapeake Bay is a long, partially mixed estuary extending from its seaward end at the Virginia capes, Cape Charles and Cape Henry, Virginia, to the mouth of the Susquehanna River at Harve de Grace, Maryland (Pritchard and Schubel, 2001). The Bay estuarine system is made up of the Bay proper and more than fifty tributaries (Pritchard and Schubel, 2001). To assess nutrient and sediment trends in these tributaries, the USGS and collaborators have been monitoring water quality at many sites for decades (Langland *et al.*, 2007). Particularly, the USGS River Input Monitoring (RIM) Program has been monitoring sites at the fall line of nine major tributaries to Chesapeake Bay since the 1980s (Figure 1.1), which collectively account for more than 90% of streamflow entering the bay from its non-tidal watershed (USGS, 2012). Because these sites are above the tidal portion of the tributaries, any trends observed there can be attributed to upstream causes (USGS, 2012). The USGS has been periodically analyzing nutrient and sediment data at the RIM sites and other sites in the CBW, and publishing loads and trends. For example, Langland *et al.* (2007) reported significantly decreasing trends in flow-adjusted annual concentration of total N (TN), total P (TP), and SS from 1985 to 2006 at about 74%, 68%, and 32% of 34 selected sites in the CBW, respectively. In addition, the Susquehanna River Basin Commission (SRBC) has also periodically

collected and analyzed monitoring data at several long-term sites through the Susquehanna Nutrient Assessment Program (SNAP) since the mid-1980s (SRBC, 2012).

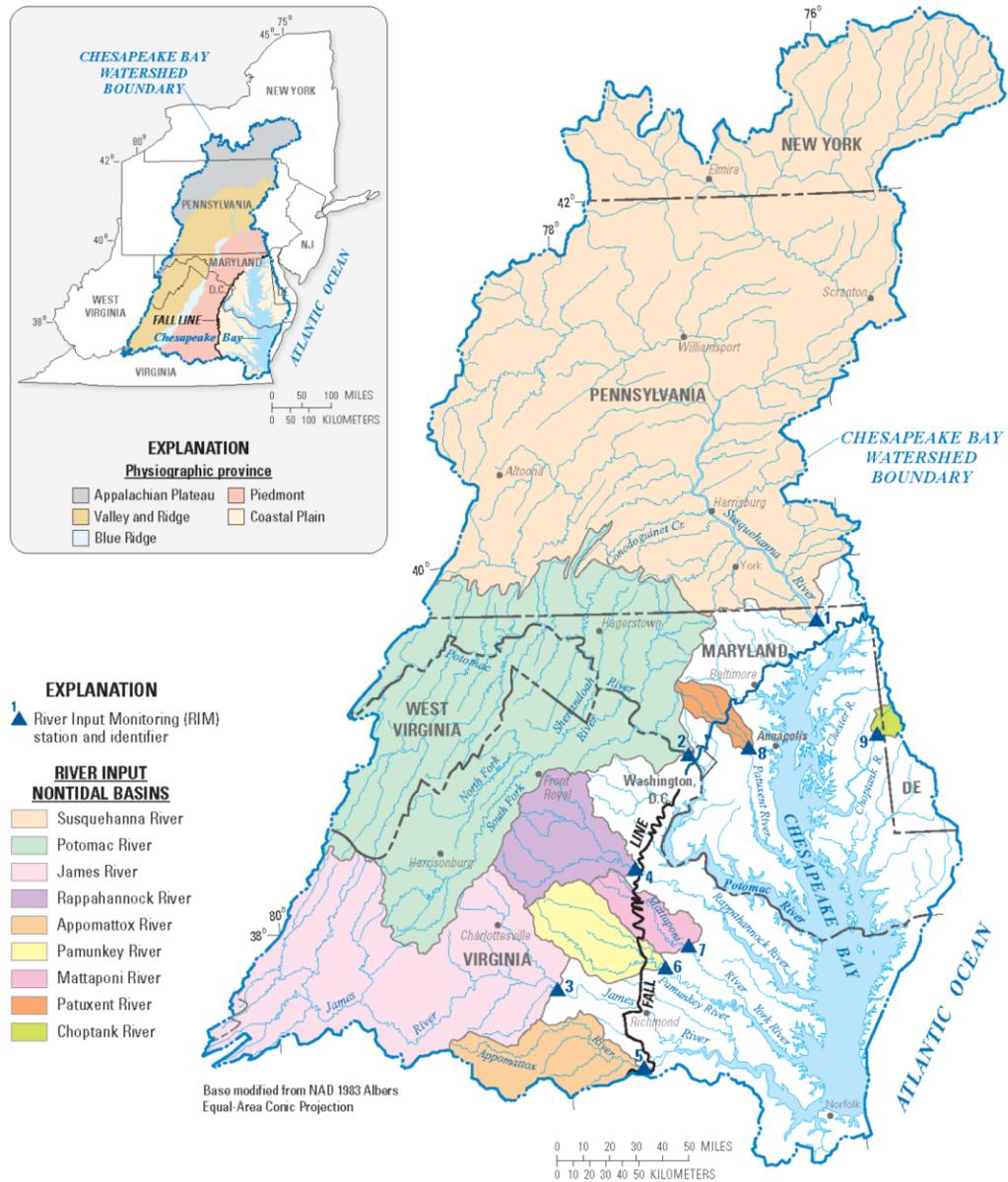


Figure 1.1. The Chesapeake Bay watershed and monitoring sites at the fall-line of the nine major tributaries – Susquehanna River (1), Potomac River (2), James River (3), Rappahannock River (4), Appomattox River (5), Pamunkey River (6), Mattaponi River (7), Patuxent River (8), and Choptank River (9). Inset shows the major physiographic provinces. This figure was reproduced from Moyer *et al.* (2012) with permission.

A common feature of the USGS and SRBC trend analyses is the adoption of annual resolution. To capture impacts of seasonality on the loading trends, such as variations in temperature and rainfall, timing of fertilizer application, and benthic recycling of P and denitrification in river channels and reservoirs, seasonal trends of nutrient and sediment need to be examined. In addition, seasonal estimates can also help identify nutrients dynamics in some critical periods, including at least the winter-spring period (*i.e.*, Jan-May) trend of TN load from the SRB in this seasonal period has been closely linked to variation of summer hypoxia in Chesapeake Bay (Hagy *et al.*, 2004).

### **1.3. Riverine Flux Estimation and Trend Analysis with the WRTDS Method**

The problem of quantifying nutrient and sediment loads and allocating them to sources is not new, nor is it unique to the Chesapeake region (Drolc and Koncan 2008; (Jarvie *et al.*, 1998; Lidén *et al.*, 1999; Stow *et al.*, 2001; Pieterse *et al.*, 2003; Harned *et al.*, 2004; Némery *et al.*, 2005; Drolc and Koncan, 2008). Overall, the issue can be broken into the following parts, each of which brings its own challenges:

- 1) detecting and evaluating historical trends in flow, concentration, and especially loading of N, P, and SS (Cohn *et al.*, 1989; Hirsch *et al.*, 1991; Hirsch *et al.*, 2010);
- 2) understanding the spatial and temporal changes in load-generating factors, such as point-source loading, air quality, and land use (Jordan *et al.*, 1997a; Lizarraga, 1997; Sharpley *et al.*, 1999; Boynton *et al.*, 2008); and
- 3) comparing loading trends and affecting factors to better understand the relations (Sprague *et al.*, 2000; Ator *et al.*, 2011; Beckert *et al.*, 2011; Hively *et al.*, 2011).

The first component is itself a complex topic, and especially so because we seek data that are somehow “normalized” with respect to those effects that do not involve long-term watershed management, such as discharge (related to rainfall and other factors), season, or other causes for short-term changes. Challenges include (1) accurately estimating the daily loading history from sparse data of concentrations; (2) separating out the underlying trend from the variability related to time, discharge, and season; and (3) understanding causative factors behind the observed long-term changes.

In regard to the first two challenges identified above, Hirsch *et al.* (2010) have recently developed a statistical method called “Weighted Regressions on Time, Discharge, and Season” or WRTDS. It uses daily river flow and less frequent (usually bi-weekly or monthly) nutrient and sediment concentration data to estimate concentrations and loadings for every day in the record, using the following principle equation:

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (1.1)$$

where  $C$  is constituent concentration,  $t$  is time in decimal years,  $Q$  is daily streamflow discharge,  $\beta_i$  are fitted coefficients, and  $\varepsilon$  is the error term. Thus, a unique set of fitted coefficients is obtained separately for each estimation day (Hirsch *et al.*, 2010). As compared to the ESTIMATOR model that has been more conventionally used (Cohn *et al.*, 1989), the primary advancements of WRTDS include (1) better description of temporal changes in concentration and load, and (2) more effective removal of the influence of random flow variation. Improvements derive in part from the fact that WRTDS does not rely on questionable assumptions about homoscedasticity of model errors, constancy of the concentration-flow relationship (or specific functional forms), and constancy of seasonal trends in concentration (Hirsch *et al.*, 2010).

Of special interest is that WRTDS can produce two types of estimates for both concentration and load – so-called “*true-condition*” and “*flow-normalized*” estimates. By integrating the full hydrological cycle, the *flow-normalized* estimates can largely remove the sometimes-dramatic influence of year-to-year random variations in streamflow (but not the seasonal variability), thus making the inter-annual trends easier to detect than they would be with the *true-condition* estimates (Hirsch *et al.*, 2010). Hirsch *et al.* (2010) have pointed out that the *true-condition* estimates are useful to help understand the real history of water quantity and quality as relevant to downstream ecological impact, whereas the *flow-normalized* estimates are more helpful to evaluate management progress in the upstream watershed – *i.e.*, with respect to influencing factors that are less related to river-flow.

To quantify uncertainty in WRTDS estimates of *true-condition* loads, *flow-normalized* loads, or slopes of trends in *flow-normalized* loads, a “bootstrap” approach has been developed very recently (Hirsch *et al.*, 2015), which is based on resampling (with replacement) from the original data and re-estimating the model. Bootstrap is a well-established statistical method for evaluating the uncertainty of complex estimators (which WRTDS certainly is) particularly when standard distributional assumptions are inappropriate (Efron and Tibshirani, 1993). This approach can provide one measure of uncertainty that can help clarify if apparent trends are really meaningful departures from past system behavior or are simply variations that can be expected to arise by chance from a stationary system. Such information is crucial to the evaluation of progress towards meeting water quality goals (Hirsch *et al.*, 2015).

Although WRTDS represents an improvement relative to some prior approaches, it is of course eligible for further testing and improvement. In the original method paper, Hirsch *et al.* (2010) listed several areas for WRTDS improvements. Some of these areas have already been well developed since then, including the “capability for analyzing censored data” and “simultaneous analysis of multiple nutrient constituents” (R. Hirsch, 2013; personal communication). In addition, the WRTDS method currently considers three factors for prediction of concentration in its basic model, *i.e.*, time, discharge, and season. Hirsch *et al.* (2010) have suggested the usefulness of additional explanatory variables that can account for flow hysteresis or watershed wetness condition. Although the terms will add complexity and potential model error to the model, they have mechanistic implications. “Flow hysteresis” should be accounted for because nutrient or sediment concentration can vary significantly, depending whether it is on the rising or falling limb of the rating curve at the time of sampling. “Watershed wetness” also matters because it can greatly influence momentum and pathway of material transport from watershed to stream. In a “dry” watershed, materials tend to be kept locally; materials in dissolved forms may enter the subsurface near its location of generation. When the watershed becomes wetter, particulate constituents can be more easily mobilized with surface runoff and dissolved constituents can be more efficiently flushed to stream due to saturation of subsurface storage. In addition, quadratic terms in discharge (in log scale) and time may also be considered as exploratory variables, which can account for the non-linear relation between concentration and discharge (both on log-scale). However, none of these variables has been heretofore evaluated (R. Hirsch, 2013; personal communication).

## 1.4. Source, Fate, and Transport of Nutrients and Sediment in Watersheds

Watershed export of nutrients and sediment is very complex due to watershed heterogeneities in source, fate, and transport (Figure 1.2). In terms of source, nutrient and sediment originate from a variety of source sectors, including primarily non-point source inputs of fertilizer, fixation, and manure; point sources inputs from urban runoffs and wastewater treatment plants; and atmospheric deposition (mainly for N). The relative contribution of each source sector varies significantly both spatially (as a function of watershed physical attributes such as land use, geology, vegetation cover, and climate) and temporally (as a function of watershed management, urban sprawl, and treatment technology improvement) (Carpenter *et al.*, 1998).

Typically, sediment and nutrients accumulate in various parts of the landscape in dry periods at low flow, and are transported as various flow paths (*i.e.*, groundwater, soil

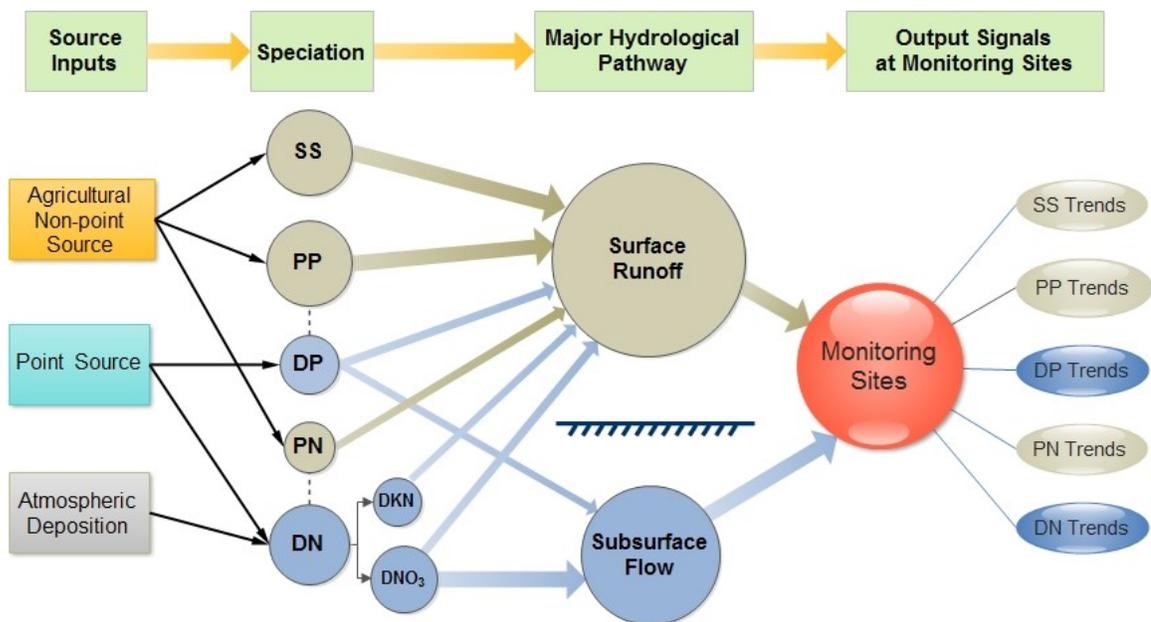


Figure 1.2. Source input, speciation, major hydrological pathway, and output signals of nitrogen, phosphorus, and suspended sediment.

water, and surface water) activate during wet periods or storm events (Shields *et al.*, 2008). These constituents are dominated by two distinct pathways, *i.e.*, surface (or overland) flow and subsurface flow. The former is important for sediment and nutrients attached to sediment particles (*e.g.*, particulate P, organic P, organic N, and ammonium), whereas the latter is important for soluble constituents such as nitrate-N and dissolved P (Holtan *et al.*, 1988; Heathwaite, 1995; Jordan *et al.*, 1997; Pionke *et al.*, 2000; Withers and Jarvie, 2008; Pärn *et al.*, 2012). Typically, P transport is more efficient in areas with more erodible soils, relatively poor drainage, and greater precipitation, whereas N transport is most efficient in areas with greater groundwater recharge, certain carbonate rocks, and less reducing conditions that may contribute to denitrification (Ator *et al.*, 2011).

Retention mechanisms also differ considerably between P and N – sedimentation may be the major retention mechanism for particulate P during overland flow (Cirimo and McDonnell, 1997; Hoffmann *et al.*, 2007), whereas during subsurface flow conditions, sorption/desorption reactions and denitrification are more important for P and nitrate, respectively (House, 2003; Withers and Jarvie, 2008). These processes are affected by a wide range of factors. P retention by sedimentation depends on flow velocity, soil characteristics (*e.g.*, water infiltration capacity, porosity and water repellence), and vegetation (Hoffmann *et al.*, 2007). P sorption/desorption reactions are affected by water pH, redox potential, soil moisture, particle size, and relative concentration in water and sediment (McDowell *et al.*, 2004). Denitrification is generally greater in poorly drained soils (Jordan *et al.*, 1997) and is highly influenced by temperature, soil texture, redox status, carbon availability, and hydraulic residence time (Cirimo and McDonnell, 1997).

## 1.5. Sediment and Nutrient Retention in Conowingo Reservoir

When considering overall riverine inputs of nutrient and sediment, it is important to recognize that such loadings can be greatly modulated by reservoirs. More specifically, reservoirs in early stages of their lifespan can effectively retain sediment and associated N and P (Jossette *et al.*, 1999; Friedl and Wüest, 2002; Medeiros *et al.*, 2011), thus providing efficient removal of N and P from streamflow, mainly through denitrification and particle deposition, respectively (Jossette *et al.*, 1999; Friedl and Wüest, 2002). In the CBW, the most studied reservoir system is probably the Lower Susquehanna River Reservoir System (LSRRS), which includes Lake Clarke, Lake Aldred, and Conowingo Reservoir (Figure 1.3) (Langland and Hainly, 1997; Langland, 2009). This system has been reported to trap about 2%, 45%, and 70% of annual TN, TP, and SS load, respectively, from Susquehanna to Chesapeake Bay (Langland and Hainly, 1997).

As the largest and most downstream reservoir in the system, Conowingo Reservoir is the only one that has not reached its sediment storage capacity (Langland, 2009). Based on assumptions in sediment input load and deposition rate, Langland (2009) estimated an additional service life of 15-20 years (from 2009) before the reservoir is filled up. By that time, annual loads of TN, TP, and SS to the Bay have been projected to increase by 2%, 70%, and 250%, respectively (Langland and Hainly, 1997). In addition, a previous study has suggested that a flow of 400,000 ft<sup>3</sup>/s (11,300 m<sup>3</sup>/s) was the “scour threshold” for Conowingo Reservoir, and that major floods above this level would further enhance sediment and nutrient delivery to Chesapeake Bay (Langland and Hainly, 1997). More recently, Hirsch (2012) detected upward trends in annual TP and SS load at the Conowingo Dam. It was hypothesized that, when reservoirs are near capacity, the water

channel would become smaller, thus resulting in faster water flow, and correspondingly, lower likelihood of deposition and higher likelihood of scour (Hirsch, 2012). Therefore, increased net scouring of sediment in Conowingo Reservoir may already be occurring at flow rates much lower than the above-cited scour threshold. Hirsch (2012) has suggested that filling processes in Conowingo Reservoir are already approaching a final asymptotic stage and that further monitoring and evaluation of the reservoir system is critical to evaluation of management plans.

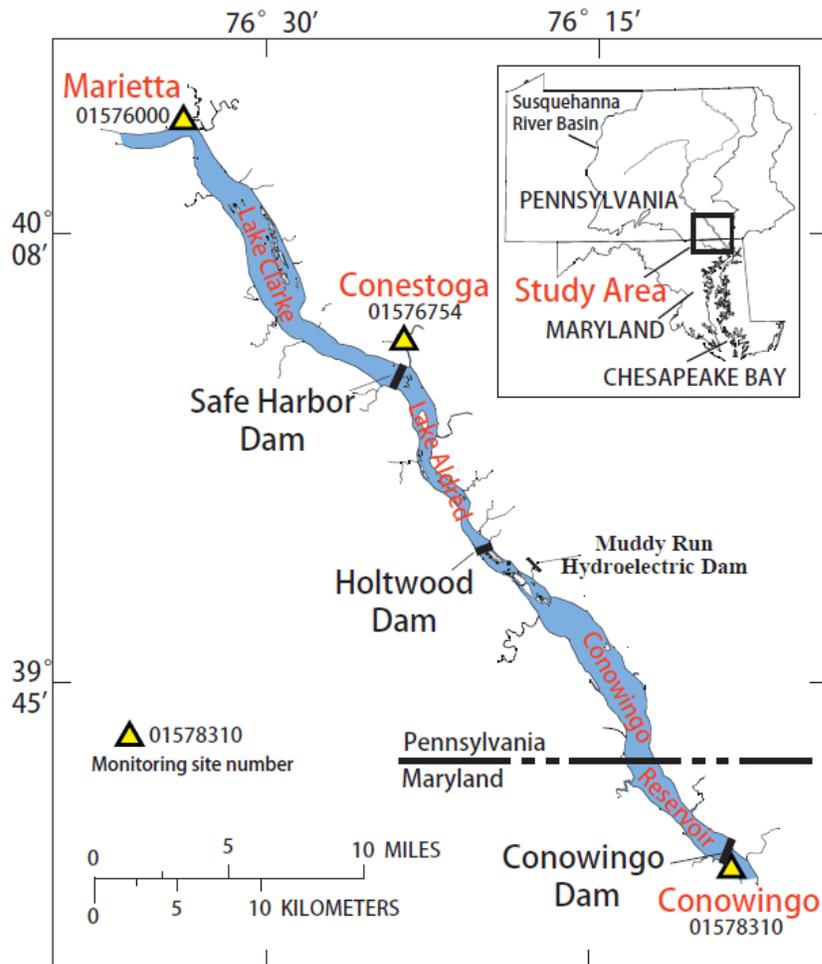


Figure 1.3. Map of the Lower Susquehanna River Reservoir System consisting of Lake Clarke, Lake Aldred, and Conowingo Reservoir. Yellow triangles indicate the three monitoring sites: Conowingo, Marietta, and Conestoga.

## 1.6. 1/f Scaling Signature in River Water-Quality Data and Implications

Recently, some new concerns have been raised about the validity of trends derived by statistical methods, from the perspective of time series analysis in the frequency domain (Kirchner and Neal, 2013). Similar to data from many other monitoring programs, water-quality data collected in the CBW typically include long-term low-frequency samplings (weekly or monthly) and episodic high-frequency samplings (daily) during storm events. In general, the high-frequency samples tend to reflect the watershed's behavior during storm events, whereas the low-frequency samples tend to reflect the watershed's average behavior (Feng *et al.*, 2004).

From the perspective of watershed hydrology, stream concentrations reflect the accumulative effects of rainfall inputs throughout the past, weighted by their fractional contribution to the present streamflow (Kirchner *et al.*, 2000; Kirchner *et al.*, 2001). For a conservative tracer originated mainly from rainfall, its streamflow concentration  $c_S(t)$  is the convolution of the rainfall concentration throughout the past  $c_R(t - \tau)$  and the travel time distribution  $h(\tau)$ , where  $\tau$  is the lag time between rainfall and runoff:

$$c_S(t) = \int_0^{\infty} h(\tau) c_R(t - \tau) d\tau \quad (1.2)$$

In addition, by the convolution theorem, Equation (1.2) implies the following:

$$C_S(f) = H(f)C_R(f) \quad (1.3)$$

$$|C_S(f)|^2 = |H(f)|^2 |C_R(f)|^2 \quad (1.4)$$

where  $f$  is frequency,  $C_S(f)$ ,  $H(f)$ , and  $C_R(f)$  are the Fourier transforms of  $c_S(t)$ ,  $h(\tau)$ , and  $c_R(t - \tau)$ , respectively, and  $|C_S(f)|^2$ ,  $|H(f)|^2$ , and  $|C_R(f)|^2$  are the corresponding

power spectra. Taking chloride (a natural tracer) as an example, the power spectrum of rainfall concentrations is nearly constant, representing white noise inputs.

In contrast, the sampled stream concentrations of chloride exhibit inverse proportionality between spectral power and frequency (Kirchner *et al.*, 2000; Kirchner *et al.*, 2001). This phenomenon, often referred to as “ $1/f^\alpha$  scaling” (with the power law slope close to  $\alpha=1$ ), arises because the watershed can store, transport, and mix solutes over a diverse range of spatial and temporal scales. In other words, the catchment can act as a fractal filter that converts the typical “white noise” inputs in rainfall to  $1/f^\alpha$  noise outputs in stream runoff (Kirchner *et al.*, 2000; Kirchner *et al.*, 2001). Moreover, recent analysis on stream data at Plynlimon (UK) suggests that  $1/f^\alpha$  scaling at this location was not limited to conservative tracers (Kirchner and Neal, 2013). Instead, it was found to be a “universal” phenomenon applicable to many other water-quality parameters, including nutrients, metals, and trace elements. Therefore, one might expect that nutrient (or sediment) concentrations in the CBW should also exhibit the  $1/f^\alpha$  scaling signature.

More importantly, the recent study by Kirchner and Neal (2013) has revealed two important implications on water-quality trend detection. First, the  $1/f^\alpha$  signals are not “self-averaging”, which implies that longer-time averages of monitoring data do not necessarily converge to a stable. This is contradictory to conventional statistics derived from the Central Limit Theorem. Second, statistically significant trends arise much more frequently on all time scales than one would expect from conventional t-statistics (Kirchner and Neal, 2013). These results suggest the necessity of carefully examining the validity of trends derived by statistical methods. Because LRD processes can induce trend-like behaviors, statistically significant trends may be falsely declared more

frequently than expected. Currently, there is still lack of understanding on LRD estimation approaches for river water-quality data, which are typically irregular and thus cannot be examined by traditional methods.

## **1.7. Objectives and Outline**

Motivated by the need to advance the scientific understanding in nutrient and sediment export from the Chesapeake Bay watershed, this dissertation has sought to apply the WRTDS method to long-term monitoring data for various major tributaries, to further explore the method's uncertainties, and to improve the method in regard to concentration and flux estimation. Specific objectives by chapter are listed below.

In Chapter 2, I analyze the long-term seasonal trends of flow-normalized nutrient and sediment loadings in Susquehanna River over the last two to three decades, both above and below the Lower Susquehanna River Reservoir System, in order to (1) evaluate progress in reduction of N, P, and SS loads from the non-tidal SRB at seasonal resolution and (2) compare the relative changes in N, P, and SS loads discharging into and emanating from the reservoirs as a means of evaluating reservoir performance in terms of sediment and nutrient retention.

In Chapter 3, I further evaluate the decadal-scale history of concentrations and fluxes at sites above and below the reservoir system between 1986 and 2013 (~30 years), in order to provide new insights on sediment and nutrient processing within the reservoir system. Specifically, three types of analyses were conducted: (1) identification of changes in concentration-discharge relationships at sites above and below the reservoir; (2) evaluation of net deposition in the reservoir system using mass-balance analysis; and (3) analysis of the effects of sediment accumulation on reservoir performance by better

accommodating effects of streamflow variability through the development of three different stationary models of the concentration relation to discharge and season.

In Chapter 4, I extend the retrospective analysis of historical data to all seven long-term monitoring locations in the Susquehanna River basin to conduct a comprehensive synthesis of (1) temporal trends of nutrient and sediment loadings at different locations of the basin and (2) spatial variations of nutrient and sediment budgets of major sub-basins. Particular focuses include comparison of long-term trends by constituent and by site, comparison of changes in watershed source inputs and those in riverine loadings, quantification of relative contributions of the sub-basins, and effects of streamflow and land use on constituent export.

In Chapter 5, I extend the retrospective analysis of historical data to all nine major non-tidal tributaries in the non-tidal Chesapeake Bay watershed, with special attention to comparisons of nutrient and sediment trends by rivers and by seasons. Specific new contributions include: (1) analysis of historical non-tidal streamflow and concentration data for all nine major tributaries, their summed loadings, and the fractional contribution of each tributary to the summed total; (2) separate analysis of multiple forms of N and P; and (3) analysis of seasonal trends.

In Chapter 6, I re-examine the changing dynamics of sediment and nutrients processing within the Lower Susquehanna River Reservoir System to address an overall concern as to whether the limited number of concentration samples at extremely high discharges may have significantly affected the regression surfaces for concentration (*i.e.*, for high, middle, and low discharges) and thus affected loading estimates, especially when samples are not evenly distributed temporally and spatially over the record.

In Chapter 7, I illustrate the several issues encountered when riverine monitoring data are commonly interpreted with log-linear concentration-discharge ( $C-Q$ ) relationships for understanding export dynamics. I propose a robust, informative, and accessible solution for the interpretation of  $C-Q$  patterns in riverine monitoring data through the recently-developed WRTDS method. Specifically, I show how the model's estimates for such coefficient can be organized and presented in ways that provide additional insights toward the interpretation of water-quality data using an example water-quality record.

In Chapter 8, I apply the WRTDS method to interpret the nature and change of  $C-Q$  relationships for nutrients and sediment at multiple long-term sites across the Chesapeake Bay watershed. In this top-down synthesis, I seek to address two main questions: (1) How does the nature of  $C-Q$  relationships (*i.e.*, dilution, chemostasis, and mobilization) vary by discharge condition and how do the patterns compare among sites and species? (2) Are  $C-Q$  relationships temporally varying (*i.e.*, non-stationary) under different discharge conditions and how do the patterns compare among sites and species?

In Chapter 9, I develop enhanced statistical models for estimation of riverine concentration and flux by simultaneously accounting for both current and antecedent discharge conditions. Specifically, the current WRTDS model is modified by adding new variables that represent antecedent discharge condition, including flow anomalies, average discounted flow, base-flow index, and flow gradient. Estimation performance of the original and modified models is evaluated for six common water-quality constituents using high-resolution (nearly daily) data at nine monitoring sites. For each site-constituent combination, concentration subsets were generated through Monte Carlo subsampling to mimic three common sampling strategies.

In Chapter 10, I systematically evaluate and compare a series of approaches for estimating LRD in irregular river water-quality time series, including (1) examination of sampling irregularity of typical river water-quality monitoring data, (2) Monte Carlo simulation of synthetic time series that contain such irregularity, and (3) comparison of approaches for estimating LRD in simulated irregular time series.

In Chapter 11, I summarize the major findings and their implications. I also discuss the applications of this research and potential areas of future research.

## **1.8. Literature Cited**

Ator, S. W., J. W. Brakebill and J. D. Blomquist, 2011. Sources, fate, and transport of nitrogen and phosphorus in the Chesapeake Bay watershed: An empirical model. U.S. Geological Survey Scientific Investigations Report 2011-5167, Reston, VA, p. 27. <http://pubs.usgs.gov/sir/2011/5167/>.

Beckert, K. A., T. R. Fisher, J. M. O'Neil and R. V. Jesien, 2011. Characterization and Comparison of Stream Nutrients, Land Use, and Loading Patterns in Maryland Coastal Bay Watersheds. *Water, Air, Soil Pollut.* 221:255-273, DOI: 10.1007/s11270-011-0788-7.

Boicourt, W. C., 1992. Influences of circulation processes on dissolved oxygen in the Chesapeake Bay. *In: Dissolved Oxygen in Chesapeake Bay*, D. Smith, M. Leffler and G. Mackiernan (Editors). Maryland Sea Grant College, College Park, MD, pp. 7-59.

Boynton, W. R., J. D. Hagy, J. C. Cornwell, W. M. Kemp, S. M. Greene, M. S. Owens, J. E. Baker and R. K. Larsen, 2008. Nutrient Budgets and Management Actions in the Patuxent River Estuary, Maryland. *Estuaries Coasts* 31:623-651, DOI:

10.1007/s12237-008-9052-9.

Boynton, W. R. and W. M. Kemp, 2000. Influence of river flow and nutrient loads on selected ecosystem processes: A synthesis of Chesapeake Bay data. *In: Estuarine Science: A Synthesis Approach to Research and Practice*, J. E. Hobbie (Editors). Island Press, Washington, D.C., pp. 269-298.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley and V. H. Smith, 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* 8:559-568, DOI: 10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2.

Chesapeake Bay Program Office, 2012. Modeling Workgroup Midpoint Assessment Priority Items. [www.chesapeakebay.net/channel\\_files/18740/modeling\\_workgroup\\_priorities\\_10-12-12.pdf](http://www.chesapeakebay.net/channel_files/18740/modeling_workgroup_priorities_10-12-12.pdf).

Cirmo, C. P. and J. J. McDonnell, 1997. Linking the hydrologic and biogeochemical controls of nitrogen transport in near-stream zones of temperate-forested catchments: a review. *Journal of Hydrology* 199:88-120, DOI: 10.1016/S0022-1694(96)03286-6.

Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.

Drolc, A. and J. Z. Koncan, 2008. Diffuse sources of nitrogen compounds in the Sava river basin, Slovenia. *Desalination* 226:256-261, DOI: 10.1016/j.desal.2007.01.244.

- Efron, B. and R. Tibshirani, 1993. *An Introduction to the Bootstrap*, Chapman & Hall, New York, NY, ISBN 0412042312
- Feng, X., J. W. Kirchner and C. Neal, 2004. Measuring catchment-scale chemical retardation using spectral analysis of reactive and passive chemical tracer time series. *Journal of Hydrology* 292:296-307, DOI: 10.1016/j.jhydrol.2004.01.012.
- Friedl, G. and A. Wüest, 2002. Disrupting biogeochemical cycles—Consequences of damming. *Aquat. Sci.* 64:55-65, DOI: 10.1007/s00027-002-8054-0.
- Hagy, J. D., W. R. Boynton, C. W. Keefe and K. V. Wood, 2004. Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658, DOI: 10.1007/bf02907650.
- Harned, D. A., J. B. Atkins and J. S. Harvill, 2004. Nutrient Mass Balance and Trends, Mobile River Basin, Alabama, Georgia, and Mississippi. *J. Am. Water Resour. Assoc.* 40:765-793, DOI: 10.1111/j.1752-1688.2004.tb04458.x.
- Heathwaite, L., 1995. Sources of eutrophication: hydrological pathways of catchment nutrient export. *In: Man's influence on freshwater ecosystems*. IAHS Publications, Wallingford, pp. 161-175.
- Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.
- Hirsch, R. M., R. B. Alexander and R. A. Smith, 1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resour. Res.* 27:803-

813, DOI: 10.1029/91WR00259.

- Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Hively, W. D., C. J. Hapeman, L. L. McConnell, T. R. Fisher, C. P. Rice, G. W. McCarty, A. M. Sadeghi, D. R. Whitall, P. M. Downey, G. T. Niño de Guzmán, K. Bialek-Kalinski, M. W. Lang, A. B. Gustafson, A. J. Sutton, K. A. Sefton and J. A. Harman Fetcho, 2011. Relating nutrient and herbicide fate with landscape features and characteristics of 15 subwatersheds in the Choptank River watershed. *Sci. Total Environ.* 409:3866-3878, DOI: 10.1016/j.scitotenv.2011.05.024.
- Hoffmann, C. C., C. Kjaergaard, J. Uusi-Kämpä, H. C. B. Hansen and B. Kronvang, 2007. Phosphorus retention in riparian buffers: review of their efficiency. *J. Environ. Qual.* 38:1942-1955, DOI: 10.2134/jeq2008.0087.
- Holtan, H., L. Kamp-Nielsen and A. O. Stuanes, 1988. Phosphorus in soil, water and sediment: an overview. *Hydrobiologia* 34:19-34, DOI: 10.1007/BF00024896.
- House, W. A., 2003. Geochemical cycling of phosphorus in rivers. *Appl. Geochem.* 18:739-748, DOI: 10.1016/s0883-2927(02)00158-0.
- Jarvie, H. P., B. A. Whitton and C. Neal, 1998. Nitrogen and phosphorus in east coast British rivers: Speciation, sources and biological significance. *Sci. Total Environ.*

210-211:79-109, DOI: 10.1016/s0048-9697(98)00109-0.

Jordan, T. E., D. L. Correll and D. E. Weller, 1997a. Nonpoint source discharges of nutrients from Piedmont watersheds of Chesapeake Bay. *J. Am. Water Resour. Assoc.* 33:631-645, <http://onlinelibrary.wiley.com/doi/10.1111/j.1752-1688.1997.tb03538.x/abstract>.

Jordan, T. E., D. L. Correll and D. E. Weller, 1997b. Relating nutrient discharges from watersheds to land use and streamflow variability. *Water Resour. Res.* 33:2579-2579, DOI: 10.1029/97wr02005.

Jossette, G., B. Leporcq, N. Sanchez and Philippon, 1999. Biogeochemical mass-balances (C, N, P, Si) in three large reservoirs of the Seine basin (France). *Biogeochemistry* 47:119-146, DOI: 10.1007/bf00994919.

Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C. Stevenson, 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* 303:1-29, DOI: 10.3354/meps303001.

Kirchner, J. W., X. Feng and C. Neal, 2000. Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature* 403:524-527, DOI: 10.1038/35000537.

Kirchner, J. W., X. Feng and C. Neal, 2001. Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations. *Journal of Hydrology* 254:82-101, DOI: 10.1016/s0022-1694(01)00487-5.

- Kirchner, J. W. and C. Neal, 2013. Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc. Natl. Acad. Sci. U. S. A.* 110:12213-12218, DOI: 10.1073/pnas.1304328110.
- Langland, M. J., 2009. Bathymetry and sediment-storage capacity change in three reservoirs on the Lower Susquehanna River, 1996-2008. *Scientific Investigations Report 2009-5110*. US Geological Survey, Reston, VA, p. 21. <http://pubs.usgs.gov/sir/2009/5110/>.
- Langland, M. J. and R. A. Hainly, 1997. Changes in bottom-surface elevations in three reservoirs on the lower Susquehanna River, Pennsylvania and Maryland, following the January 1996 flood - implications for nutrient and sediment loads to Chesapeake Bay. *Water-Resources Investigations Report 97-4138*. US Geological Survey, Lemoyne, PA, p. 34. [pa.water.usgs.gov/reports/wrir97-4138.pdf](http://pubs.usgs.gov/reports/wrir97-4138.pdf).
- Langland, M. J., D. L. Moyer and J. Blomquist, 2007. Changes in streamflow, concentrations, and loads in selected nontidal basins in the Chesapeake Bay Watershed, 1985-2006. U.S. Geological Survey Open File Report 2007-1372, Reston, VA, p. 68. <http://pubs.usgs.gov/of/2007/1372/>.
- Lidén, R., A. Vasilyev, P. Stålnacke, E. Loigu and H. B. Wittgren, 1999. Nitrogen source apportionment—a comparison between a dynamic and a statistical model. *Ecol. Model.* 114:235-250, DOI: 10.1016/s0304-3800(98)00146-x.
- Linker, L. C., R. A. Batiuk, G. W. Shenk and C. F. Cerco, 2013. Development of the Chesapeake Bay Watershed Total Maximum Daily Load Allocation. *J. Am. Water Resour. Assoc.* 49:986-1006, DOI: 10.1111/jawr.12105.
- Lizarraga, J. S., 1997. Estimation and analysis of nutrient and suspended- sediment loads

- at selected sites in the Potomac River Basin, 1993-95. US Geological Survey Water-Resources Investigations Report 97-4154, Baltimore, MD, p. 23.
- Malone, T. C., L. H. Crocker, S. E. Pike and B. W. Wendler, 1988. Influences of river flow on the dynamics of phytoplankton production in a partially stratified estuary. *Mar. Ecol. Prog. Ser.* 48:235-249.
- McDowell, R. W., B. J. F. Biggs, A. N. Sharpley and L. Nguyen, 2004. Connecting phosphorus loss from agricultural landscapes to surface water quality. *Chem. Ecol.* 20:1-40, DOI: 10.1080/02757540310001626092.
- Medeiros, P. R. P., B. A. Knoppers, G. H. Cavalcante and W. F. L. d. Souza, 2011. Changes in nutrient loads (N, P and Si) in the São Francisco estuary after the construction of dams. *Braz. Arch. Biol. Technol.* 54:387-397, DOI: 10.1590/s1516-89132011000200022.
- Murphy, R. R., W. M. Kemp and W. P. Ball, 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* 34:1293-1309, DOI: 10.1007/s12237-011-9413-7.
- Némery, J., J. Garnier and C. Morel, 2005. Phosphorus budget in the Marne Watershed (France): urban vs. diffuse sources, dissolved vs. particulate forms. *Biogeochemistry* 72:35-66, DOI: 10.1007/s10533-004-0078-1.
- Pärn, J., G. Pinay and Ü. Mander, 2012. Indicators of nutrients transport from agricultural catchments under temperate climate: A review. *Ecol. Indicators* 22:4-15, DOI: 10.1016/j.ecolind.2011.10.002.
- Pieterse, N. M., W. Bleuten and S. E. Jørgensen, 2003. Contribution of point sources and diffuse sources to nitrogen and phosphorus loads in lowland river tributaries.

- Journal of Hydrology* 271:213-225, DOI: 10.1016/s0022-1694(02)00350-5.
- Pionke, H. B., W. J. Gburek and A. N. Sharpley, 2000. Critical source area controls on water quality in an agricultural watershed located in the Chesapeake Basin. *Ecol. Eng.* 14:325-335, DOI: 10.1016/S0925-8574(99)00059-2.
- Pritchard, D. W. and J. R. Schubel, 2001. Human influences on physical characteristics of the Chesapeake Bay. *In: Discovering the Chesapeake: the history of an ecosystem*, P. D. Curtin, G. S. Brush and G. W. Fisher (Editors). The Johns Hopkins University Press, Baltimore, MD, pp. 60-82.
- Sharpley, A. N., W. J. Gburek, G. Folmar and H. B. Pionke, 1999. Sources of phosphorus exported from an agricultural watershed in Pennsylvania. *Agric. Water Manage.* 41:77-89, DOI: 10.1016/s0378-3774(99)00018-9.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- Shields, C. A., L. E. Band, N. Law, P. M. Groffman, S. S. Kaushal, K. Savvas, G. T. Fisher and K. T. Belt, 2008. Streamflow distribution of non-point source nitrogen export from urban-rural catchments in the Chesapeake Bay watershed. *Water Resour. Res.* 44, DOI: 10.1029/2007wr006360.
- Sprague, L. A., M. J. Langeland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. U.S. Geological Survey Water-Resources Investigations Report 00-4218, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).

SRBC, 2012. Sediment and nutrient assessment program.

Stow, C. A., M. E. Borsuk and D. W. Stanley, 2001. Long-term changes in watershed nutrient inputs and riverine exports in the Neuse River, North Carolina. *Water Res.* 35:1489-1499, DOI: 10.1016/S0043-1354(00)00402-4.

USEPA, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment.

<http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html>.

USGS, 2012. Chesapeake Bay river input monitoring program.

Withers, P. J. A. and H. P. Jarvie, 2008. Delivery and cycling of phosphorus in rivers: a review. *Sci. Total Environ.* 400:379-395, DOI: 10.1016/j.scitotenv.2008.08.002.

*Page intentionally left blank*

## **Chapter 2. Long-term Seasonal Trends of Nitrogen, Phosphorus, and Suspended Sediment Load from the Non-tidal Susquehanna River Basin to Chesapeake Bay<sup>1</sup>**

### **Abstract**

Reduction of nitrogen (N), phosphorus (P), and suspended sediment (SS) load has been a principal focus of Chesapeake Bay watershed management for decades. To evaluate the progress of management actions in the Bay's largest tributary, the Susquehanna River, we analyzed the long-term seasonal trends of flow-normalized N, P, and SS load over the last two to three decades, both above and below Conowingo Reservoir. Our results indicate that annual and decadal-scale trends of nutrient and sediment load generally followed similar patterns in all four seasons, implying that changes in watershed function and land use had similar impacts on nutrient and sediment load at all times of the year. Above Conowingo Reservoir, the combined loads from the Marietta and Conestoga stations indicate general trends of N, P, and SS reduction in the Susquehanna River basin, which can most likely be attributed to a suite of management actions on point, agricultural, and stormwater sources. In contrast, upward trends of SS and particulate-associated P and N were generally observed below Conowingo Reservoir since the mid-1990s. Our analyses suggest that

---

<sup>1</sup> This chapter (Abstract through Section 2.7) has been published as: Zhang, Q.; Brady, D. C.; Ball, W. P., Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* **2013**, *452-453*, 208-221, doi: [10.1016/j.scitotenv.2013.02.012](https://doi.org/10.1016/j.scitotenv.2013.02.012). Co-author Brady was involved in results interpretation and editing. Co-author Ball was involved in hypothesis development, study design, results interpretation, and editing. Copyright 2013 Elsevier. Reproduced/modified by permission of Elsevier. All figures, tables, and data were created by Qian Zhang unless otherwise indicated. Section, table, and figure numbers have been modified. References have also been re-formatted for consistency among chapters. Some minor substantive changes from the published manuscript have been made and are identified with footnotes.

(1) the reservoir's capacity to trap these materials has been diminishing over the past two to three decades, and especially so for P since the mid-1990s, and that (2) the reservoir has already neared its sediment storage capacity. These changes in reservoir performance will pose significant new kinds of challenges to attainment of total maximum daily load goals for the Susquehanna River basin, and particularly if also accompanied by increases in storm frequency and intensity due to climate change. Accordingly, the reservoir issue may need to be factored into the proper establishment of regulatory load requirements and the development of watershed implementation plans.

## **2.1. Introduction**

Chesapeake Bay has experienced persistent summertime hypoxia in its bottom waters that has been attributed to a combination of anthropogenic nutrient inputs from the watershed (Malone *et al.*, 1988; Boynton and Kemp, 2000; Hagy *et al.*, 2004; Kemp *et al.*, 2005; Murphy *et al.*, 2011) and naturally occurring vertical stratification (Boicourt, 1992; Pritchard and Schubel, 2001; Murphy *et al.*, 2011). On one hand, high nutrient inputs – primarily nitrogen (N) and phosphorus (P) – can stimulate phytoplankton growth that can exert considerable biochemical oxygen demand when the algal matter sinks to the deep channel (Boynton and Kemp, 2000; Cloern, 2001; Kemp *et al.*, 2009). On the other hand, freshwater flow acts to strengthen the water column stratification that can isolate the deep water hypoxic zones, thus preventing oxygen replenishment from the surface water (Goodrich *et al.*, 1987; Boicourt, 1992; Pritchard and Schubel, 2001). In addition, suspended sediment (SS) can reduce light penetration and thus inhibit the growth of beneficial submerged aquatic vegetation (Brakebill *et al.*, 2010). Of the two influences, anthropogenic inputs and stratification, the more attainable means of

controlling hypoxia is the reduction of N, P, and SS load, and this has been a principal focus of Chesapeake Bay watershed (CBW) management for decades. In 2010, the strength of this endeavor was increased, with the introduction of total maximum daily loads (TMDLs) for N, P, and SS (US Environmental Protection Agency, 2010).

To aid the assessment of reduction progress, the U.S. Geological Survey (USGS) and collaborators have been collecting and analyzing water quality data at many monitoring sites in the CBW for decades (Sprague *et al.*, 2000; Langland *et al.*, 2007). For example, the USGS River Input Monitoring (RIM) Program has been monitoring streamflow and water quality at nine stations at the fall-line of major tributaries since the mid-1980s (US Geological Survey, 2012a). In a comprehensive study, (Langland *et al.*, 2007) detected significant decreasing trends in flow-adjusted annual concentration of total N (TN), total P (TP), and SS from 1985 to 2006 at about 74, 68, and 32 percent of the 34 monitoring sites in the CBW, respectively.

As part of on-going efforts to analyze load trends in the major tributaries to Chesapeake Bay, this study focuses on the Susquehanna River because it is the largest tributary in terms of freshwater discharge (60%), TN load (62%), and TP load (34%) (Belval and Sprague, 1999). Encouragingly, McGonigal (2010) detected significantly decreasing trends in flow-adjusted annual concentration of TN, TP, and SS at most monitoring sites in the Susquehanna River basin (SRB) from 1986 to 2009 (Figure 2.1). Consistent with these findings, a later study by Langland *et al.* (2012a) also reported generally decreasing trends of flow-adjusted annual concentration of these pollutants in the SRB from 1985 to 2010. In addition, Langland *et al.* (2012a) reported improving conditions of TN and SS, but non-significant changes in TP at the Conowingo station at

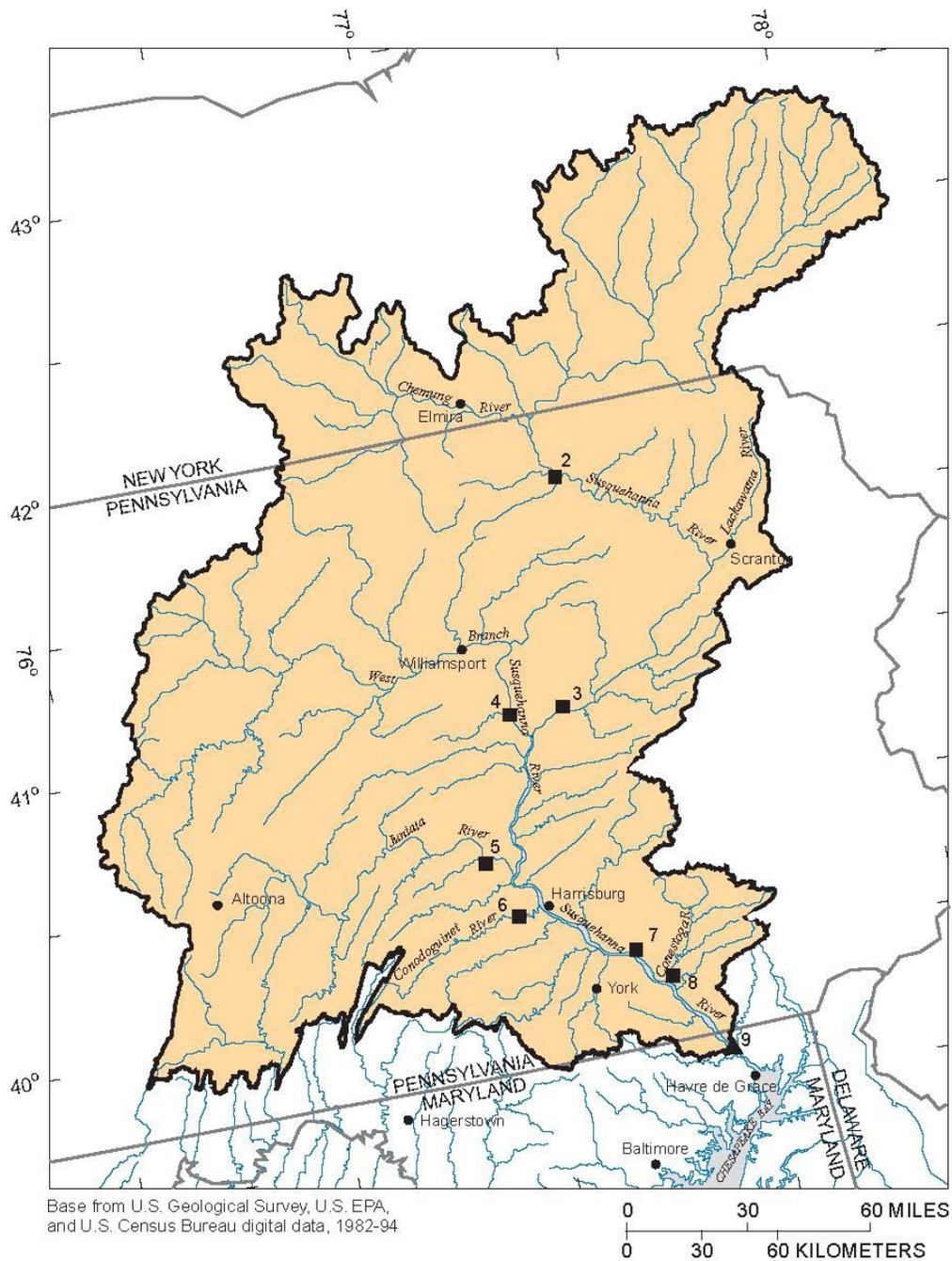


Figure 2.1. Map of the Susquehanna River basin and long-term monitoring sites. Sites in this study include two main-stem stations, Marietta (No. 7) and Conowingo (No. 9), and one tributary station, Conestoga (No. 8). The non-tidal Susquehanna River basin, shaded in yellow, covers portions of New York, Pennsylvania, and Maryland. This figure was reproduced from Figure 6 in Sprague *et al.* (2000) with permission, with the original numbering of the monitoring sites.

the fall-line of the Susquehanna River. A common feature of these studies is the adoption of annual resolution in trend analyses. However, in order to capture impacts of seasonality such as variations in temperature and rainfall, fertilizer application, and benthic recycling of P and denitrification in river channels and reservoirs, seasonal trends of these anthropogenic pollutants need to be investigated, and attention needs to be given to calculations not only of concentration but also of load, which is more complex because of the need to temporally match concentration with flow.

In regard to prior estimates of historical load, the USGS has been using a tool developed in 1989 called ESTIMATOR (Cohn *et al.*, 1989) to compute and report daily nutrient and sediment load in Chesapeake Bay tributaries. In recent work, Murphy *et al.* (2011) combined some of these loading estimates for the Susquehanna River at Conowingo with some interpolations and extrapolations of upstream data that had been previously developed by Hagy *et al.* (2004) to plot a 60-year history of Susquehanna winter-spring (*i.e.*, Jan-May) TN load, which has been of special interest because of its reported correlation with summer-time oxygen depletion in Chesapeake Bay (Hagy *et al.*, 2004). The 60-year TN load history (*cf.* Figure 3c of Murphy *et al.* (2011)), although replete with flow-related inter-annual variations, suggested a general decreasing trend of Jan-May TN load in recent decades. In addition, Murphy *et al.* (2011) has reported correlation between the Jan-May TN loads and bay hypoxia during the late-summer period. (By contrast, long-term trends in early-summer hypoxia were found to be correlated with some long-term flow-unrelated trends of increasing stratification.) Because the reported correlation is dependent on good understanding of the TN load history, it is important that such trends be evaluated using the best tools available.

When considering overall riverine inputs of nutrient and sediment, it is also important to recognize that such loadings can be greatly complicated by impacts of sediment retention and release in reservoirs. In particular, reservoirs in early stages of their lifespan can effectively retain sediment and associated N and P (Jossette *et al.*, 1999; Friedl and Wüest, 2002; Medeiros *et al.*, 2011), thus providing efficient removal of N and P from streamflow, mainly through denitrification and particle deposition, respectively (Jossette *et al.*, 1999; Friedl and Wüest, 2002). In the SRB, the most studied reservoir system is probably the Lower Susquehanna River Reservoir System, which includes Lake Clarke, Lake Aldred, and Conowingo Reservoir (Figure 2.2) (Langland and Hainly, 1997; Langland, 2009). The system has been reported to trap about 2, 45, and 70 percent of annual TN, TP, and SS load, respectively, from the Susquehanna River to the bay (Langland and Hainly, 1997), thus alleviating the pollutant load considerably. As the most downstream and the largest reservoir in the system, Conowingo Reservoir (hereafter, “the reservoir”) is the only one that was reported not to have reached its sediment storage capacity (SSC) (Langland, 2009). Based on assumptions in SS input load and sediment deposition rate in the reservoir, Langland (2009) estimated an additional service life of 12-17 years (from 2012) before the reservoir would be filled up. By that time, the annual load of TN, TP, and SS to the bay has been projected to increase by 2, 70, and 250 percent, respectively (Langland and Hainly, 1997). In this context, previous study has suggested that a flow of 400,000 ft<sup>3</sup>/s (11,300 m<sup>3</sup>/s) was the “scour threshold” for Conowingo Reservoir, and that major floods above this level would further increase the nutrient and sediment load delivery to Chesapeake Bay (Langland and Hainly, 1997). More recently, Hirsch (2012) has analyzed discharge and water-

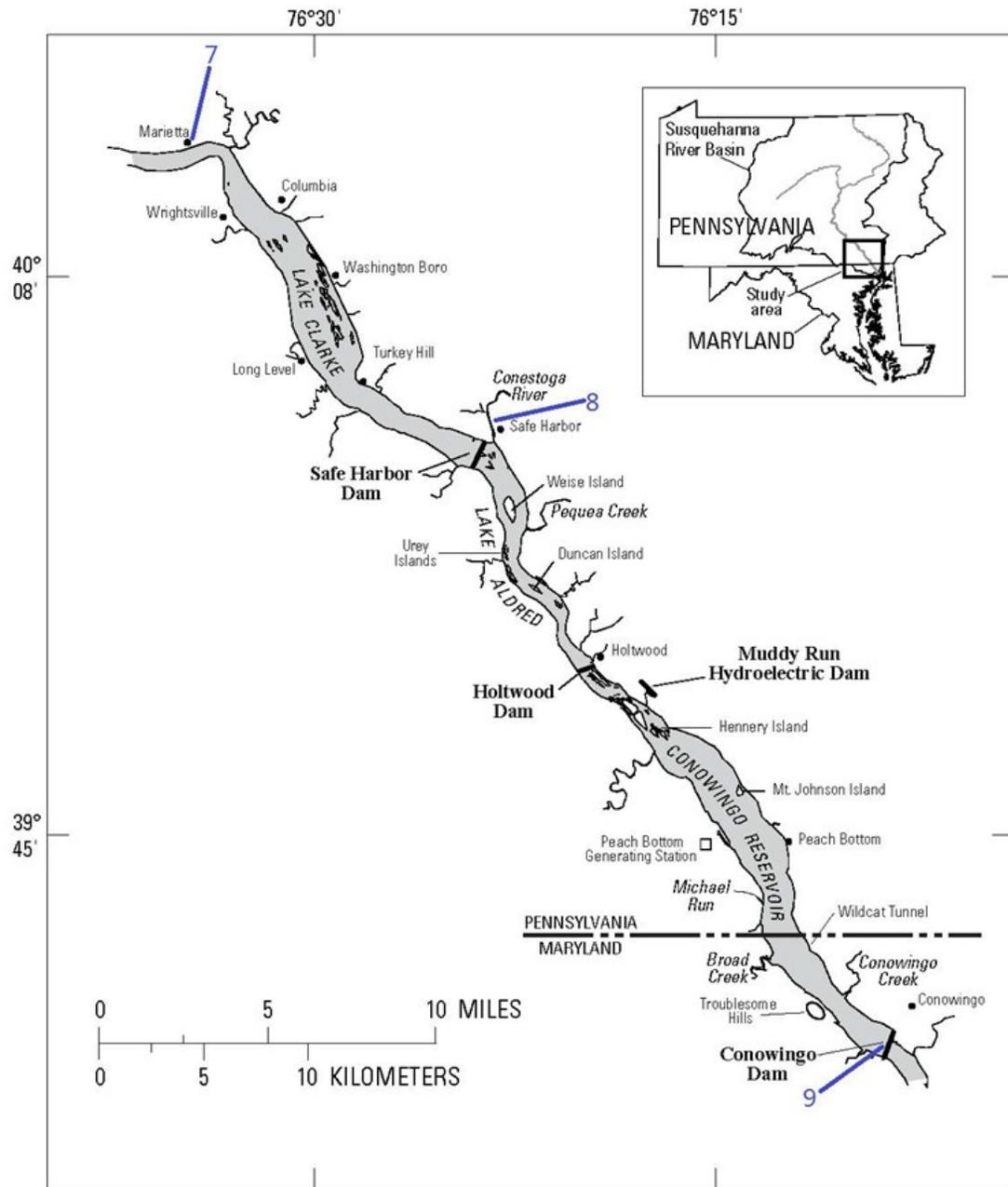


Figure 2.2. Map of Lower Susquehanna River Reservoir System and the study sites. The reservoir system consists of Lake Clarke, Lake Aldred, and Conowingo Reservoir. The Marietta station (No. 7) is just above the reservoirs and the Conowingo station (No. 9) is just below the reservoirs. The Conestoga station (No. 8) monitors streamflow from the Conestoga River, a major tributary to the Susquehanna River. This figure was adapted from Figure 1 in Langland (2009) with permission. See Figure 2.1 for locations of the three sites in the non-tidal Susquehanna River basin.

quality data at the Conowingo Dam, and detected upward trends in annual TP and SS load. The author has evaluated some hypotheses and presented a “scour hypothesis” as a most likely explanation to the observed trends. This hypothesis states that when reservoirs are near capacity, the water channel would become smaller, thus resulting in faster water flow, and correspondingly, higher likelihood of scour (Hirsch, 2012a). Therefore, increased net scouring of sediment in Conowingo Reservoir may already be occurring at flow rates much lower than the above-cited scour threshold (Blankenship, 2012). Hirsch (2012a) has suggested that filling processes in Conowingo Reservoir are already approaching a final asymptotic stage and that further monitoring and evaluation of the reservoir performance are critical to evaluation of management plans.

In the above context, the work described herein was undertaken concurrently with the efforts of Hirsch (2012a) and with the goal of more closely examining many issues raised in that work and similarly motivated by prior studies. Particular new contributions of this work are as follows:

- (1) analysis of multiple N and P species to examine potential differences as related to particulate versus dissolved fractions;
- (2) direct comparison of above- and below- reservoir data at multiple sites to provide a mass balance analysis across the reservoir;
- (3) application of a new and reportedly more accurate loading estimation method called “weighted regressions on time, discharge, and season (WRTDS) (Hirsch *et al.*, 2010)” to develop the first set of estimates for upstream sites;
- (4) analysis of flow-normalized trends to better understand long-term trends independent of random streamflow. The WRTDS method is capable of providing

both “true-condition” and “flow-normalized” estimates of concentration and load.

In comparison, most previous studies (*e.g.*, Langland *et al.*, 2007; McGonigal, 2010; Langland *et al.*, 2012) have used the ESTIMATOR model, which is not able to produce the flow-normalized trends;

- (5) analysis of seasonal loads and trends using both the “true-condition” and “flow-normalized” approaches. Most previous studies of SRB sites (*e.g.*, Langland *et al.*, 2007; Hirsch *et al.*, 2010; McGonigal, 2010; Hirsch, 2012a; Langland *et al.*, 2012) have focused on analyses of annual load and trend.

More specifically, we have examined the long-term seasonal history of N, P, and SS loads in the Susquehanna River, both above and below Conowingo Reservoir, through the following three broad types of analysis:

- (1) reconstruction of our best understanding of the long-term history of concentrations and loads of nitrate (NO<sub>x</sub>) and TN from the Susquehanna River to Chesapeake Bay for a 67-year period (1945-2011), in terms of both “true-condition” and “flow-normalized” estimates, using the latest available method (WRTDS) and the longest available records of concentration and flow at the Conowingo station;
- (2) estimation of seasonal “flow-normalized” loads for multiple species of nutrients and sediment for two locations just above the reservoir system for a 26-year period (1986-2011) over which relevant concentration and flow data are available – major species studied include SS, TP, particulate P (PP), dissolved P (DP), TN, particulate N (PN), and dissolved N (DN); and

- (3) similar estimation of seasonal “flow-normalized” loads for the same species of nutrients and sediment at the Conowingo station for a 34-year period (1978-2012) over which relevant data are available.

Our objectives in undertaking these analyses were as follows:

- (1) to compare long-term trends in N loading with prior estimates published by Murphy *et al.* (2011) and to thus verify whether the long-term trends are still apparent; and
- (2) to evaluate progress in reduction of N, P, and SS load from the non-tidal SRB at seasonal resolution; and
- (3) to compare the relative changes in N, P, and SS loads discharging into and emanating from the reservoir at seasonal resolution, thus allowing an evaluation of reservoir performance and service life in terms of sediment and nutrient retention.

## **2.2. Methods**

### **2.2.1. Study Sites**

The RIM station at the Conowingo Dam is about 10 miles from the Susquehanna River mouth and receives 99% of the streamflow from the SRB (Belval and Sprague, 1999). This station is also located at the river fall-line, a physical fall that provides distinct separation of the tidal and non-tidal basins. Upstream, six additional sites at Towanda, Danville, Lewisburg, Newport, Marietta, and Conestoga have been monitored by the Susquehanna River Basin Commission (SRBC) through the Susquehanna Nutrient Assessment Program (SNAP) since the mid-1980s (Susquehanna River Basin

Commission, 2012). Since all the stations are above the fall-line (*i.e.*, not influenced by tides), trends observed there can be used to assess nutrient and sediment reduction progress in their respective upstream watersheds within the SRB (Sprague *et al.*, 2000).

Sites examined in the present study include the RIM station at Conowingo and two SNAP stations at Marietta and Conestoga. The Marietta station is the most downstream SNAP station on the river mainstem and represents the vast majority (~96%) of the watershed area represented by the Conowingo station and with a median streamflow that is slightly higher (Table 2.1). However, one major distinction between the two stations is their location relative to the Lower Susquehanna River Reservoir System – Marietta is upstream and Conowingo is downstream of the reservoirs (Figure 2.2). The Conestoga station on the Conestoga River (a major tributary to the Susquehanna located between Marietta and Conowingo; Figure 2.2), monitors surface runoff from the small but heavily agricultural Conestoga basin (Table 2.1). In general, the combined nutrient and sediment load from the Marietta and Conestoga stations represents a majority of input to the reservoirs, whereas load at Conowingo represents the output. Comparisons between the input and output are thus well suited for examining the possible impacts of the reservoirs on long-term seasonal trends of nutrient and sediment loads.

Table 2.1. Details of the study sites. <sup>a</sup>

USGS ID	Station name	Upstream land area (mi <sup>2</sup> )	Upstream land use (percent)				Flow statistics <sup>b</sup> (cubic feet per second)		
			Urban	Agricultural	Forested	Other	Min	Median	Max
01576000	Susquehanna River at Marietta, PA	25,990	4	30	64	2	24,370	36,280	63,560
01576754	Conestoga River at Conestoga, PA	470	8	54	37	1	217	664	1140
01578310	Susquehanna River near Conowingo, MD	27,100	2	29	67	2	23,560	35,575	65,540

<sup>a</sup> modified from Table 8 in Sprague *et al.* (2000) with permission

<sup>b</sup> calculated based on annual average flow data from 1985 to 2010 (US Geological Survey, 2012b)

### 2.2.2. Statistical Methods

Because of concomitant constraints on labor, time, and funding, water-quality samples have been collected only once or several times each month at the study sites. Therefore, appropriate statistical methods are required to make predictions for unsampled days. Selection of best methods of estimation for nutrient and sediment concentrations and loads based on available monitoring data has been an important topic of discussion since at least the late 1980s (Cohn *et al.*, 1989). Prior to 2015,<sup>2</sup> the USGS has applied an estimation tool known as the ESTIMATOR model (Cohn *et al.*, 1989) to estimate daily nutrient and sediment concentration and load in Chesapeake Bay tributaries. More recently, however, Hirsch *et al.* (2010) have described the need for new statistical methods that can both (a) better describe temporal variations in concentration and load, and (b) more effectively remove the influence of random flow variation. More importantly, the new methods should not rely on questionable assumptions such as a constant concentration-flow relationship, constant seasonal trends in concentration, and the existence of specific functional forms of these trends (Hirsch *et al.*, 2010). Hirsch *et al.* (2010) incorporated these considerations into the development of the WRTDS method, which has recently been applied to several large data sets and which has been fully described elsewhere, *e.g.*, Hirsch *et al.* (2010), Sprague *et al.* (2011). For the convenience of readers here, I have briefly summarized the basic structure and application of WRTDS in Appendix A1.

---

<sup>2</sup> At the time of our 2013 publication (Zhang *et al.*, 2013), the ESTIMATOR method was still in use. The underlined text is an updated statement.

The WRTDS method produces two types of estimates for both concentration and load – so-called “true-condition” and “flow-normalized” estimates, as described in more detail in Appendix A1. Hirsch *et al.* (2010) have pointed out that the true-condition estimates are useful to help understand the real history of riverine nutrient (or sediment) and downstream ecological impact, whereas the flow-normalized estimates are more helpful to evaluate management progress in the watershed – *i.e.*, for identifying long-term trends without confounding complications associated with inter-annual flow variability.<sup>3</sup> The flow-normalization algorithm, described in more detail in Appendix A1, can greatly remove the sometimes dramatic influence of random variations in streamflow by linking the estimation to the full history of hydrological flows over long-term cycles, thus rendering longer-term inter-annual trends easier to detect and understand than they would be with true-condition estimates. For these reasons, we have focused most of our attention in this work to analyses of flow-normalized load.

One major assumption of the flow-normalization method is the stationarity of streamflow time series during the study period, as more fully discussed elsewhere, *e.g.*, Hirsch *et al.* (2010), Sprague *et al.* (2011). In this regard, one should be aware that flow-normalized estimates can potentially be misleading if stationarity is violated – that is, if the probability distribution of streamflow on a given day of the year has changed significantly over time, and if such change has been able to exert substantial impacts on the relations between flow and water quality (Hirsch *et al.*, 2010). In the mid-Atlantic region where Chesapeake Bay is located, for example, one might have concerns that

---

<sup>3</sup> Relative to our published paper (Zhang *et al.*, 2013), the underlined text is a more accurate statement of the intended concept.

watershed development has altered the “flashiness” of streamflow (Jarnagin, 2007), and that these changes, if they exist, could challenge the validity of the stationarity assumption. At present, however, we have no means to further explore this issue. According to Hirsch *et al.* (2010), there is currently no formal procedure to defend or reject the appropriateness of flow-normalization algorithm in light of nonstationarity of streamflow,<sup>4</sup> and this is an area where future research is needed to improve WRTDS (Hirsch *et al.*, 2010).

Another issue to consider for any given application of WRTDS is the selection of “half-window widths” for the estimation process, as described in Appendix A1. In this study, the half-window widths were defined as 10 years and 0.5 years for time and season, respectively. For the discharge dimension, the window was selected such that, for a given discharge  $Q$  (as reference), positive weights would be assigned only to discharges falling between  $Q_{\min} = Q/\exp(2)$  and  $Q_{\max} = Q\exp(2)$ , or in other words,  $\ln(Q_{\max}/Q) = \ln(Q/Q_{\min}) = 2$ . In their analyses of Chesapeake Bay tributaries including Susquehanna, Hirsch *et al.* (2010) used the above half-window widths and considered them as appropriate based on testing. We agreed with this assessment – preliminary independent analysis with our own data *set also* suggested that this subjective choice would have little impact on model estimates and load trends, so long as reasonable values are assumed, within ranges suggested by Hirsch *et al.* (2010).

### 2.2.3. Data Compilation and Analyses

---

<sup>4</sup> Relative to our published paper (Zhang *et al.*, 2013), the underlined text is a more accurate statement of the intended concept.

We collected streamflow and water-quality data at Conowingo (1978-2011) from the USGS National Water Information System Web Interface (USGS-01578310; US Geological Survey, 2012b), and at Marietta (1986-2011) and Conestoga (1984-2011) from the SRBC SNAP website (Susquehanna River Basin Commission, 2012). The collected water quality data included information for eight nutrient and sediment constituents, namely, SS, TP, DP, TN, DN, dissolved orthophosphate (DOP), dissolved nitrate plus nitrite ( $\text{DNO}_x$ ), and dissolved ammonia plus organic N (DKN). We implemented the WRTDS method using the statistical package R (R Development Core Team, 2011) to produce the flow-normalized estimates for every day in the period of record for each species. The flow-normalized daily estimates of load in units of  $\text{kg day}^{-1}$  were used to calculate the seasonal averages of load for each of the four seasons, defined as January-March, April-June, July-September, and October-December, respectively. In addition, since there was no measurement of PP or PN, their seasonal loads were inferred by subtracting DP and DN from TP and TN seasonal loads, respectively. Thus, the signals of particulate and dissolved fractions could be separated. Similarly, dissolved hydrolysable P (DHP), or the “non-labile” fraction of DP (refer Table 1 in Neal *et al.* (2010) for terminology), was inferred by subtracting DOP from DP. In contrast to DP, individually measured data were directly available for DN and  $\text{DNO}_x$  to the present date and for DKN up to May 1995, after which DKN concentration in water samples has been reported as the difference between measured DN and  $\text{DNO}_x$  concentrations. For each of the four seasons studied, we observed that the DN loads estimated using WRTDS on measured DN data fell between 95% and 105% of the values (for the same season and location) that were calculated from the sum of estimated  $\text{DNO}_x$  load plus DKN load.

To reconstruct the 67-year history of NO<sub>x</sub> concentration and load at the Conowingo station (1945-2011), our first step was to close the data gaps in streamflow discharge (1945-1968) and NO<sub>x</sub> concentration (1945-1978) based on upstream data at Harrisburg (USGS-01570500; US Geological Survey, 2012b; see Figure 2.1 for location). We first compiled the streamflow data at Harrisburg from 1945 to 1968 and converted them to Conowingo flow data using the ratio reported by Hagy *et al.* (2004) (*i.e.*, Conowingo flow = 10/9 x Harrisburg flow). We then compiled the NO<sub>x</sub> concentration data at Harrisburg from 1945 to 1978 and converted them to NO<sub>x</sub> concentration at Conowingo using monthly ratios reported by Hagy *et al.* (2004). These manipulated records, together with observational data (1968-2011 for flow; 1978 to 2011 for NO<sub>x</sub> concentration), constituted the 67-year full records at Conowingo (Figure 2.3a-d). On that basis, we estimated the true-condition and the flow-normalized estimates for NO<sub>x</sub> at Conowingo from 1945 to 2011.

In addition, to reconstruct the 67-year history of true-condition TN load at Conowingo (1945-2011), we needed to convert the pre-1978 true-condition NO<sub>x</sub> load to TN load, due to lack of TN concentration data for that period. We first developed linear regression models relating monthly TN to NO<sub>x</sub> load at Conowingo for each month of the year based on available TN and NO<sub>x</sub> load estimates from 1981 to 2010 (Table A1 in Appendix A2). We then used these linear models to convert the pre-1978 true-condition NO<sub>x</sub> load (described above) to true-condition TN load in respective months. These monthly TN loads were then combined with directly estimated (post-1978) monthly TN loads to reconstruct the 67-year monthly TN loads from 1945 to 2011. Finally, the Jan-

May true-condition TN loads at Conowingo were obtained by averaging the monthly TN loads from January to May in each year.

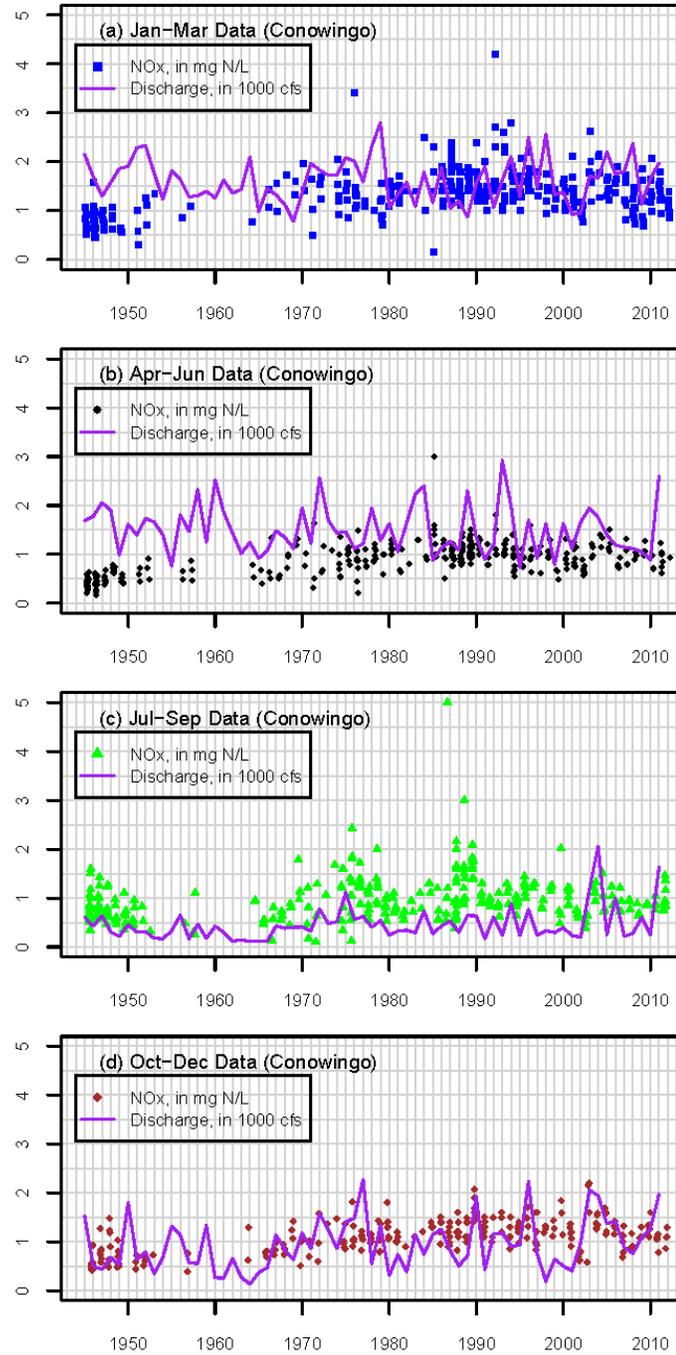


Figure 2.3. Observed data of NO<sub>x</sub> concentration and seasonal streamflow discharge in (a) Jan-Mar, (b) Apr-Jun, (c) Jul-Sep, and (d) Oct-Dec, in the Susquehanna River at the Conowingo station for the period from 1945 to 2011.

## **2.3. Results and Discussion**

In the sections below, I present our results in regard to the three major sets of tasks and objectives identified in Section 2.1. First, I present the 67-year analysis of NO<sub>x</sub> and TN trends at Conowingo in terms of both “true-condition” and “flow-normalized” results. Second, we analyze the combined loads of SS, P, and N from the Marietta and Conestoga stations to evaluate progress of management actions in the non-tidal SRB above the reservoir. This watershed covers portions of New York (NY), Pennsylvania (PA), and Maryland (MD). Finally, I present and discuss the seasonal trends of SS, P, and N load at the Conowingo station to examine the evolving behavior of the reservoir in modulating sediment and nutrient load at seasonal resolution.

### **2.3.1. History of NO<sub>x</sub> and TN Load at the Conowingo Station (1945-2011)**

#### ***2.3.1.1. Results***

Our retrospective analyses of the 67-year record of Susquehanna River NO<sub>x</sub> concentration and load are presented in Figure 2.4. In regard to the concentration results, the true-condition (Figure 2.4a) and the flow-normalized (Figure 2.4b) estimates both show similar annual- and decadal-scale trends among all four seasons, with a steady rise from 1945 to around 1990, followed by a steady decline. In regard to the load results, trends in the true-condition loads (Figure 2.4c) are difficult to discern, owing to the high degree of inter-annual variability. Removal of this influence is in fact a primary motivation for the consideration of the flow normalized loads (Figure 2.4d), which show more clear trends (see “Discussion” below). In general, trends in the flow-normalized loads are similar to the trends in the flow-normalized estimates of concentration.

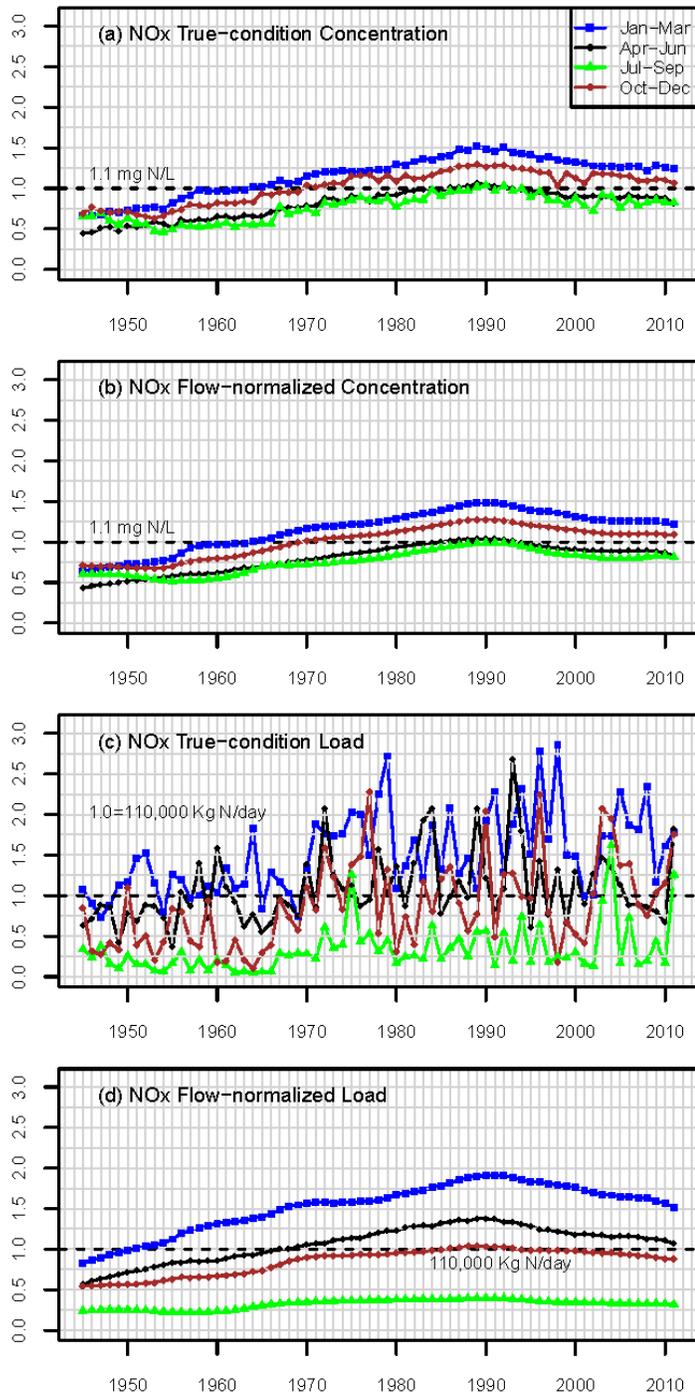


Figure 2.4. Seasonal averages of NO<sub>x</sub> (a) true-condition concentration, (b) flow-normalized concentration, (c) true-condition load, and (d) flow-normalized load in the Susquehanna River at the Conowingo station. All estimates have been normalized by the median of respective long-term annual averages at the Conowingo station (locating at y = 1.0 in each panel).

As previously explained, the lack of raw data for TN concentration prevents a similar calculation of flow-normalized concentration or load. For this constituent, our interest is primarily in the “true-condition” load estimates during the period of Jan-May, for purposes of comparison with values used in earlier analyses by Murphy *et al.* (2011). These results are shown in Figure 2.5, together with the values from Murphy *et al.* (2011) and also differences between the two sets of loads. In general, results using either method exhibit a similar long-term trend, with generally much lower peak loads prior to 1970, increased variability since about 1980, and a general trend of stabilized or decreasing loads since that time.

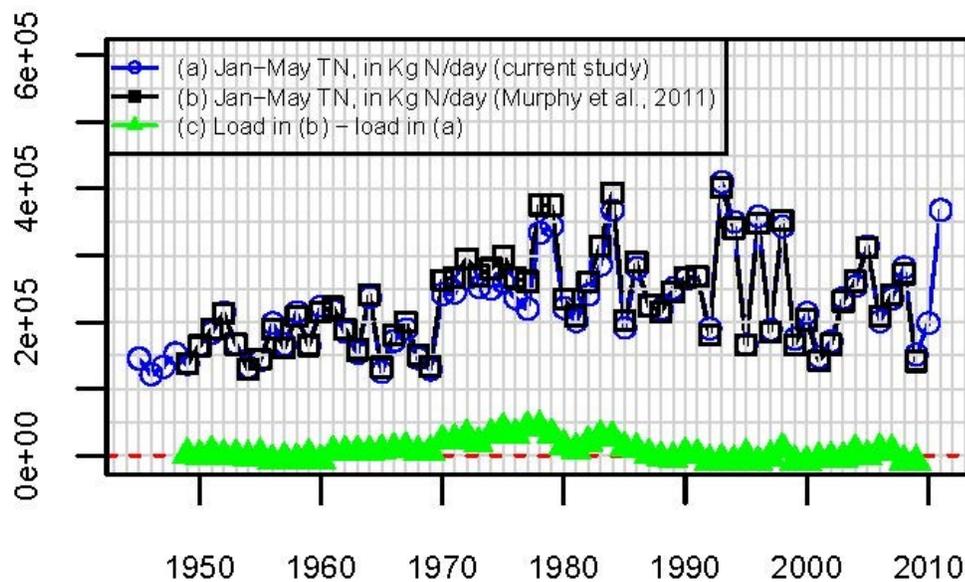


Figure 2.5. Estimates of “true-condition” Jan-May TN load in the Susquehanna River at the Conowingo station for the period 1945 to 2011. Plot (a) shows TN loads obtained in this study. Plot (b) shows TN loads reported by Murphy *et al.* (2011) for the period 1949 to 2009. For that study, TN loads prior to 1980 were obtained using regression equations between TN and NO<sub>x</sub> load developed by Hagy *et al.* (2004), and TN loads from 1981 to 2009 were directly obtained from the USGS RIM Program website (US Geological Survey, 2012a). Plot (c) shows differences in estimates between (a) and (b).

### **2.3.1.2. Discussion**

The NO<sub>x</sub> results presented in Figure 2.4 show similar trends in all four seasons. More generally, however, such consistency in seasonal trends should not necessarily be expected. In fact, our on-going study of trends in other tributaries of Chesapeake Bay has revealed substantially different trends in some cases. For example, from the late 1970s to the early 1990s, the flow-normalized NO<sub>x</sub> concentrations in the Potomac River show a trend of slight decline in Jan-Mar, but strong upward trends in the other three seasons (data not shown).

Perhaps the most important point to observe from Figure 2.4 is the manner in which the flow-normalized NO<sub>x</sub> loads (Figure 2.4d) remove the effects of the highly variable streamflow during each season (shown as solid lines in Figure 2.3). In this regard, Figure 2.4d reveals a smooth trend in load change that is similar in its basic aspects to the trends in estimated concentration (Figure 2.4a and 4b) and devoid of the flow-induced variations evident in Figure 2.4c. From this example, we can see that the flow-normalized loads are more helpful to evaluate progress of management actions in the watershed. Differences between trends in flow-normalized loadings and in flow-normalized concentration are presumably the result of flow influences on concentration and resulting effects on the regressions that account for flow. Although the exact meanings of these differences are complex to understand, it has been suggested that trends in flow-normalized concentration are more representative of changes in point sources, which are presumed to be less heavily influenced by flow (Hirsch *et al.*, 2010).

As noted in Section 2.1, the 67-year true-condition TN load history at Conowingo was reconstructed using the WRTDS method in order to verify whether Jan-May trends

reported by Murphy *et al.* (2011) would still be observed. Our results (Figure 2.5) confirm that true-condition estimates of TN loading with WRTDS are generally similar to those previously assumed by Murphy *et al.* (2011). Overall, the WRTDS estimates range between 0.84 and 1.09 of the previously reported TN load values, with largest differences occurring in the 1970s.

### 2.3.2. History of SS, P, and N Load from the Marietta and Conestoga Stations (1986-2011)

#### 2.3.2.1. Results

The combined flow-normalized SS loads from Marietta and Conestoga show consistently downward trends in all four seasons (Figure 2.6), with decelerated reduction or slight rise between the mid-1990s to the mid-2000s.

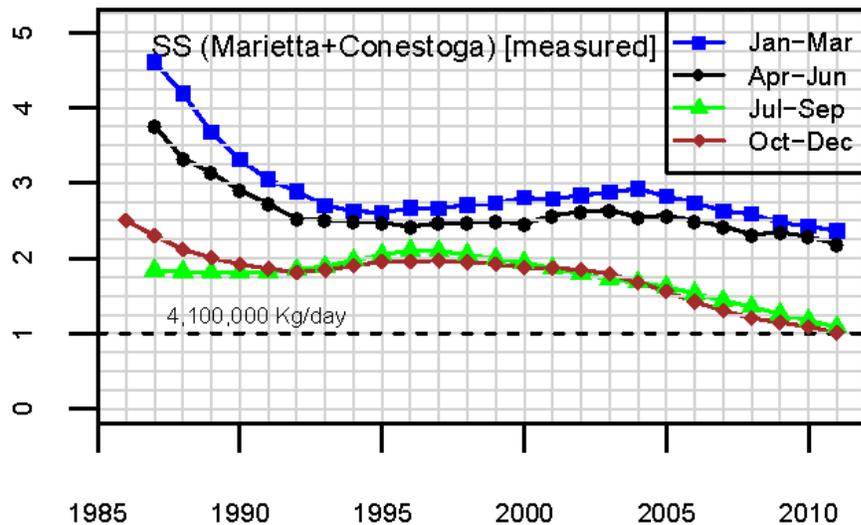


Figure 2.6. Seasonal averages of flow-normalized SS load from the Marietta and Conestoga stations. All loads have been normalized by the median of long-term annual SS loads at the Conowingo station (locating at  $y = 1.0$ ).

The combined flow-normalized TP loads from Marietta and Conestoga also show downward trends in all four seasons, with decelerated reduction from the mid-1990s to the early 2000s (Figure 2.7a). Within TP, PP contributed to the TP reduction only since around 2000 (Figure 2.7b), whereas DP contributed to the TP reduction throughout the study period (Figure 2.7c). Within DP, DOP increased consistently until 2002 and started to contribute to the DP reduction thereafter (Figure 2.7d), whereas DHP decreased substantially in the earlier period (up until 2002) and has remained low since that time (Figure 2.7e).

The combined flow-normalized TN loads from Marietta and Conestoga also show consistently downward trends in all four seasons (Figure 2.8a). Within TN, PN decreased rapidly until the late 1990s, and thereafter showed slight rise in Jan-Mar and Apr-Jun but continual reduction in Jul-Sep and Oct-Dec (Figure 2.8b). DN shows downward trends in all four seasons (Figure 2.8c), similar to the TN trends previously noted. Both DN and PN contributed to the TN reduction until the late 1990s, thereafter primarily DN contributed to the TN reduction. The two fractions of DN, *i.e.*, DNO<sub>x</sub> (Figure 2.8d) and DKN (Figure 2.8e), were estimated separately using available data and both show downward trends throughout the study period.

#### **2.3.2.2. Discussion**

The flow-normalized SS, P (TP, PP, DP), and N (TN, PN, DN) loads from Marietta and Conestoga all show downward trends from 1986 to 2011, suggesting that management controls have been effective in reducing watershed inputs of these pollutants in the non-tidal SRB above the reservoir.

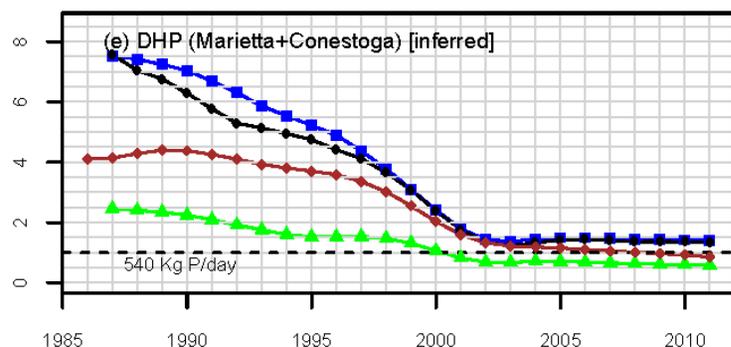
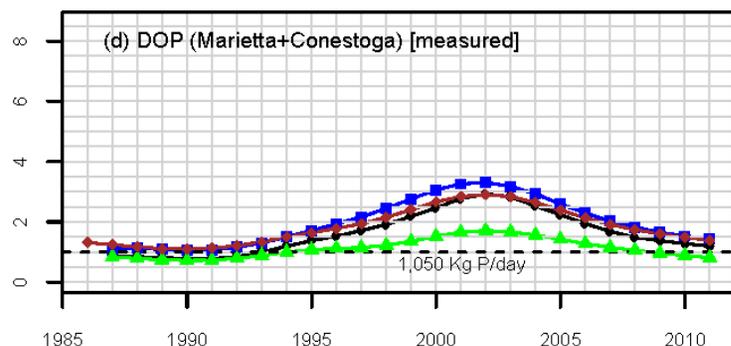
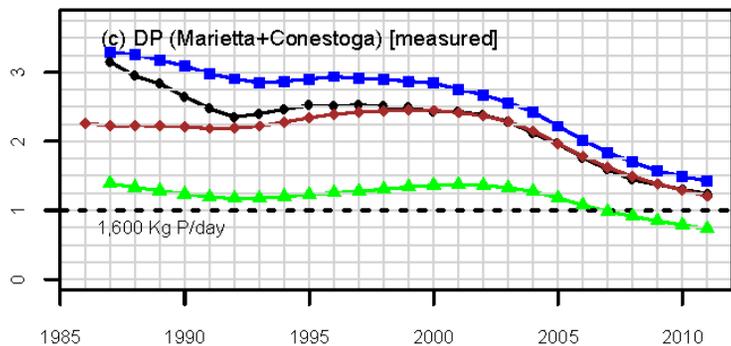
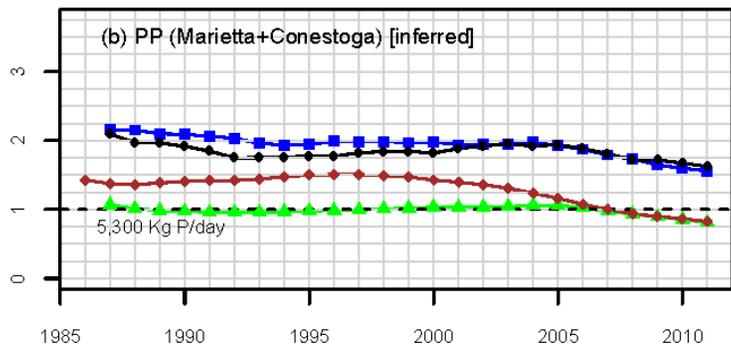
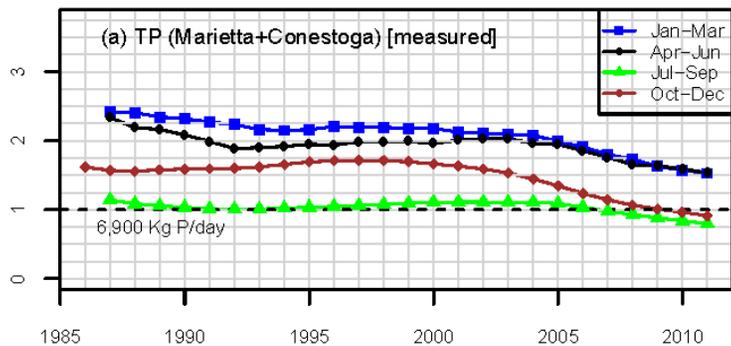


Figure 2.7. Seasonal averages of flow-normalized load of (a) TP, (b) PP, (c) DP, (d) DOP (dissolved orthophosphate), and (e) DHP (dissolved hydrolysable P) from the Marietta and Conestoga stations. All loads have been normalized by the median of respective long-term annual loads at the Conowingo station at the reservoir outlet (locating at  $y = 1.0$  in each panel).

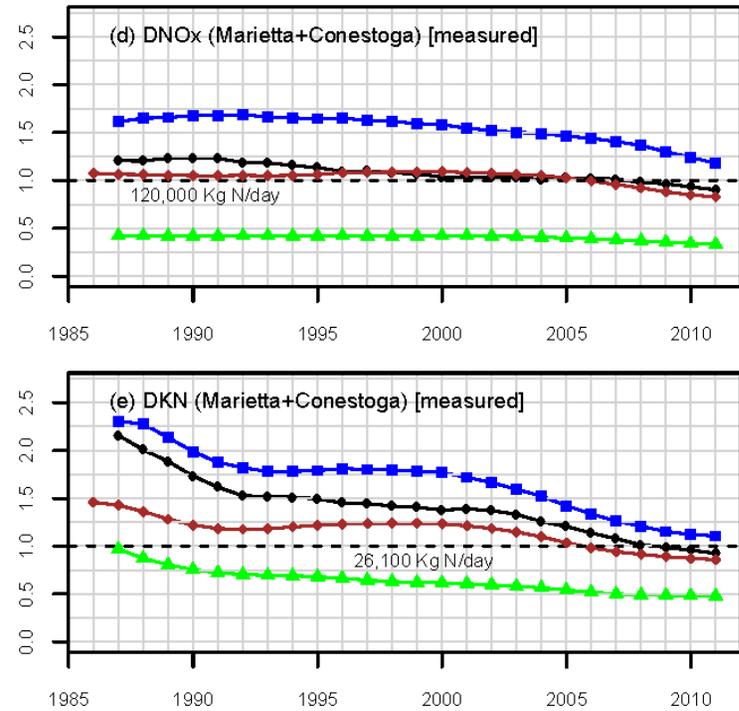
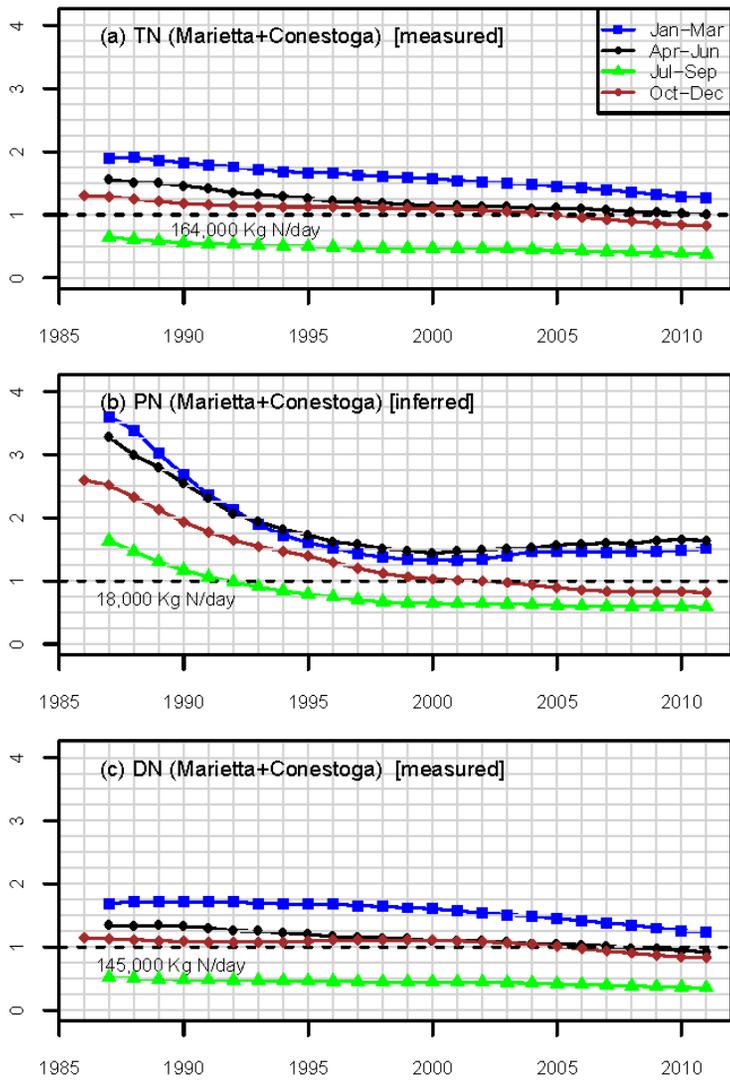


Figure 2.8. Seasonal averages of flow-normalized load of (a) TN, (b) PN, (c) DN, (d) DNO<sub>x</sub>, and (e) DKN from the Marietta and Conestoga stations. All loads have been normalized by the median of respective long-term annual loads at the Conowingo station at the reservoir outlet (locating at  $y=1.0$  in each panel).

To more fully understand the above-noted “positive progress,” it is useful to review factors affecting the source, transport, and transformation of nutrient and sediment in the SRB. In terms of source, Ator *et al.* (2011) reported that mean annual TN loads from the non-tidal SRB, calculated for the year 2002 by using statistical representations of long-term mean hydrological conditions, originated mainly from non-point source inputs of fertilizer, fixation, and manure (58%), followed by atmospheric deposition (20%), urban sources (12%), and point sources (10%). Using the same method, TP loads were reported to have originated from point sources (39%) as well as from fertilizer and manure (32%), followed by erosion of rocks (22%) and urban sources (7%). In addition, Brakebill *et al.* (2010) determined agricultural land as the greatest overall source and urban development as the source with highest yield (load per unit area) for SS in the CBW. Once the N and P are generated from these sources, they can be temporarily stored in the system (*e.g.*, land surface, riparian buffer, river channels, reservoirs, *etc*), transformed chemically or biologically (*e.g.*, plant uptake, mineralization and denitrification of N, precipitation of P, *etc*), or transported downstream (Brakebill *et al.*, 2010; Ator *et al.*, 2011). SS, however, exhibits a more conservative behavior since it cannot be readily transformed (Brakebill *et al.*, 2010).

In the last few decades, a suite of management practices have been implemented to control N, P, and SS load in the SRB, some focusing on reduction by controlling pollutant transport or transformation, but more focusing on control at the pollutant source. An overview of source-based historical management strategies in the SRB is provided in Appendix A3. Overall, it is likely that the source-based management strategies and associated controls on transport and transformation processes, were

responsible for the downward nutrient and sediment trends in the SRB at locations above the reservoir. Indeed, Brakebill *et al.* (2010) have suggested that effective SS control measures should include both source reduction (*e.g.* settlement ponds, soil conservation practices, riparian buffers, *etc*) and streambank protection (*e.g.* directing erosive flow, flood-plain stabilization, *etc*). However, identification of the extent of implication and relative contribution of these different management actions is well beyond the scope of our current study.

### **2.3.3. History of SS, P, and N Load at the Conowingo Station (1978-2011)**

#### **2.3.3.1. Results**

The flow-normalized SS loads at Conowingo show generally “fall-and-then-rise” trends in all four seasons (Figure 2.9). In Jan-Mar, Apr-Jun, and Jul-Sep, SS load generally decreased until around 1990, stabilized for about one decade, and has then increased rapidly since the late 1990s. In Oct-Dec, SS load displayed much weaker variation. Overall, the SS load at Conowingo has been digressing increasingly far from the TMDL goal in recent years in all seasons.

The flow-normalized TP loads at Conowingo show very similar “fall-and-then-rise” trends in all four seasons (Figure 2.10a), closely following the SS trend. Overall, the TP load at Conowingo has also digressed increasingly far from the TMDL goal. The effect is clearly related to particulate species – PP shows the same “fall-and-then-rise” trend (Figure 2.10b), whereas DP shows downward trends in all four seasons (Figure 2.10c). Both DP and PP contributed to the TP reduction until the mid-1990s, and PP alone contributed to the TP rise thereafter. Within DP, DOP shows downward trends particularly in the early period (Figure 2.10d), whereas DHP shows bi-modal patterns in

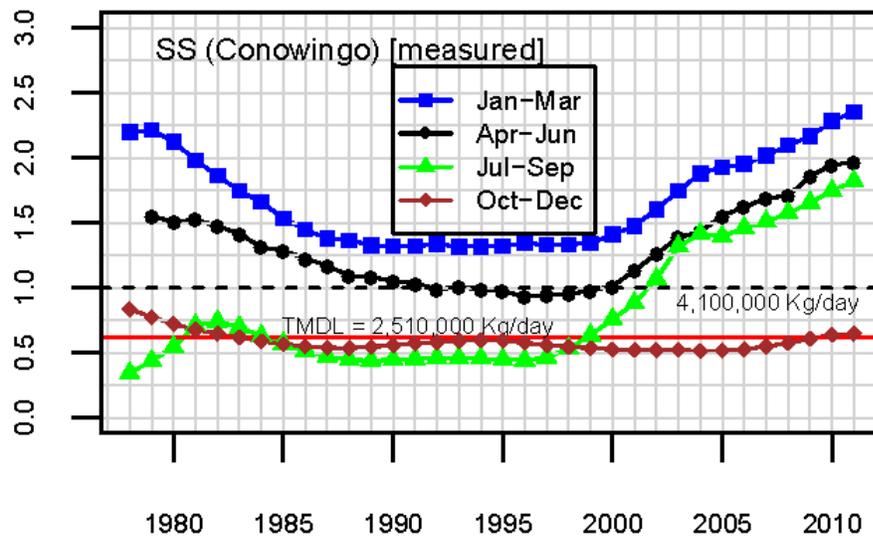


Figure 2.9. Seasonal averages of flow-normalized SS load in the Susquehanna River at the Conowingo station. All loads have been normalized by the median of long-term annual SS loads at Conowingo (locating at  $y = 1.0$ ). The TMDL of 2,510,000 kg/day set for Susquehanna River (US Environmental Protection Agency, 2010) is inserted for comparison.

all four seasons (Figure 2.10e).

The flow-normalized TN loads at Conowingo also show long-term trends that are similar among all four seasons, but opposite to those of SS and TP (*i.e.*, “rise-and-then-fall”), with the peak load occurring in the late 1980s (Figure 2.11a). Overall, the TN load at Conowingo has been brought closer and closer to the TMDL goal in recent years. Within TN, PN shows upward trends in Jan-Mar, Apr-Jun, and Jul-Sep (Figure 2.11b), whereas DN shows similar trends as those of TN (Figure 2.11c). Both DN and PN contributed to the TN rise until the late 1980s, and DN alone contributed to the TN reduction thereafter. Within DN,  $DNO_x$  shows “rise-and-then-fall” trends (Figure 2.11d), and DKN shows similar trends but with the start of the “fall” occurring 3-7 years earlier than  $DNO_x$  (Figure 2.11e).

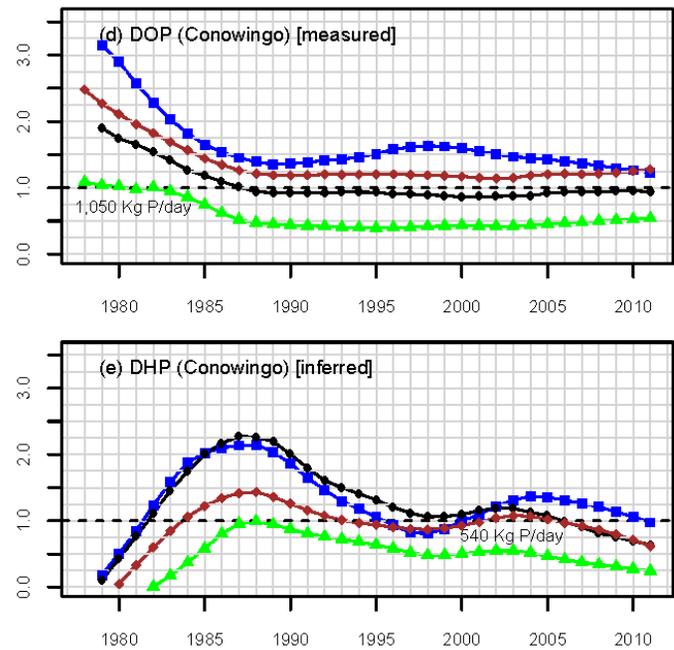
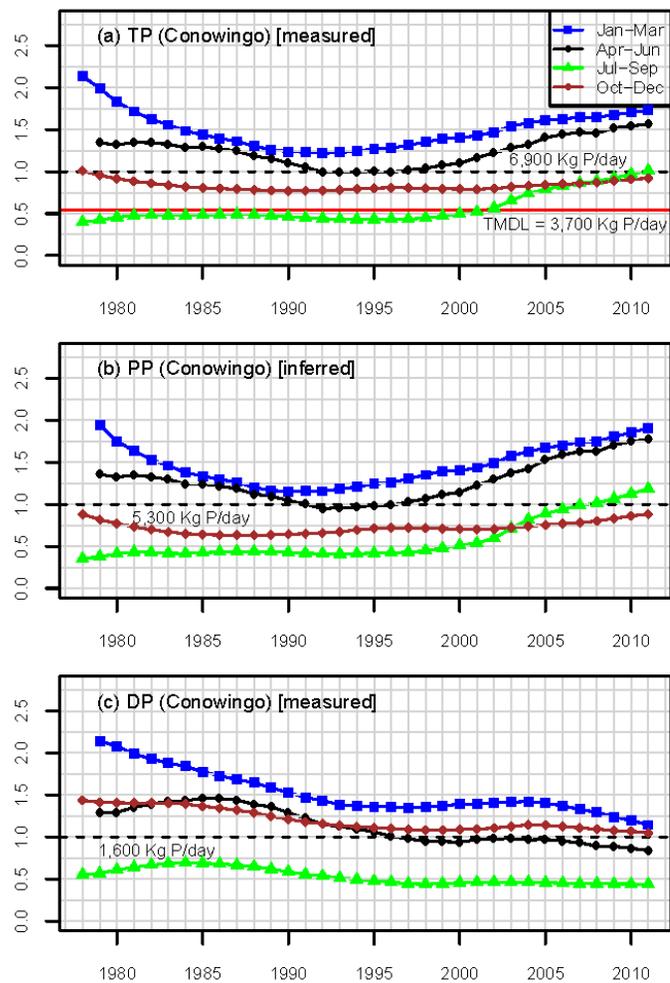


Figure 2.10. Seasonal averages of flow-normalized load of (a) TP, (b) PP, (c) DP, (d) DOP (dissolved orthophosphate), and (e) DHP (dissolved hydrolysable P) in the Susquehanna River at the Conowingo station. All loads have been normalized by the median of respective long-term annual loads at Conowingo (locating at  $y = 1.0$  in each panel). The TMDL of 3,700 kg P/day<sup>5</sup> set for Susquehanna River (US Environmental Protection Agency, 2010) is inserted in (a) for comparison.

<sup>5</sup> Relative to our published paper (Zhang *et al.*, 2013), the underlined text is a correction. (Note that the figure legend is correct and unchanged.)

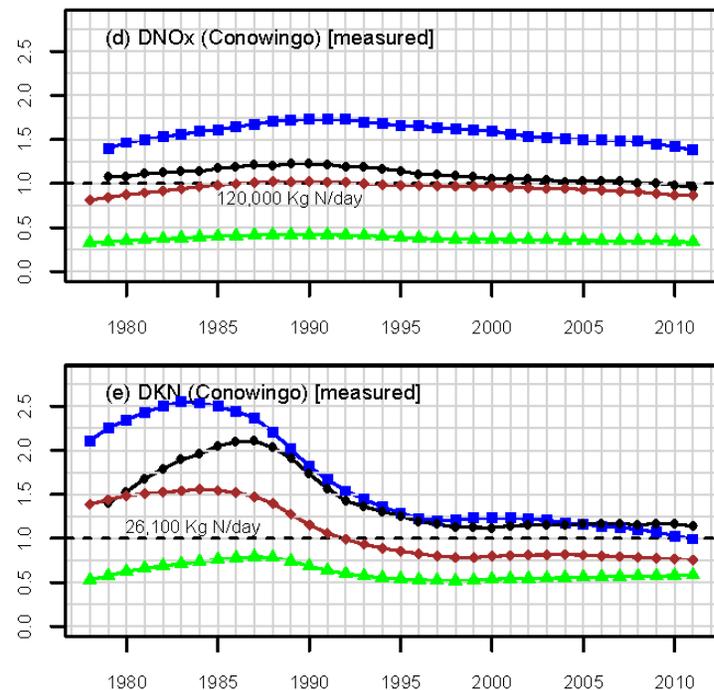
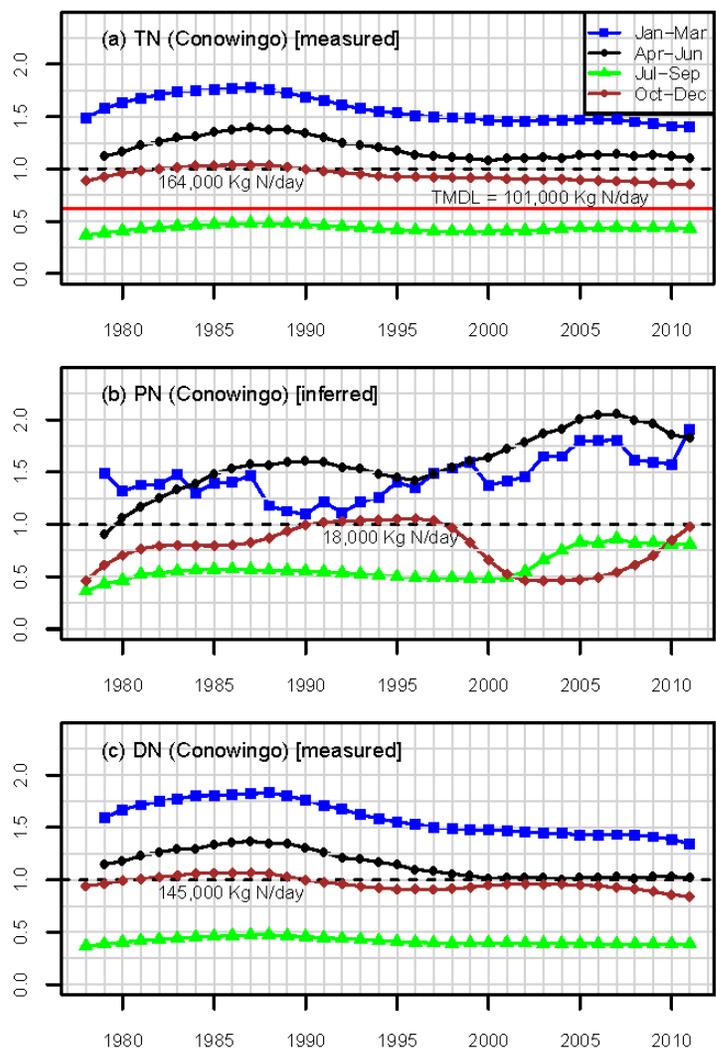


Figure 2.11. Seasonal averages of flow-normalized load of (a) TN, (b) PN, (c) DN, (d) DNO<sub>x</sub>, and (e) DKN in the Susquehanna River at the Conowingo station. All loads have been normalized by the median of respective long-term annual loads at Conowingo (locating at  $y=1.0$  in each panel). The TMDL of 101,000 kg N/day set for Susquehanna River (US Environmental Protection Agency, 2010) is inserted in (a) for comparison.

### **2.3.3.2. Discussion**

#### *Dissolved and particulate nutrient fractions at the Conowingo station*

As with the previously discussed “above-reservoir” results (Section 2.3.2), the below-reservoir results also show an overall trend of reducing load for both DN and DP. On the other hand, however, there is a clear upward trend of SS load at the Conowingo station since the late 1990s and accompanying increases in PP and PN loads. In terms of TN, there is still a trend of overall decline in all seasons because PN is a small portion of the TN and so PN has not reversed the progress achieved through DN reduction. For phosphorus, however, PP is the major fraction of TP, and the recent rise of PP has caused the TP to rise since the mid-1990s.

In terms of impact on the bay, the observed increases in particulate fractions of P and N are of concern. Although these particulate species are not as immediately available for algal consumption as are the DN (*e.g.* DNO<sub>x</sub>) and DP (*e.g.* DOP) species, a portion of the particulate species can undergo decomposition and generate bioavailable N and P to sustain algae growth (Kemp and Boynton, 1984). Such generation of bioavailable nutrients from particulate phases can be strongly promoted at conditions of high temperature (Kemp and Boynton, 1984) and low dissolved oxygen concentration (Boynton *et al.*, 1996), which are coupled characteristics of Chesapeake Bay in summer.

#### *The reservoir's role in sediment and particulate nutrient retention*

Considering that the watershed monitored by Conowingo has almost identical streamflow, watershed area, and land use pattern to that monitored by the Marietta and Conestoga stations (Table 2.1), the deteriorating situation of SS load at Conowingo can be largely attributed to the impact of the reservoir, as evidenced by a comparison between

these loads and those observed upstream (Figure 2.6). In fact, the reservoir appears to have been gradually losing its sediment storage capacity (SSC) especially since the mid-1990s, with concurrent effects on PP and PN. Correspondingly, the upward trends of PP and PN load at Conowingo suggest negative progress in particulate nutrient control for the overall non-tidal SRB, which can be largely attributed to the impact of the reservoir, as evidenced by comparisons between these loads and those observed upstream (Figure 2.7b and 2.8b). Coupled with the gradually diminishing SSC, the reservoir seems to be trapping less PP and PN than in the early years. The seasonal reservoir effluent trends at Conowingo are consistent with those observed by Hirsch (2012a; see Figure 13, Figure 17) using annual load estimates, and our new analysis of upstream data now further support his suggestion that recent changes reflect alterations in reservoir performance.

*Trends in rate of change in sediment inventory (storage) within the reservoir (1986-2011)*

To further explore the evolving behavior of the reservoir in modulating N, P, and SS load, we considered the reservoir as the control volume (CV), the combined load from Marietta and Conestoga as the input, and the load at Conowingo as the output. For simplicity, we ignored the watershed processes within the CV (*i.e.*, a small watershed area below Marietta (site No. 7 in Figure 2.1) and Conestoga (site No. 8 in Figure 2.1) and above Conowingo (site No. 9 in Figure 2.1), which corresponds to roughly 2.4% of the total watershed area above Conowingo). We then used the difference between our flow-normalized estimates of the SS input and output rates to roughly represent the rate of change in SS inventory within the reservoir, thus reflecting rates of storage or release. Similarly, we evaluated the rates of inventory change for PP and PN within the reservoir.

Changes in SS load across the reservoir showed net storage of SS in most years in all four seasons (Figure 2.12a), but the reservoir's capacity to trap new SS input has been gradually diminishing since the beginning of the record, albeit with an apparent plateau in net storage rate occurring in the 1990s. On a net basis, these flow-normalized results suggest that the reservoir may have started to actually lose SS in Jul-Sep since 2007 and in Jan-Mar since 2011, and it appears to be on a trajectory to start losing SS in Apr-Jun and Oct-Dec soon. Note that net loss of sediment is presumably related to scouring and that the Jul-Sep values since 2007 are likely the result of the historical hurricanes and storms that occur predominantly in this season. In fact, the flow-normalized estimate of net loss of SS in Jul-Sep for 2011 was sufficient to exceed the estimates of net storage in the other seasons, such that the estimate for that year would represent an overall annual net loss and for the first time in the history of the Conowingo Dam. Similarly, we observed gradually diminishing capacity of the reservoir for trapping new input of PP (Figure 2.12b) and PN (Figure 2.12c) in all four seasons in the last 26 years. For PP, the flow-normalized results imply net loss from the reservoir in Jul-Sep since 2007, Jan-Mar since 2008, and Apr-Jun and Oct-Dec since 2011. For PN, there has been net loss from the reservoir in Jan-Mar since 1997, Apr-Jun since 1998, Jul-Sep since 2003, and Oct-Dec since 2010. In terms of the corresponding estimates of annual change across the reservoir, flow-normalized output reach the estimates of input for PN and PP in 2003 and 2009, respectively, but with PP showing a much more rapid rise since the late 1990s (Figure 2.12b). The different trajectories of PP and PN (Figure 2.12) are possibly related to differences in the size fractions with which P and N are predominantly associated in the inlet, the outlet, or both (Hirsch, 2012b). In addition, there may also be substantial

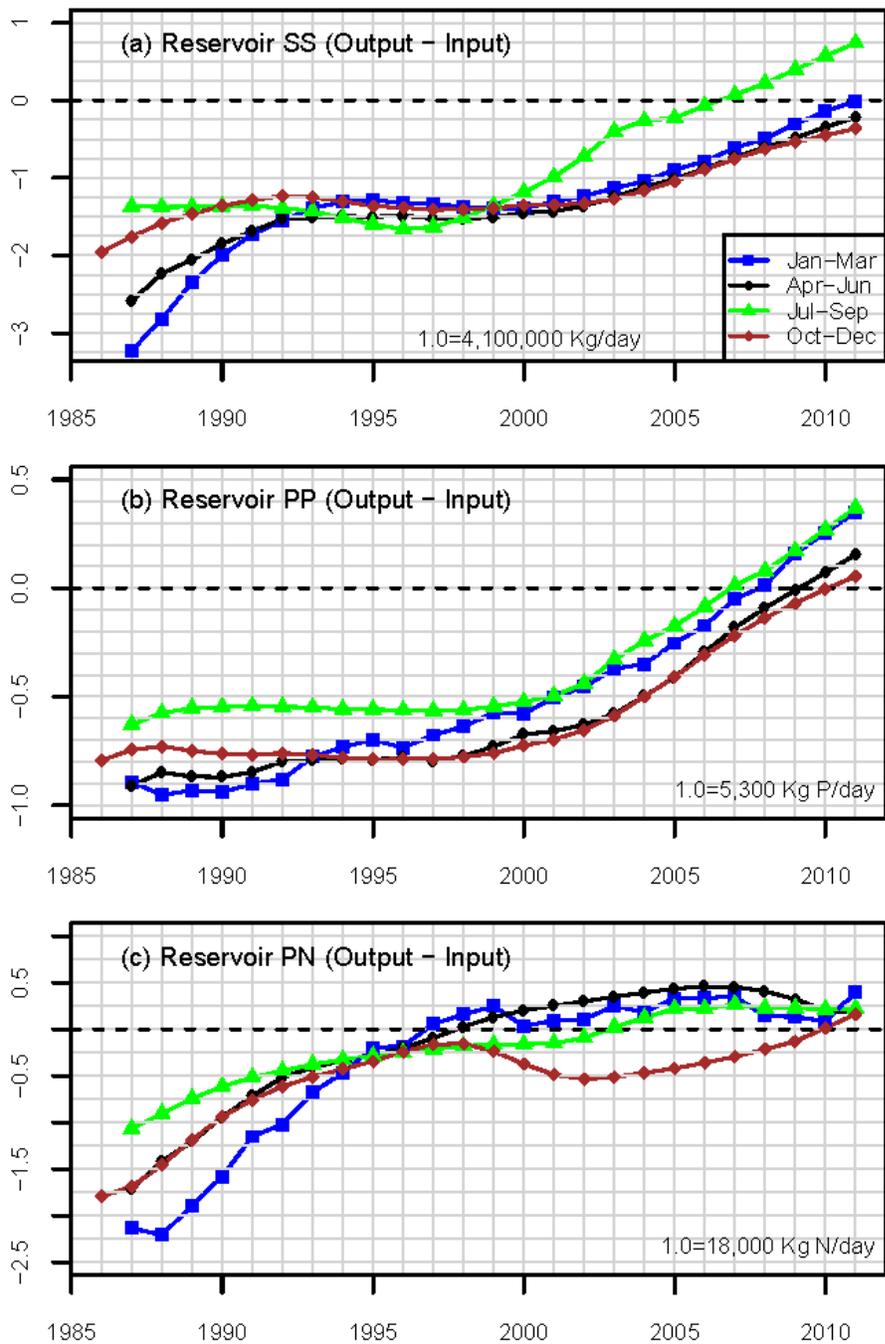


Figure 2.12. Rates of storage change in (a) SS, (b) PP, and (c) PN within Conowingo Reservoir based on flow-normalized load. All rates of change have been calculated as the differences between the loads at Conowingo (system output) and the combined loads from Marietta and Conestoga (system input), and then normalized by the median of respective long-term annual loads at the Conowingo station.

differences in regard to the overall (net) transformation of N and P between dissolved and particulate fractions, as might be expected from fundamental dissimilarities in the biogeochemical processes that affect each nutrient.

It should be noted that sediment or particulate nutrient retention in the reservoir in any given year would be affected by a combination of highly dynamic and complex processes. For example, short-term changes in sediment storage can occur due to scouring in storm events (Langland, 2009; Hirsch, 2012a) or due to short-term changes in reservoir stratification and biochemistry. In this regard, for example, the Jul-Sep results may have been especially influenced by some storm events in recent years, and the net SS loss in 2011 does not necessarily mean that this situation will continue in the coming years. Overall, such complications can confound the ability of the method to evaluate the “capacity to trap materials” under normal flow conditions. In addition, it is especially important to recall that flow-normalized trends in input and output loadings do not reflect the best estimate of “true conditions” for any given year. Nevertheless, it is evident from Figure 2.12 that there is a clear decadal-long trend of steady decline in the reservoir’s ability to trap sediment and particulate nutrients and there is reason for serious concern that the reservoir may be already at or near its storage capacity.

#### *Cumulative SS deposition in the reservoir (1987-2010)*

In order to assess the reservoir’s remaining SSC, one must consider both the available capacity in the reservoir and the on-going rate of sediment deposition. Considering the latter issue first, we note that WRTDS-based “true-condition” calculations of upstream and downstream SS loadings provide new estimates of annual SS deposition in the reservoir that are useful for comparisons against prior estimates made by others using other methods. In previous studies, Langland (2009) reported  $1.47 \times 10^7$  U.S. tons of SS

deposition from 1996 to 2008 (averaging  $1.23 \times 10^6$  tons/year) using a bathymetry mapping method, and  $1.69 \times 10^7$  tons of SS deposition for the same period (averaging  $1.41 \times 10^6$  tons/year) using monitored loading estimates. In comparison, our true-condition estimates suggest an average deposition loading of  $1.55 \times 10^6$  tons/year for 1996 to 2008, which matches reasonably well with those reported by Langland (2009), being roughly 27% and 10% higher, respectively.

In terms of the remaining SSC in the reservoir, this was reported to be  $4.2 \times 10^7$  tons in 1996 based on bathymetry mapping (Langland, 2009). Our true-condition estimate of cumulative SS deposition from 1996 to 2010 is  $2.0 \times 10^7$  tons. Thus, ~47% of the 1996 capacity had already been consumed and there was only  $2.2 \times 10^7$  tons of remaining SSC as of 2010, which was about 11% of the original 1928 SSC ( $2.04 \times 10^8$  tons) reported by Langland and Hainly (1997). This result clearly indicates that the reservoir is approaching its SSC.

#### *Evolving behavior of the reservoir in sediment and nutrient retention*

In addition, we investigated the evolving pattern of seasonal SS concentration in relation to river discharge at Marietta (reservoir inlet) and Conowingo (reservoir outlet), as an alternative method to examine the evolving behavior of the reservoir in SS retention. For each season, we selected the middle day as representative of the season. At Marietta, the SS concentration vs. discharge relationships appear to be similar for the selected dates between 1990 and 2010 in all four seasons (Figure 2.13a-d). At Conowingo, however, this relationship has gradually shifted upward since the beginning of the study period in all four seasons (Figure 2.13e-h). The results suggest that, for a given flow condition, there are higher SS concentrations in recent years than in the early years in all four seasons. In addition, a significant shift in this relationship appears to

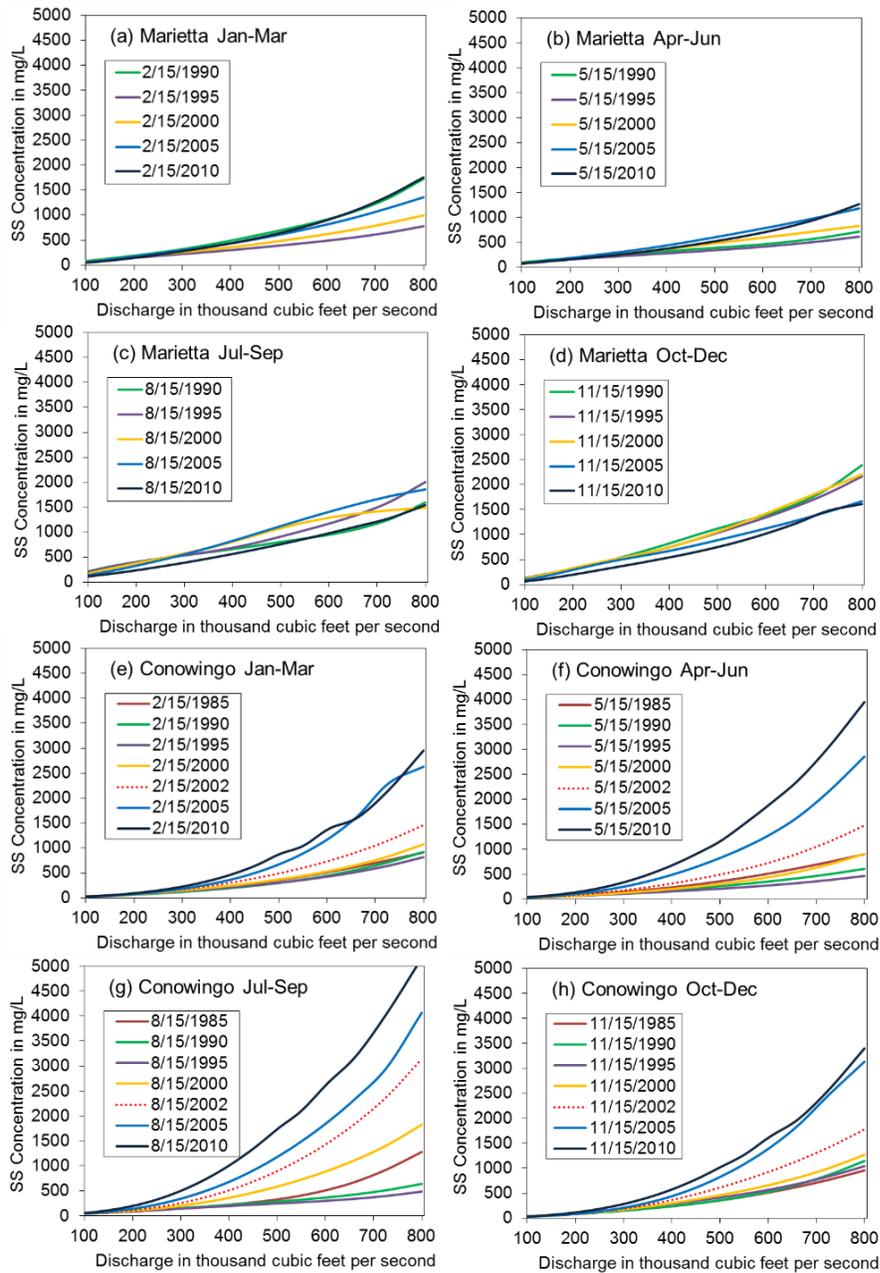


Figure 2.13. The evolving patterns of SS concentration vs. river discharge in each season at (a-d) the Marietta and (e-h) the Conowingo stations. A 5-year interval was selected to show the evolution. Patterns in 2002 (dashed line) were also added in (e) to (h) to aid comparison. (Note: as further discussed in Chapter 6 – see especially Figure 6.2 and related discussion, data are sparse at the high flow range (non-existent above 592,000 cfs) and plots are extremely uncertain in the extrapolated range beyond 400,000 cfs.)<sup>6</sup>

<sup>6</sup> Relative to our published paper (Zhang *et al.*, 2013), the underlined text is a cautious note.

have occurred around 2000 in all four seasons, consistent with our previous inference that the reservoir has been losing its SSC especially since the late 1990s (Figure 2.12). Similar patterns were also observed for PN and PP (data not shown). Thus, the concentration results (Figure 2.13) tend to confirm our earlier conclusion that the reservoir is now less efficiently trapping SS, PP, and PN than in the earlier years, and particularly so during high flow conditions. Considering only the date of September 1 as a point of comparison among years, Hirsch (2012a) has also observed rising patterns of TN, TP, and SS as a function of discharge (Hirsch, 2012a; see Figures 8, 12, 16).

#### *Summary and broader implications*

The above analyses have two important implications. First, the flow-normalized reservoir input and output trends (Figure 2.12) suggest that the reservoir has been steadily losing its storage capacity for SS, PP, and PN over the past two to three decades, and especially so for phosphorus since the mid-1990s. Second, both these trends and the concentration vs. discharge plots (Figure 2.13) show that the reservoir is becoming increasingly sensitive to scour events and that it has neared its storage capacity. Despite earlier predictions that the reservoir may not reach its total SSC until 2024–2029 (Langland, 2009), it is evident that increasingly substantial amounts of SS, PP, and PN are already entering Chesapeake Bay as the result of major reductions in reservoir performance toward sediment retention. Moreover, one might expect such increases to be further intensified if there are more frequent and intense major storms as the result of changing climate (Rabalais *et al.*, 2009; Najjar *et al.*, 2010).

On a seasonal basis, these findings complement the annual estimates recently provided by Hirsch (2012a). The current study adds additional information about flow-normalized seasonal trends of multiple nutrient and sediment species, with a special

focus on directly comparing above- and below-reservoir loading estimates as a means of considering long-term trends in reservoir performance.

Although recent rises in loadings of particulate-based nutrients have been at least in part counter-acted by reductions in the more readily available dissolved species, the changes in reservoir performance will pose significant new kinds of challenges to attainment of TMDL goals for the SRB. In this regard, our results reinforce recommendations recently made by Hirsch (2012a) -- *i.e.*, that these changes need to be factored into the proper establishment of regulatory load requirements and the development of watershed implementation plans. As better described elsewhere (Susquehanna River Basin Commission Sediment Task Force, 2002), a wide range of riverine, upland, and reservoir management options will need to be considered for controlling the sediment load in the non-tidal SRB.

#### **2.4. Conclusions**

This paper presents the results of our analyses of long-term seasonal trends of flow-normalized N, P, and SS loads from the non-tidal Susquehanna River to Chesapeake Bay. Major findings include:

- Long-term trends of flow-normalized N, P, and SS load generally followed similar patterns in all four seasons, implying that changes in watershed function and land use had similar impacts on nutrient and sediment load at all times of the year.
- 67-year concentration and load histories of NO<sub>x</sub> at the fall-line of the Susquehanna River show a steady rise from 1945 to 1990 and a steady decline thereafter.

- Flow-normalized loads of N, P, and SS have been generally reduced in the SRB above Conowingo Reservoir in the last 26 years, which can most likely be attributed to a suite of management control actions on point, agricultural, and stormwater sources.
- Flow-normalized loads of SS, PP, and PN at the outlet of Conowingo Reservoir have been generally rising since the mid-1990s. The reservoir's capacity to trap these materials has been diminishing, and it has neared its sediment storage capacity. These important changes will pose significant new kinds of challenges to attainment of TMDL goals for the SRB.

## **2.5. Supporting Information**

Supporting information to this chapter is provided in Appendix A. All derived data from this work, along with the raw river monitoring data, are stored at the publicly accessible Johns Hopkins University Data Archive via <http://dx.doi.org/10.7281/T1KW5CX5> (Zhang and Ball, 2014).

## **2.6. Acknowledgements**

This research was funded by the U.S. Water Environment Research Federation under Grant No. U4R09 and the National Science Foundation under Award No. 0854329. The authors benefited greatly from suggestions and comments from anonymous reviewers and Robert Hirsch (USGS), and would also like to thank Rebecca Murphy for useful discussions. This work would not have been possible without the streamflow and concentration data made available through the USGS National Water Quality Assessment Program and the SRBC Susquehanna Nutrient Assessment Program. The authors also want to thank Mike Langland (USGS) and Lori Sprague (USGS) for permission to adapt Table 2.1 and Figures 2.1-2.2 from their reports.

## 2.7. Literature Cited

- Ator, S. W., J. W. Brakebill and J. D. Blomquist, 2011. Sources, fate, and transport of nitrogen and phosphorus in the Chesapeake Bay watershed: An empirical model. *Scientific Investigations Report 2011-5167*. US Geological Survey, Reston, VA, p. 27.  
<http://pubs.usgs.gov/sir/2011/5167/>.
- Belval, B. L. and L. A. Sprague, 1999. Monitoring nutrients in the major rivers draining to Chesapeake Bay. *Water-Resources Investigations Report 99-4238*. US Geological Survey, p. 8. [http://va.water.usgs.gov/online\\_pubs/WRIR/99-4238/wrir\\_99\\_4238\\_text.pdf](http://va.water.usgs.gov/online_pubs/WRIR/99-4238/wrir_99_4238_text.pdf).
- Blankenship, K., 1999. NY sewage treatment plant to install BNR, boost Bay efforts. *Chesapeake Bay Journal* (March 1999).
- Blankenship, K., 2012. Conowingo Dam releasing pollutants at more frequent rate. *Chesapeake Bay Journal* (June 2012).
- Boicourt, W. C., 1992. Influences of circulation processes on dissolved oxygen in the Chesapeake Bay. *In: Dissolved Oxygen in Chesapeake Bay*, D. Smith, M. Leffler and G. Mackiernan (Editors). Maryland Sea Grant College, College Park, MD, pp. 7-59.
- Boynton, W. R. and W. M. Kemp, 2000. Influence of river flow and nutrient loads on selected ecosystem processes: A synthesis of Chesapeake Bay data. *In: Estuarine Science: A Synthesis Approach to Research and Practice*, J. E. Hobbie (Editors). Island Press, Washington, D.C., pp. 269-298.
- Boynton, W. R., W. M. Kemp, J. M. Barnes, L. L. Matteson, F. M. Rohland, L. L. Megdeberger and B. J. Weaver, 1996. Ecosystem Processes Component, Level 1, Interpretive Report No. 12. *CBL Ref. No. 95-039*. Chesapeake Bay Laboratory, University of Maryland, Solomons.
- Brakebill, J. W., S. W. Ator and G. E. Schwarz, 2010. Sources of suspended-sediment flux in streams of the Chesapeake Bay Watershed: A regional application of the SPARROW model. *J. Am. Water Resour. Assoc.* 46:757-776, DOI: 10.1111/j.1752-

1688.2010.00450.x.

Chesapeake Bay Program, 1998. Chesapeake Bay watershed model application and calculation of nutrient and sediment loadings, Appendix F: Point source loadings. Annapolis, MD, p.

693. [http://www.chesapeakebay.net/content/publications/cbp\\_12313.pdf](http://www.chesapeakebay.net/content/publications/cbp_12313.pdf).

Chesapeake Executive Council, 1988. Baywide nutrient reduction strategy: an agreement commitment report. Annapolis, MD.

[http://www.chesapeakebay.net/content/publications/cbp\\_12114.pdf](http://www.chesapeakebay.net/content/publications/cbp_12114.pdf).

Cloern, J. E., 2001. Our evolving conceptual model of the coastal eutrophication problem. *Mar. Ecol. Prog. Ser.* 210:223-253, DOI: 10.3354/meps210223.

Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.

Friedl, G. and A. Wüest, 2002. Disrupting biogeochemical cycles—Consequences of damming. *Aquat. Sci.* 64:55-65, DOI: 10.1007/s00027-002-8054-0.

Goodrich, D. M., W. C. Boicourt, P. Hamilton and D. W. Pritchard, 1987. Wind-induced destratification in Chesapeake Bay. *J. Phys. Oceanogr.* 17:2232-2240, DOI: 10.1175/1520-0485(1987)017<2232:widicb>2.0.co;2.

Hagy, J. D., W. R. Boynton, C. W. Keefe and K. V. Wood, 2004. Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658, DOI: 10.1007/bf02907650.

Hirsch, R. M., 2012a. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.

Hirsch, R. M., 2012b. Personal communication.

Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water*

*Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.

- Jarnagin, S. T., 2007. Historical analysis of the relationship of streamflow flashiness with population density, imperviousness, and percent urban land cover in the Mid-Atlantic region. *Internal Report APM 408*. US Environmental Protection Agency, Reston, VA, p. 103. [http://www.epa.gov/esd/land-sci/pdf/Jarnagin\\_2007\\_Historical\\_Analysis\\_of\\_Streamflow\\_Flashiness.pdf](http://www.epa.gov/esd/land-sci/pdf/Jarnagin_2007_Historical_Analysis_of_Streamflow_Flashiness.pdf).
- Jossette, G., B. Leporcq, N. Sanchez and Philippon, 1999. Biogeochemical mass-balances (C, N, P, SI) in three large reservoirs of the Seine basin (France). *Biogeochemistry* 47:119-146, DOI: 10.1007/bf00994919.
- Kemp, M. W. and W. R. Boynton, 1984. Spatial and temporal coupling of nutrient inputs to estuarine primary production: the role of particulate transport and decomposition. *Bull. Mar. Sci.* 35:522-535.
- Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C. Stevenson, 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* 303:1-29, DOI: 10.3354/meps303001.
- Kemp, W. M., J. M. Testa, D. J. Conley, D. Gilbert and J. D. Hagy, 2009. Temporal responses of coastal hypoxia to nutrient loading and physical controls. *Biogeosciences* 6:2985-3008, DOI: 10.5194/bg-6-2985-2009.
- Kogelmann, W. J., H. S. Lin, R. B. Bryant, D. B. Beegle, A. M. Wolf and G. W. Petersen, 2004. A statewide assessment of the impacts of phosphorus-index implementation in Pennsylvania. *J. Soil Water Conserv.* 59:9-18, <http://www.jswconline.org/content/59/1/9.short>.
- Langland, M. J., 2009. Bathymetry and sediment-storage capacity change in three reservoirs on the Lower Susquehanna River, 1996-2008. U.S. Geological Survey Scientific Investigations Report 2009-5110, Reston, VA, p. 21. <http://pubs.usgs.gov/sir/2009/5110/>.

- Langland, M. J., 2009. Bathymetry and sediment-storage capacity change in three reservoirs on the Lower Susquehanna River, 1996-2008. *Scientific Investigations Report 2009-5110*. US Geological Survey, Reston, VA, p. 21. <http://pubs.usgs.gov/sir/2009/5110/>.
- Langland, M. J., J. D. Blomquist, D. L. Moyer and K. E. Hyer, 2012. Nutrient and suspended-sediment trends, loads, and yields and development of an indicator of streamwater quality at nontidal sites in the Chesapeake Bay watershed, 1985-2010. U.S. Geological Survey Scientific Investigations Report 2012-5093, Reston, VA, p. 26. <http://pubs.usgs.gov/sir/2012/5093/pdf/sir2012-5093.pdf>.
- Langland, M. J. and R. A. Hainly, 1997. Changes in bottom-surface elevations in three reservoirs on the lower Susquehanna River, Pennsylvania and Maryland, following the January 1996 flood - implications for nutrient and sediment loads to Chesapeake Bay. *Water-Resources Investigations Report 97-4138*. US Geological Survey, Lemoyne, PA, p. 34. [pa.water.usgs.gov/reports/wrir97-4138.pdf](http://pubs.usgs.gov/reports/wrir97-4138/pdf).
- Langland, M. J., D. L. Moyer and J. Blomquist, 2007. Changes in streamflow, concentrations, and loads in selected nontidal basins in the Chesapeake Bay Watershed, 1985-2006. U.S. Geological Survey Open File Report 2007-1372, Reston, VA, p. 68. <http://pubs.usgs.gov/of/2007/1372/>.
- Lemunyon, L. J. and G. R. Gilbert, 1993. The concept and need for a phosphorus assessment tool. *J. Prod. Agric.* 6:483-486, <http://www.mendeley.com/research/concept-need-phosphorus-assessment-tool/>.
- Litke, D. W., 1999. Review of phosphorus control measures in the United States and their effects on water quality. *Water-Resources Investigations Report 99-4007*. US Geological Survey, Denver, CO, p. 43. <http://pubs.usgs.gov/wri/wri994007/>.
- Malone, T. C., L. H. Crocker, S. E. Pike and B. W. Wendler, 1988. Influences of river flow on the dynamics of phytoplankton production in a partially stratified estuary. *Mar. Ecol. Prog. Ser.* 48:235-249.
- McGonigal, K. H., 2010. Nutrients and suspended sediment transported in the Susquehanna River

- Basin, 2009, and trends, January 1985 through December 2009. *Publication No. 272*.  
Susquehanna River Basin Commission, p. 36.  
[http://www.srbc.net/pubinfo/techdocs/Publication\\_234/NutrientIntro.pdf](http://www.srbc.net/pubinfo/techdocs/Publication_234/NutrientIntro.pdf).
- Medeiros, P. R. P., B. A. Knoppers, G. H. Cavalcante and W. F. L. d. Souza, 2011. Changes in nutrient loads (N, P and SI) in the São Francisco estuary after the construction of dams. *Braz. Arch. Biol. Technol.* 54:387-397, DOI: 10.1590/s1516-89132011000200022.
- Murphy, R. R., W. M. Kemp and W. P. Ball, 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* 34:1293-1309, DOI: 10.1007/s12237-011-9413-7.
- Najjar, R. G., C. R. Pyke, M. B. Adams, D. Breitburg, C. Hershner, M. Kemp, R. Howarth, M. R. Mulholland, M. Paolisso, D. Secor, K. Sellner, D. Wardrop and R. Wood, 2010. Potential climate-change impacts on the Chesapeake Bay. *Estuar. Coast. Shelf Sci.* 86:1-20, DOI: 10.1016/j.ecss.2009.09.026.
- Neal, C., H. P. Jarvie, P. J. A. Withers, B. A. Whitton and M. Neal, 2010. The strategic significance of wastewater sources to pollutant phosphorus levels in English rivers and to environmental management for rural, agricultural and urban catchments. *Sci. Total Environ.* 408:1485-1500, DOI: 10.1016/j.scitotenv.2009.12.020.
- New York State Department of Environmental Conservation, 2007. New York State Tributary Strategy for Chesapeake Bay Restoration.  
[http://www.dec.ny.gov/docs/water\\_pdf/cbaystratfinal.pdf](http://www.dec.ny.gov/docs/water_pdf/cbaystratfinal.pdf).
- Pennsylvania Department of Environmental Protection, 2004. Pennsylvania's Chesapeake Bay Tributary Strategy. <http://www.elibrary.dep.state.pa.us/dsweb/Get/Version-45267/3900-BK-DEP1656.pdf>.
- Pritchard, D. W. and J. R. Schubel, 2001. Human influences on physical characteristics of the Chesapeake Bay. *In: Discovering the Chesapeake: the history of an ecosystem*, P. D. Curtin, G. S. Brush and G. W. Fisher (Editors). The Johns Hopkins University Press, Baltimore, MD, pp. 60-82.

- R Development Core Team, 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rabalais, N. N., R. E. Turner, R. J. Diaz and D. Justić, 2009. Global change and eutrophication of coastal waters. *ICES J. Mar. Sci.* 66:1528-1537, DOI: 10.1093/icesjms/fsp047.
- Sharpley, A. N., T. Daniel, T. Sims, J. Lemunyon, R. Stevens and R. Parry, 2003. Agricultural phosphorus and eutrophication, 2<sup>nd</sup> ed. *Agricultural Research Service, ARS-149*. US Department of Agriculture, p. 44.  
<http://ddr.nal.usda.gov/dspace/bitstream/10113/26693/1/CAT30907360.pdf>.
- Sprague, L. A., R. M. Hirsch and B. T. Aulenbach, 2011. Nitrate in the Mississippi River and its tributaries, 1980 to 2008: are we making progress? *Environ. Sci. Technol.* 45:7209-7216, DOI: 10.1021/es201221s.
- Sprague, L. A., M. J. Langland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. *Water-Resources Investigations Report 00-4218*. US Geological Survey, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).
- Susquehanna River Basin Commission, 2012. Sediment and nutrient assessment program.
- Susquehanna River Basin Commission Sediment Task Force, 2002. Sediment Task Force recommendations. *Publication No. 221*. Susquehanna River Basin Commission, p. 21.  
[http://www.srbc.net/programs/docs/sedi\\_task\\_force\\_rec221.pdf](http://www.srbc.net/programs/docs/sedi_task_force_rec221.pdf).
- Tukey, J. W., 1977. *Exploratory data analysis*. Reading, MA, Addison-Wesley, ISBN 0201076160
- US Department of Agriculture and US Environmental Protection Agency, 1999. Unified National Strategy for Animal Feeding Operations. <http://www.epa.gov/npdes/pubs/finafost.pdf>.
- US Environmental Protection Agency, 2000. Fact Sheet 1.0 - Stormwater Phase II Final Rule: An Overview. *Stormwater Phase II Final Rule Fact Sheet Series: EPA 833-F-00-001*.
- US Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for

Nitrogen, Phosphorus and Sediment.

<http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html>.

US Geological Survey, 2012a. Chesapeake Bay river input monitoring program.

US Geological Survey, 2012b. Surface-water data for the nation.

Weld, J. L., R. L. Parsons, D. B. Beegle, A. N. Sharpley, W. J. Gburek and W. R. Clouser, 2002.

Evaluation of phosphorus-based nutrient management strategies in Pennsylvania. *J. Soil Water Conserv.* 57:448-454, <http://www.jswconline.org/content/57/6/448.short>.

Zhang, Q. and W. P. Ball, 2014. Data associated with Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. Johns Hopkins University Data Archive, Baltimore, MD.

<http://dx.doi.org/10.7281/T1KW5CX5>, DOI: 10.7281/T1KW5CX5.

*Page intentionally left blank*

### Chapter 3. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation<sup>7</sup>

#### Abstract

Reduction of suspended sediment (SS), total phosphorus (TP), and total nitrogen is an important focus for Chesapeake Bay watershed management. The Susquehanna River, the bay's largest tributary, has drawn attention because SS loads from behind Conowingo Dam (near the river's mouth) have been rising dramatically. To better understand these changes, we evaluated histories of concentration and loading (1986-2013) using data from sites above and below Conowingo Reservoir. First, observed concentration-discharge relationships show that SS and TP concentrations at the reservoir inlet have declined under most discharges in recent decades, but without corresponding declines at the outlet, implying recently diminished reservoir trapping. Second, best estimates of mass balance suggest decreasing net deposition of SS and TP in recent decades over a wide range of discharges, with cumulative mass generally dominated by the 75<sup>th</sup>~99.5<sup>th</sup> percentile of daily Conowingo discharges. Finally, stationary models that better accommodate effects of riverflow variability also support the conclusion of diminished trapping of SS and TP under a range of discharges that includes those well below the

---

<sup>7</sup> This chapter (Abstract through Section 3.7) has been published as: Zhang, Q.; Hirsch, R. M.; Ball, W. P., Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* **2016**, *50*, (4), 1877-1886, doi: [10.1021/acs.est.5b04073](https://doi.org/10.1021/acs.est.5b04073). Co-authors Hirsch and Ball were both involved in hypothesis development, study design, results interpretation and editing. Copyright 2016 ACS. Reproduced/modified by permission of ACS. All figures, tables, and data were created by Qian Zhang unless otherwise indicated. Section, table, and figure numbers have been modified. References have also been re-formatted for consistency among chapters. Some minor substantive changes from the published manuscript have been made and are identified with footnotes.

literature-reported scour threshold. Overall, these findings suggest that decreased net deposition of SS and TP has occurred at sub-scour levels of discharge, which has significant implications for the Chesapeake Bay ecosystem.

### **3.1. Introduction**

To alleviate summertime hypoxia in Chesapeake Bay and associated impacts on estuarine ecology, reduction of nitrogen (N), phosphorus (P), and sediment loadings has been a long-term focus of Chesapeake Bay watershed management (Hagy *et al.*, 2004; Kemp *et al.*, 2005; Murphy *et al.*, 2011). This endeavor has been recently reinforced with the promulgation of total maximum daily loads (TMDLs) (U.S. Environmental Protection Agency, 2010) and state-wide efforts to establish watershed implementation plans (Linker *et al.*, 2013; Shenk and Linker, 2013).

Among Chesapeake Bay's tributaries, Susquehanna River is the largest (Hagy *et al.*, 2004; Murphy *et al.*, 2011) and is one of nine that account for over 90% of non-tidal discharge (Shenk and Linker, 2013). Of this 9-river non-tidal fraction, which has been modeled as accounting for ~77% of total freshwater discharge to the Bay (1991-2000; G. Shenk, personal communication) (Shenk and Linker, 2013), the Susquehanna has contributed ~62% of flow, ~65% of total nitrogen (TN), ~46% of total phosphorus (TP), and ~41% of suspended sediment (SS), as based on measured flows and estimated loads over the period 1979 to 2012 (Zhang *et al.*, 2015). The relatively lower fractional contributions of TP and SS reflect retention within the Lower Susquehanna River Reservoir System (LSRRS), which consists of Lake Clarke (formed in 1931), Lake Aldred (formed in 1910), and Conowingo Reservoir (formed in 1928) (Figure 3.1) (Reed and Hoffman, 1997; Langland, 2009; Langland, 2015). In general, reservoirs in early

stages of operation can effectively remove sediment and particulate P and N, mainly through particle deposition and burial (Jossette *et al.*, 1999; Friedl and Wüest, 2002), and for N, possible denitrification (Kemp *et al.*, 1990; Testa *et al.*, 2013). Relative removal rates among constituents are also affected by escaping particles that are finer and therefore higher in P concentration than those retained (Hainly *et al.*, 1995; Horowitz *et al.*, 2012) and by incoming N, which is predominantly dissolved and in contrast to P,

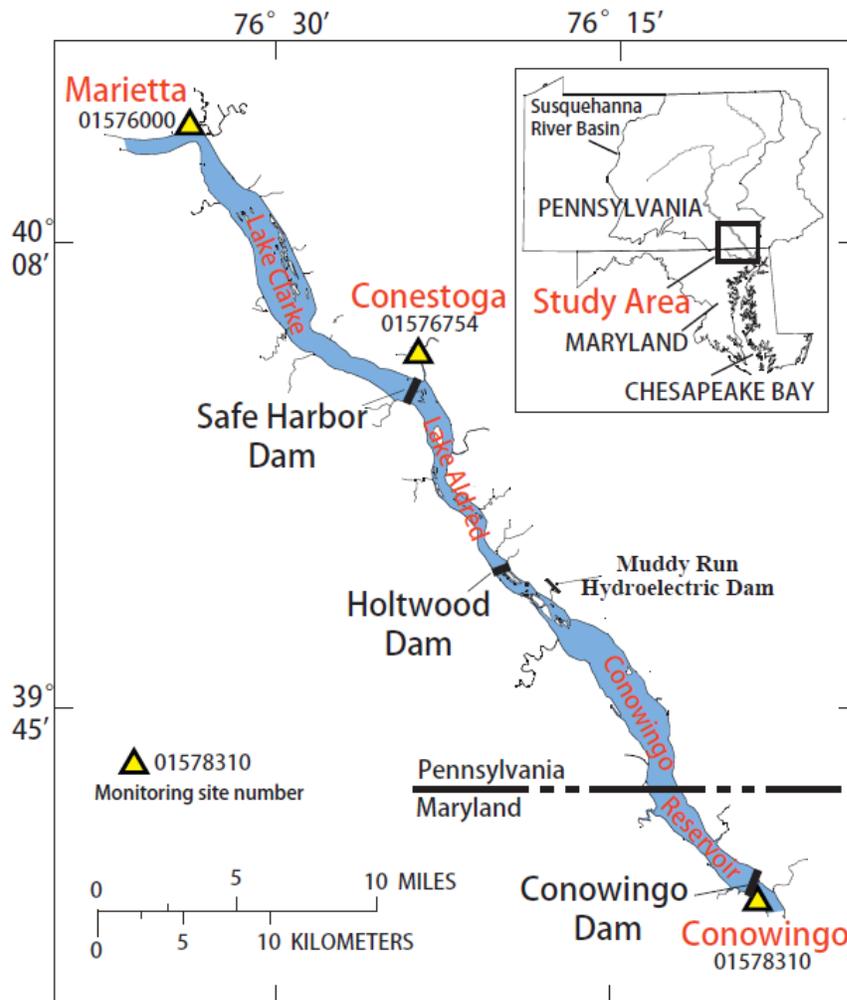


Figure 3.1. Map of the Lower Susquehanna River Reservoir System consisting of Lake Clarke, Lake Aldred, and Conowingo Reservoir. Yellow triangles indicate the three monitoring sites: Conowingo, Marietta, and Conestoga. (See Table B1 for site details.) This figure was modified after Figure 1 in Langland (2015) with simplifications.

which is predominantly bound to particles (Hainly *et al.*, 1995; Hirsch, 2012; Zhang *et al.*, 2013). In this regard, estimates suggest that the LSRRS has historically trapped about 70%, 45%, and 2% of SS, TP, and TN loads, respectively (Langland and Hainly, 1997). Unfortunately, however, Lake Clarke and Lake Aldred have been effectively filled for several decades and the largest and most downstream reservoir, Conowingo, is reaching the end of its effective life for sediment removal (Langland and Hainly, 1997; Langland, 2009; Langland, 2015), as supported by a growing body of evidence documenting substantial recent decline in net trapping of SS and particulate nutrients (Hirsch, 2012; Zhang *et al.*, 2013). In this regard, statistical evaluations of trends for both dissolved and particulate species across the LSRRS have suggested that input loadings of all species have declined since the 1980s, but that output loadings of SS and particulate nutrients have trended upward since the mid-1990s (Zhang *et al.*, 2013). Langland and Hainly (1997) have estimated that once Conowingo Reservoir reaches its sediment storage capacity and assuming no change in the inputs of SS and TP to the reservoir, average annual loads of SS and TP flowing past Conowingo Dam would increase by about 250% and 70%, respectively, as compared with loads observed before the reservoir neared its sediment storage capacity. In this context, a “SS scour threshold” of  $\sim 11,300 \text{ m}^3/\text{s}$  ( $400,000 \text{ ft}^3/\text{s}$ ) was reported in 1978, which has likely decreased in more recent years (Gross *et al.*, 1978). As reservoir storage approaches capacity, there are remaining questions about the new “dynamic equilibrium” that will occur. In particular, little is known about the intra-annual phenology of sediment discharges, that is, the relative importance of increased magnitude and frequency of scour events at very high discharges

and the decrease in sediment deposition during the much more frequent times of moderate to high discharges.

In the broader context of the non-tidal Chesapeake watershed, recent work on its nine major tributaries has documented general rising trends in SS and particulate nutrient loads (Zhang *et al.*, 2015). These trends are not well understood but may relate to (1) land-related practices such as land clearance and urbanization (Gellis *et al.*, 2008; Brakebill *et al.*, 2010), (2) removal of small mill dams (Walter and Merritts, 2008; Merritts *et al.*, 2011), and (3) increasing erosion of river bed and bank sediments (Brakebill *et al.*, 2010; Massoudieh *et al.*, 2013), particularly under the condition of increasing storm intensity (Karl and Knight, 1998). Within this context, the Susquehanna rise was estimated to have contributed ~92% and ~68% to the total (nine non-tidal tributaries) summed rise of SS and TP, respectively, during the decade of 2002-2012 (Zhang *et al.*, 2015).

To provide new insights on sediment and nutrient processing within the LSRRS, we have further evaluated the history of concentration and loading from sites above and below the LSRRS for the period between 1986 and 2013 (~30 years) using available data from streamflow and concentration monitoring by the U.S. Geological Survey (USGS) and the Susquehanna River Basin Commission (SRBC). Specifically, we performed three types of analyses on SS, TP, and TN with increasing use of statistical modeling:

- (1) Identification of changes in concentration-discharge relationships at sites above and below the reservoir using observed data – see Sections 3.2.1 and 3.3.1 for methods and results, respectively;
- (2) Evaluation of net deposition in the reservoir system using mass-balance analysis

based on best estimates of loadings – see Sections 3.2.2 and 3.3.2; and

- (3) Analysis of the effects of sediment accumulation on reservoir performance by better accommodating effects of streamflow variability through the development of three different historical stationary models of the concentration relation to discharge and season– see Sections 3.2.3 and 3.3.3.

These analyses have been made possible by the decadal-scale historical output and input data and the recent development of statistical modeling approaches. To our knowledge, such mass-balance analyses of long-term sediment and nutrient accumulations have heretofore not been conducted on any major reservoir system that is similarly close to the end of its period of effective sediment trapping.

## **3.2. Data and Methods**

### **3.2.1. Study Sites and Data**

The non-tidal Susquehanna River Basin covers portions of New York, Pennsylvania, and Maryland. It comprises four physiographic provinces, namely, Appalachian Plateaus (58% of the area), Valley and Ridge (32%), Piedmont (9%), and Blue Ridge (1%) (Zhang *et al.*, 2015). Land uses in this watershed comprise forested (67%), agricultural (29%), urban (2%), and other (2%) (Sprague *et al.*, 2000).

There are three monitoring sites located in the vicinity of the LSRRS (Figure 3.1; Table B1 in Appendix B1). The Conowingo site (drainage area: 70,189 km<sup>2</sup>) at the system outlet is at Conowingo Dam, which is on Susquehanna’s fall-line and about 10 miles from the river mouth at Havre de Grace, MD. This site has been monitored by the USGS since the 1970s and represents discharge from over 99% of the Susquehanna

watershed (Belval and Sprague, 1999). Two upstream sites, Marietta and Conestoga, have been monitored by the USGS (U.S. Geological Survey, 2014a) for streamflow and by the SRBC (Susquehanna River Basin Commission, 2014) for water quality since the mid-1980s. The Marietta site (drainage area: 67,314 km<sup>2</sup>) is on the mainstem and represents the majority (~96%) of the watershed represented by Conowingo. The Conestoga site (drainage area: 1,217 km<sup>2</sup>) monitors runoff from the small but heavily agricultural Conestoga tributary. At the three sites, annual average streamflow per unit area is comparable, with minimum values of 0.27-0.32 m/year and maximum values of 0.85-0.98 m/year (Table B1).<sup>8</sup>

### **3.2.2. Analysis with Standard WRTDS Models**

Selection of methods for estimating constituent concentration and loading based on low-frequency monitoring data has been an important area of hydrological research. Recently, Hirsch *et al.* (2010) have developed a method called “Weighted Regressions on Time, Discharge, and Season” (WRTDS). WRTDS provides improvements over prior methods, *e.g.*, ESTIMATOR (Cohn *et al.*, 1989), because it does not rely on assumptions about homoscedasticity of model errors, constancy of seasonal trends in concentration, or a fixed concentration-flow relationship (Hirsch *et al.*, 2010). In regard to homoscedasticity, ESTIMATOR invokes an assumption of constant residual error across all seasons and discharges and hence has a bias correction factor (BCF) that is also constant across all seasons and discharges. By contrast, WRTDS takes into account the substantial differences among these errors and the BCFs are calculated accordingly

---

<sup>8</sup> Relative to our published paper (Zhang *et al.*, 2016), the underlined text is a correction.

(Moyer *et al.*, 2012; Hirsch, 2014). Consequently, WRTDS estimates can better represent the changing seasonal and flow-related patterns and are more resistant to the problem of load-estimation bias. WRTDS has been used in a wide range of regional to national water-quality studies (Hirsch *et al.*, 2010; Hirsch, 2012; Moyer *et al.*, 2012; Zhang *et al.*, 2013; Hirsch, 2014; Hirsch and De Cicco, 2015; Zhang *et al.*, 2015).

We have applied WRTDS to estimate concentrations and loads for every day in the record based on daily streamflow ( $Q$ ) and more sparse concentration ( $C$ ) data. WRTDS was implemented using the  $R$  package called EGRET (Exploration and Graphics for RivEr Trends) (Hirsch and De Cicco, 2015). For a particular site, such estimation is performed in four steps. First, WRTDS establishes a set of evenly-spaced grid points on a surface defined by time ( $t$ ) and  $\log(Q)$ . Grid values for the time and discharge dimensions were selected in accordance with a standard grid design described in the user manual (*c.f.* pages 40, 46-47) (Hirsch and De Cicco, 2015). Coarser and finer grid resolutions were tested during the development of the WRTDS model. Results were shown to be insensitive to resolutions finer than this standard grid design. Second, for each grid point, WRTDS develops a separate weighted-regression model using observed data and estimates  $C$ :

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (3.1)$$

where  $\beta_i$  are fitted coefficients and  $\varepsilon$  is the error term. The log-log relationship between  $C$  and  $Q$  is used, because it has been well established that such formulation can provide a generally better fit than a linear relationship (e.g., Hirsch *et al.*, 1991; Helsel and Hirsch,

2002; Qian, 2010).<sup>9</sup> More importantly, the residuals from this model, fitted to the log of concentration, generally approximate normal distributions quite well, with only limited exceptions at the extremes of the distributions (Figures B1-B3 in Appendix B2-I). Step 2 thus results in an estimated concentration regression “surface” as functions of  $t$  and  $\log(Q)$  – see examples for Conowingo in Figure B4 (Appendix B2-II). In Step 3, concentration for each day in the record is estimated using a bilinear interpolation of this surface, with proper accommodation of re-transformation bias (Hirsch and De Cicco, 2015). Finally, the estimated  $C$  is multiplied by daily  $Q$  to estimate daily loading. To alleviate potential edge effects for years near the start or end of the record, the updated EGRET package (version 2.2.0) was used in this work; see the user manual (*c.f.* pages 17-18, 41) (Hirsch and De Cicco, 2015).

For each of the three sites, we implemented the standard WRTDS model to produce daily “true-condition” loading estimates (Hirsch *et al.*, 2010) for SS, TP, and TN, respectively. Residual analysis indicates that the residuals have no structural relationship with time, discharge, or season. (See Figures B5-B7 in Appendix B2-II.) To obtain uncertainty estimates on daily and annual loadings for each species at each site, we have followed the method of Hirsch *et al.* (2015) which involves resampling (with replacement) of the raw concentration data to obtain 100 realizations of representative data sets and associated WRTDS-based estimates of daily and annual loadings. Our approach is more fully described in Appendix B2-III, which includes uncertainty results for annual loadings (Figure B8).

---

<sup>9</sup> Relative to our published paper (Zhang *et al.*, 2016), the underlined text is a more accurate statement of the intended concept.

The loading estimates were then used for mass-balance analysis. Specifically, mass loading rates at Conowingo were used to represent reservoir output and those at Marietta and Conestoga were summed to represent the vast majority of reservoir input (97.6% of drainage area). Nonetheless, reservoir input was further adjusted by including an estimated contribution from the small unmonitored area above Conowingo and below Marietta and Conestoga. This estimate was made using Conestoga loadings and the appropriate drainage area ratio. Finally, output loads were subtracted from input loads to determine reservoir net deposition.

### **3.2.3. Analysis with Stationary WRTDS Models**

Inter-annual comparisons of loading and net deposition based on standard WRTDS models are influenced by the particular time history of discharges that happened in a given year as well as the concentration regression surface (concentration as a function of time and discharge). To better isolate and reveal the changes that have occurred in the concentration regression surface (which we presume to reflect changes in reservoir system function), we developed three historical “stationary” WRTDS models. (See Figure B9 in Appendix B2-IV.) The term “stationary” in our context means that a temporally-invariant regression surface was assumed to be applicable over the entire period of record (*i.e.*, a “stationary” probability function of concentration conditioned on discharge and season). By comparing results based on regression surfaces obtained from three decadal separated years, applied to the same streamflow record, we are able to isolate the effect of the change in the regression surface itself. Because all three histories developed use the exact same streamflow record, any observed differences should better

represent fundamental differences in reservoir system function among the three selected years.

First, we selected three one-year-wide  $C$  versus  $t$ ,  $\log(Q)$  regression surfaces from the standard WRTDS model. For simplicity but without losing generality, we selected the 1990, 2000, and 2010 annual surfaces (Figure B9). Second, we separately repeated each of these 1-year surfaces to fill in the entire time-span to produce three different “stationary” surfaces for the entire record. Finally, these three period-of-record stationary surfaces were respectively used in conjunction with the actual history of daily discharge to estimate daily loadings, using the interpolation approach as described in Section 3.2.2. Note that the difference among the three stationary models is captured by the selected surfaces, which are considered to represent water-quality conditions at the study site in those selected years. We performed this stationary-model analysis on SS, TP, and TN at each of the three sites. To provide estimates of the uncertainty of these modeled results, we have resampled (with replacement) the raw concentration data ten times to obtain concentration data replicates and used each replicate to develop the three (*i.e.*, 1990-, 2000-, and 2010-surface based) stationary models. The resulting range of load estimates forms uncertainty bands that are subsequently shown as dashed lines on plots. In addition, as with the standard WRTDS estimates, we have also conducted mass-balance analysis on the stationary-model estimates to evaluate net deposition.

### **3.3. Results and Discussion**

#### **3.3.1. Temporal Changes in Concentration-Discharge Relationships**

The manner in which nutrient and sediment concentrations vary with streamflow (*i.e.*,

$C$ - $Q$  relationships) reflect the relative role of dilution in comparison to mechanisms of dissolution, erosion, and transport (Evans and Davies, 1998; House and Warwick, 1998; Lefrançois *et al.*, 2007; Richardson, 2012; Burt *et al.*, 2015) and temporal changes in such relationships can therefore be useful indicators of changes in system function. In this section, I present an analysis of this type for the two mainstem sites: Conowingo (reservoir outlet) and Marietta (reservoir inlet).  $C$ - $Q$  scatterplots were constructed for three 9-year periods, 1987-1995 (“P<sub>1</sub>”), 1996-2004 (“P<sub>2</sub>”), and 2005-2013 (“P<sub>3</sub>”), as shown in Figure 3.2 and Figure B10 of Appendix B3. Non-parametric LOWESS (locally weighted scatterplot smoothing) curves are shown to better visualize the  $C$ - $Q$  relationships.

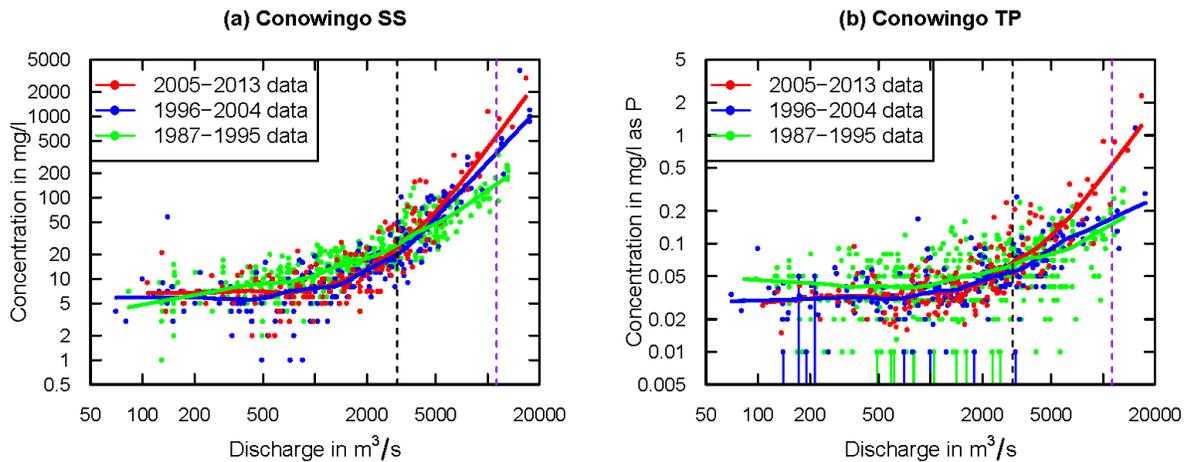


Figure 3.2. Observed concentration-discharge ( $C$ - $Q$ ) relationships for suspended sediment (SS) and total phosphorus (TP) in the Susquehanna River at Conowingo, MD, for three separate periods between 1987 and 2013. Solid lines are fitted LOWESS (locally weighted scatterplot smoothing) curves. Vertical black dashed lines in (a)-(b) correspond to 3000 m<sup>3</sup>/s. Vertical purple dashed lines in (a)-(b) indicate the literature scour threshold of 11300 m<sup>3</sup>/s (400000 ft<sup>3</sup>/s). Data points with vertical solid lines in (b) indicate left-censored concentration samples. (Detection limit varied with samples.)

For SS, the  $C-Q$  curves are generally convex upward. At Conowingo (Figure 3.2a),  $C-Q$  relationships are similar among the three periods at low discharges ( $< 3,000 \text{ m}^3/\text{s}$ ), but the curves at higher discharges are clearly more elevated at later periods (*i.e.*,  $P_2$  and  $P_3$ ). By contrast, the curves at Marietta (Figure B10b) show clear decline at later time ( $P_3$ ) at low discharges ( $< 3,000 \text{ m}^3/\text{s}$ ) but negligible difference between  $P_2$  and  $P_3$  at higher discharges. For TP, the patterns are generally similar to SS (Figure 3.2b and Figure B10d). For TN, Conowingo shows similar  $C-Q$  relationships among the three periods at most discharges, except that at the very high discharges ( $> 7,000 \text{ m}^3/\text{s}$ )  $P_2$  and  $P_3$  show higher concentrations than  $P_1$  (Figure B10e). At Marietta,  $P_3$  shows lower concentrations than  $P_1$  and  $P_2$  at most discharges (Figure B10f).

Overall, the  $C-Q$  relationships for Marietta show lower concentrations in SS, TP, and TN in  $P_3$  than  $P_1$  and  $P_2$  at most discharges, suggesting decreased inputs of these constituents from various sources (*e.g.*, agricultural, point, atmospheric, and stormwater sources) (Sprague *et al.*, 2000; Zhang *et al.*, 2013). Such promising changes at Marietta, however, have not been propagated across the reservoirs to emerge at Conowingo, where trends are reversed at high discharges. These results tend to corroborate previous reports of decreasing net trapping (Hirsch, 2012; Zhang *et al.*, 2013; Langland, 2015) and highlight the critical role of the reservoir for sediment and nutrient retention and storage.

### **3.3.2. Changes in Net Deposition: Analysis of Loads from Standard WRTDS Models**

In this section, I present mass-balance analysis of loadings across the reservoir using estimates from *standard* WRTDS models to estimate changes in net deposition in the reservoir (Section 3.2.2). While these estimates are currently our best approximation of historical loadings and can provide useful indication of reservoir function, we remark that

they are model outputs subject to limitations of sample availability and complications of inter-annual flow variability. In the latter regard, we also further analyze the results using *stationary* WRTDS models (Section 3.3.3).

### ***3.3.2.1. Cumulative SS Deposition in the Reservoir***

To consider recent sediment deposition rates in the context of the reservoir's ~85 years of service, we reconstructed a long-term record of cumulative deposition behind Conowingo Dam. Specifically, estimates of annual net deposition between 1987 and 2013 were obtained by applying mass-balance analysis on upstream and downstream estimates derived using the standard WRTDS models. These results were then interpreted on a storage-volume basis by assuming that the 2008 bathymetry-based estimate of sediment capacity (Langland, 2015) is correct. Our estimates of annual net deposition can then be used to estimate capacity in years before and after 2008 and these estimates can then be compared with other bathymetry data. The results are shown as green points in Figure B11 of Appendix B4-I. The resulting curve is very consistent with the 1996 bathymetry result and close to the 1990 bathymetry result, giving us some confidence in the method for the pre-2008 period.

Based on this curve, cumulative deposition behind Conowingo Dam has followed a concave shape between 1987 and 2013, suggesting a declining rate of net deposition during this period. Taking the year 2000 as a dividing-point, the estimated average rate of net deposition was  $\sim 1.70 \times 10^6$  tons/year for 1987-2000 but only  $\sim 0.72 \times 10^6$  tons/year for 2000-2010. By contrast, the 1929-1987 average rate was  $\sim 2.24 \times 10^6$  tons/year, based on the 1987 estimate and the original storage capacity (Figure B11 and Table B2; Appendix B4-I). Thus, the average rates during 1987-2000 and 2000-2010 are only 76% and 32%,

respectively, of the 1929-1987 rate. Overall, these results strongly suggest that annual rates of deposition have decreased over time, particularly in recent years. These findings are consistent with the literature (Hirsch, 2012; Zhang *et al.*, 2013; Langland, 2015).

It is noteworthy that the reconstructed curve from 2008 to 2011 (green points in Figure B11) did not match the 2011 bathymetry-based capacity, which was conducted immediately after Tropical Storm (TS) Lee (Gomez and Sullivan Engineers, 2012). In particular, bathymetry suggests continued net deposition between 2008 and 2011, whereas WRTDS results suggest net scour in 2011. This inconsistency may reflect inaccuracies in either type of measurement and perhaps relate to both under-estimated input loadings owing to missed input sampling during three extremely highflow days in 2011 (*i.e.*, September 8<sup>th</sup>, 9<sup>th</sup>, and 11<sup>th</sup>) and to over-estimated output loadings at the high discharge end. Given the currently available data, we have conducted an uncertainty analysis to assign a 95% confidence interval on the cumulative deposition between 1987 and 2013, as shown with blue and purple points in Figure B11. This uncertainty interval increases dramatically during years of high flow such that it is able to capture the 2011 bathymetry as well as the earlier bathymetry results. In order to further examine the WRTDS results in the context of bathymetry, it is critical to conduct both new bathymetry surveys and continued monitoring of reservoir input and output, with more emphasis on fully capturing representative flow conditions.

### ***3.3.2.2. Changes in Net Deposition over Time***

To better understand decadal-scale changes in net deposition of sediment and nutrients, we quantified output/input ratios (O/I) for SS, TP, and TN and prepared boxplots of these ratios for each complete year (*i.e.*, 1987-2013) (Figure 3.3). The ratios

were calculated from 35-day moving averages of input and output loads to reflect travel time between Marietta and Conowingo. (Estimation of physical travel time in the LSRRS and rationalization of the selection of 35 days for the averaging period are described in Appendix B4-II. This analysis reveals that the general trend is insensitive to averaging period for selections between 1 and 35 days; see Figures B12-B14.)

O/I results based on 35-day moving averages of input and output are provided in Figure 3.3, including both a single realization of WRTDS modeling of the original data (Figure 3.3a-3.3c) and uncertainty analysis based on the 100 realizations obtained as described above (Section 3.2.2). Average annual median values of O/I, and their 95%

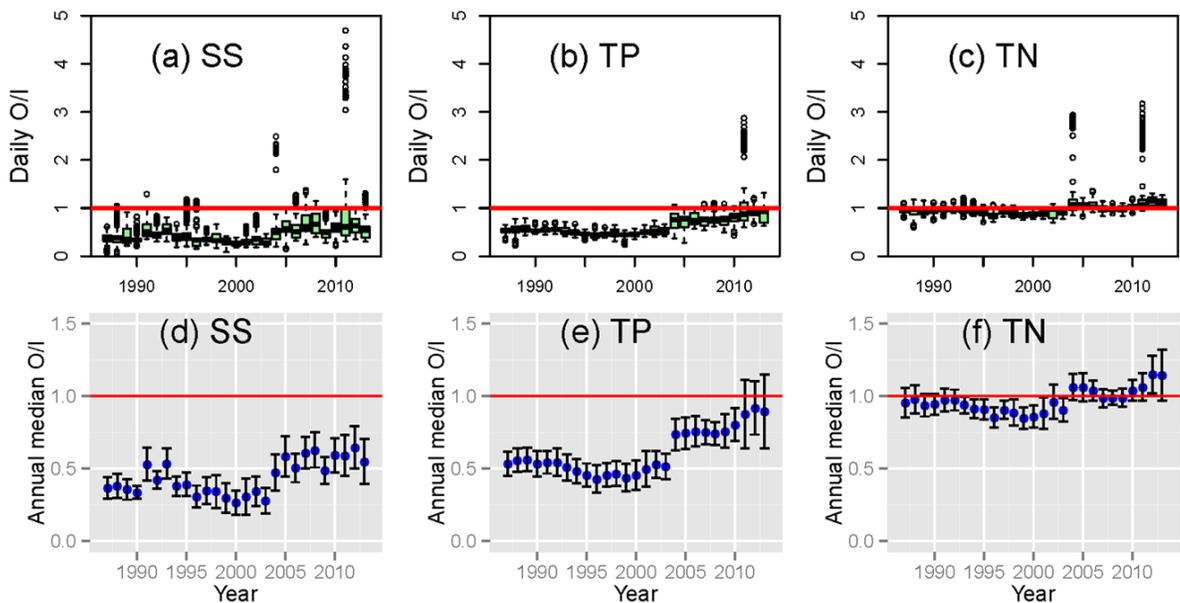


Figure 3.3. Estimated output/input ratios (O/I) based on 35-day moving averages of WRTDS-estimated output and input loadings. Plots (a)-(c) are annual boxplots for (a) suspended sediment (SS), (b) total phosphorus (TP), and (c) total nitrogen (TN). (Note that each boxplot represents 365 daily data points and that  $O/I < 1.0$  reflects net deposition.) Plots (d)-(f) show the results of uncertainty analysis based on 100 synthetic data sets – see text. Plots show the averages of annual median values (blue dots) and the 95% confidence intervals (black error bars).

confidence intervals, are shown in Figures 3.3d-3.3f. Average numbers of excursions above 1.0, as well as their confidence intervals, are provided in Figure B15 (Appendix B4-III).

In regard to SS and TP, Figures 3.3a-3.3b reveal that annual median O/I has been rising, particularly since the 2000s. In addition, Figures 3.3d-3.3e confirm that the trends in annual median O/I are qualitatively maintained based on the 100 realizations, but with overlapping confidence intervals. Figures 3.3a-3.3b further reveal that annual 75<sup>th</sup> percentiles are approaching (SS) or exceeding (TP) 1.0, with substantial excursions in certain years (Figures 3.3a-3.3b; B15a-B15h). Overall, these results confirm the aforementioned trends of declining rate of net deposition. Given that streamflow O/I has been stable at ~1.0 in 1987-2013 (not shown), these patterns in estimated SS and TP ratios most likely relate to diminished reservoir net trapping efficiency. The fact that TP ratios are generally greater than SS ratios suggests that decreasing retention in recent years is more pronounced for the finer (and more P-enriched) sediments, as should be expected since these fine sediments are more likely to be less well retained and more easily remobilized as reservoir infill continues.

In regard to TN, Figures 3.3c and 3.3f reveal that O/I has been generally stable with a median ~1.0, but with a notable rise in recent years. Importantly, however, Figure 3.3c and Figure B15i-B15l also reveal substantial numbers of low-level excursions above 1.0, particularly in recent years. The different patterns in TP and TN O/I ratios reflect effective trapping of TP, which is dominated by PP, and generally inefficient removal of TN, which is dominated by dissolved N (DN) and particularly nitrate (Langland and Hainly, 1997; Hirsch, 2012; Zhang *et al.*, 2013). (Although removal of nitrate by algal

growth and deposition might occur and some denitrification is possible, deposited particulate N [PN] can also be recycled as DN through bacterial activity in sediments (Kemp *et al.*, 1990; Testa *et al.*, 2013). Because of DN dominance, TN ratios are not far below 1.0 and the annual numbers of excursion based on modeled estimates are actually much higher for TN than for SS and TP (Figure B15). Differences in settling properties of PN- and PP-laden particles could also affect results.<sup>10</sup> In general, both the recent rise in TN O/I and the rising number of excursions above 1.0 reflect an increasingly larger quantity and fraction of PN in the reservoir output that deserves further study and management consideration.

One potential concern in regard to the ratio trends discussed above is that sampling dates varied between Marietta and Conowingo, with more highflow dates sampled at Conowingo, and that WRTDS surfaces may vary among sites in a way that could bias results. We examined the potential impact of this sampling issue by considering equally censored data at both sites. (See details in Appendix B4-IV and Figures B16-B18 therein.) Results show that our basic conclusions are unaffected by differences in sampling patterns (Figure B19).

### ***3.3.2.3. Changes in Reservoir Trapping Efficiency as a Function of Flow***

The above-noted changes in reservoir trapping have attracted considerable attention from managers and there has been a popular tendency to consider extreme storm events as the major concern for the future (The Lower Susquehanna River Watershed

---

<sup>10</sup> Relative to our published paper (Zhang *et al.*, 2016), the underlined text is a more accurate statement of the intended concept.

Assessment Team, 2014). While extreme flows have indeed been important sources of sediment discharge, decreased net trapping by the reservoir is a more complex story that is relevant under many flow conditions. To better understand this problem, we evaluated sediment and nutrient loadings in five flow classes defined according to percentiles of daily Conowingo discharge, namely, Q<sub>1</sub> (0<sup>th</sup> to 25<sup>th</sup> percentile), Q<sub>2</sub> (25<sup>th</sup> to 50<sup>th</sup>), Q<sub>3</sub> (50<sup>th</sup> to 75<sup>th</sup>), Q<sub>4</sub> (75<sup>th</sup> to 99.5<sup>th</sup>), and Q<sub>5</sub> (99.5<sup>th</sup> to 100<sup>th</sup>). Note that Q<sub>1</sub> to Q<sub>4</sub> each contain ~25% of the days in 1987-2013 but these flows are well below the literature-reported scour threshold of 11,300 m<sup>3</sup>/s (Gross *et al.*, 1978). Q<sub>5</sub> contains the days with the 0.5% highest flows (50 out of 9,862 days in the period of record), ranging from 7,674 to 20,077 m<sup>3</sup>/s, with the high end representing the highest daily discharge observed during TS Lee in 2011.

This flow-classified analysis reveals that SS and TP O/I ratios have increased since the early 2000s across all classes, as shown in Figures B20 and B21 (Appendix B4-V), respectively, and despite the stable ~1.0 ratio for flow. For TN, the general loading trends have been slightly downward for both input and output over the most record of 1987-2013 (Figure B22). TN O/I, however, has risen slightly since ~2000 for Q<sub>1</sub>-Q<sub>4</sub> and greatly for Q<sub>5</sub>. The latter reflects the fact that PN losses have become much more significant in recent years, as confirmed through separate analysis of PN (data not shown). Because time averaging is not possible for flow-class data, daily-based ratios are shown in Figures B20-B22. As previously noted, our comparative analysis suggests that these ratios also accurately reveal the broad trend in net deposition, despite the travel time issue.

In the above regard, it is noteworthy that Q<sub>4</sub> contains a large number of highflow days that, although well below the previously-documented scour threshold, represent the

highest cumulative discharge volume ( $V_w$ ) and cumulative loading for almost all constituents for both reservoir input and output in 1987-2013 (Figure 3.4). For input (Figure 3.4a), Q4 has contributed over half of  $V_w$  (55%), SS (64%), TP (59%), and TN (55%). For output (Figure 3.4b), Q4 has contributed over half of  $V_w$  (54%), TN (53%), and TP (53%), although its SS contribution (35%) is less than Q5 (58%) due to very high estimates of SS load for extremely highflow days that belong to Q5. Interestingly, Q5's contribution of SS at the outlet (58%) is much higher than that of TP (25%), owing to the facts that SS during high flows have a greater percentage of larger size fractions than at

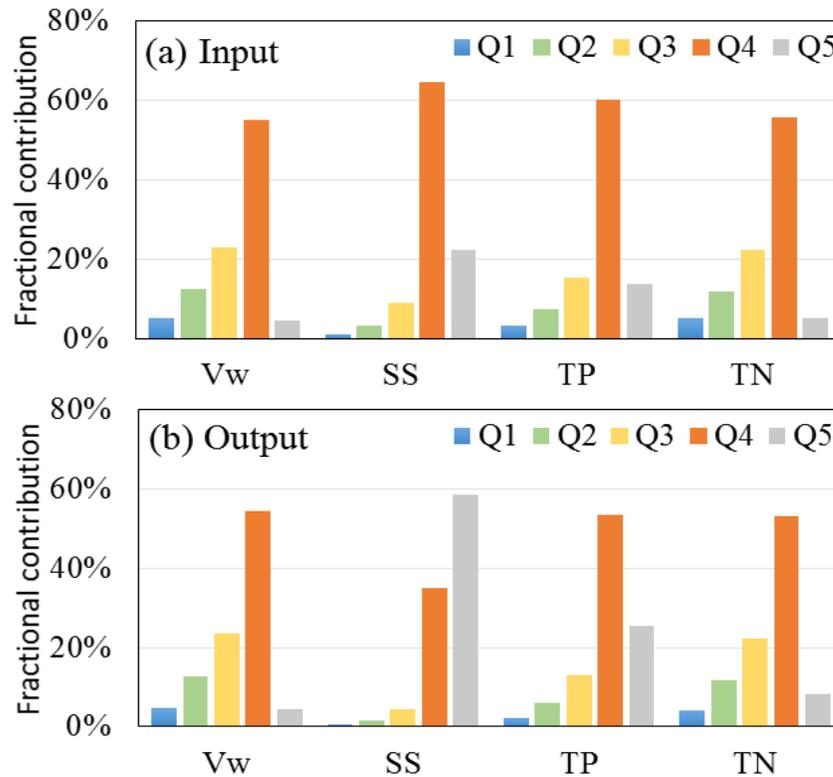


Figure 3.4. Fractional contribution of total streamflow volume ( $V_w$ ) and nutrient and sediment mass discharges in the Susquehanna River at Conowingo by each of the five flow classes for reservoir input (a) and output (b) in 1987-2013. (Ranges of the five flow classes – Q1: 25~396  $m^3/s$ ; Q2: 399~787  $m^3/s$ ; Q3: 790~1,464  $m^3/s$ ; Q4: 1,467~7,646  $m^3/s$ ; Q5: 7,674~20,077  $m^3/s$ .)

low flows and that larger particles have lower specific surface area and lower P:SS ratios. Nonetheless, our major conclusion from this analysis is that Q<sub>4</sub> has dominated the absolute mass of nutrient delivery and has contributed a major part of SS delivery through the reservoir, despite the fact that flows in Q<sub>4</sub> are generally insufficient to cause major scour.

### **3.3.3. Changes in Net Deposition: Analysis of Loads from Stationary WRTDS**

#### **Models**

In this section, we expand our analysis to consider cases where we force our WRTDS regression surfaces to be *stationary* (Section 3.2.3). The objective is to remove potential effects of inter-annual streamflow variability on the validity of our conclusions in regard to changes in reservoir trapping efficiency. For brevity, the 1990-, 2000-, and 2010-surface based stationary models are hereafter referred to as M<sub>1</sub>, M<sub>2</sub>, and M<sub>3</sub>, respectively.

#### ***3.3.3.1. Frequency Plots of Ranked Loadings***

One way of comparing the three stationary models is to rank the estimated loadings and count the number of days per year that these loadings exceed certain values under each model. Such probability-of-exceedance plots are shown for all constituents at each count value (*i.e.*, 0-365 days/year) in Figures B23-B31 of Appendix B5-I. These figures also contain three enlarged portions for visual clarity, that is, 0-15, 100-115, and 200-215 days/year. For brevity, we focus on the beginning portion (*i.e.*, the highest 15 days/year); these sub-plots are shown in Figure 3.5.

For SS at Conowingo (reservoir output), M<sub>3</sub> shows dramatically higher loading than M<sub>1</sub> and M<sub>2</sub> for the highest 15 days/year (Figure 3.5a). For example, Conowingo (reservoir outlet) SS loading is estimated to exceed 50 million kg/day with a frequency of

3, 4, and 7 days/year under M<sub>1</sub>, M<sub>2</sub>, and M<sub>3</sub>, respectively. Moreover, the uncertainty bands of M<sub>3</sub> (shown as red dashed lines) are always above those of M<sub>2</sub> and M<sub>1</sub> (blue and green dashed lines), indicating that the noted rise with M<sub>3</sub> is robust. Further along the x-axis, M<sub>3</sub> becomes less distinguishable from M<sub>1</sub> and M<sub>2</sub> (Figure B23, c-d). For Marietta (reservoir inlet), M<sub>3</sub> is always below M<sub>1</sub> and M<sub>2</sub> (Figures 3.5b, B24), indicating declined loading based on the most recent watershed condition (2010) than prior conditions (1990 and 2000). These results are consistent with the aforementioned conclusion that upstream watershed load has declined even while outlet load has risen. For net rates of SS storage (*i.e.*, net deposition), M<sub>3</sub> has evidently decreased in relative to M<sub>1</sub> and M<sub>2</sub> for the highest 15 days/year (Figure 3.5c). For instance, under M<sub>1</sub>, SS net deposition exceeds 115 million kg/day with a frequency of two days/year. At this same frequency the net deposition exceeds 110 million kg/day under M<sub>2</sub> and 40 million kg/day under M<sub>3</sub>. Further along the x-axis, the M<sub>3</sub> model also shows less net deposition than the other two models (Figure B25, c-d).

Results for TP are generally similar to those of SS (Figures 3.5d-3.5f). For example, Conowingo TP loading is estimated to exceed 50,000 kg/day on about 5 days/year under M<sub>1</sub>, 6 days/year under M<sub>2</sub>, and 10 days/year under M<sub>3</sub> (Figure 3.5d). For reservoir input, M<sub>3</sub> is always below M<sub>1</sub> and M<sub>2</sub> (Figures 3.5e, B27), indicating decreased TP loading based on the most recent watershed condition (2010) than prior conditions. For TP net deposition, M<sub>3</sub> has evidently decreased in relative to M<sub>1</sub> and M<sub>2</sub> (Figure 3.5f). For instance, out of 2 days per year, net TP deposition exceeds 85,000 kg/day under M<sub>1</sub>, 90,000 kg/day under M<sub>2</sub>, but only 25,000 kg/day under M<sub>3</sub>. Further along the x-axis, M<sub>3</sub> also shows less net deposition than the other two models (Figures B28, c-d).

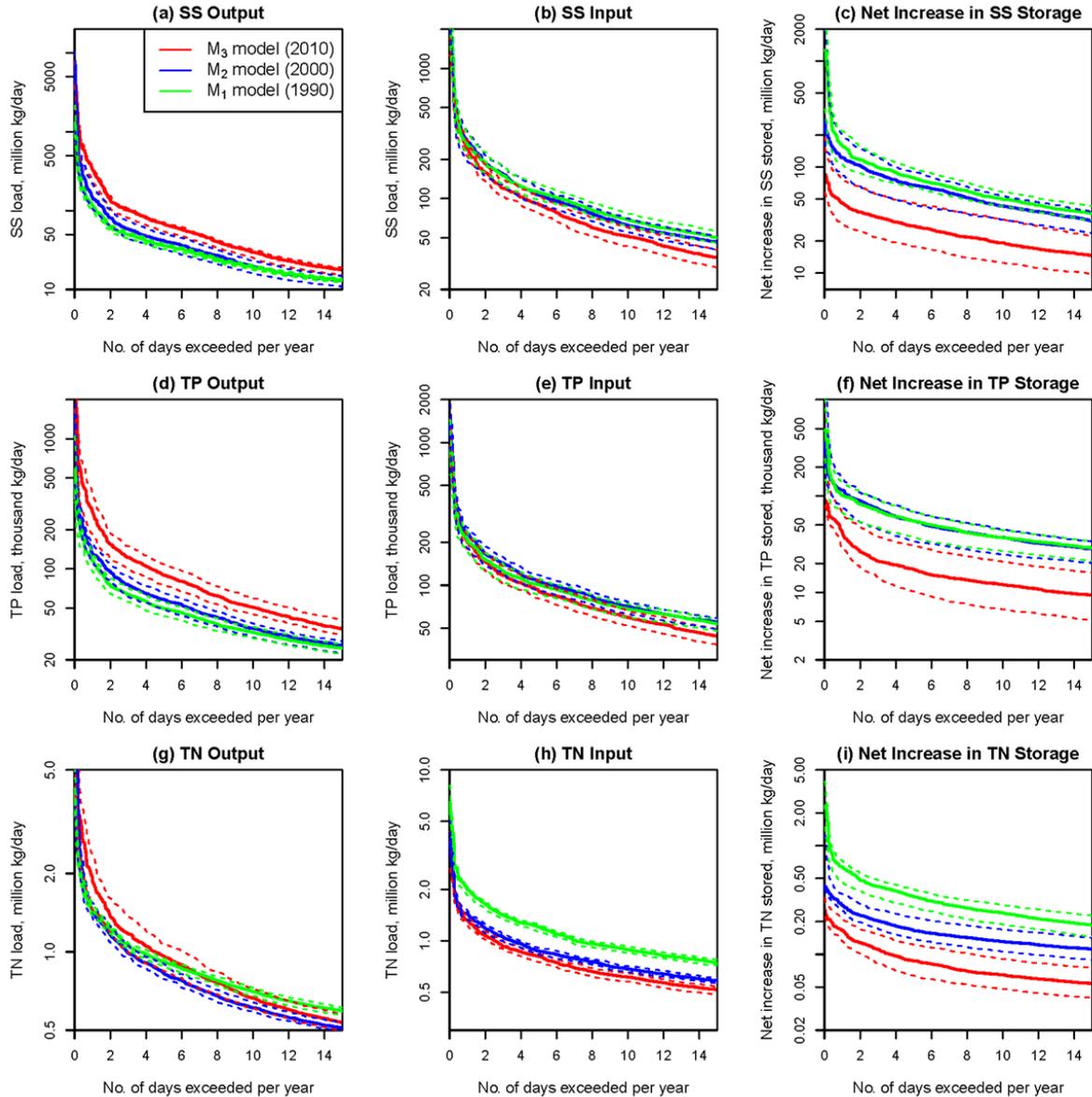


Figure 3.5. Frequency plots of ranked loadings for reservoir output, input, and net increase in storage for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) based on the three historical stationary WRTDS models. For output (panels a, d, g) and input (panels b, e, h), dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. For net increase in storage (panels c, f, i), dashed lines indicate the 95% confidence intervals based on 100 sets of net storage estimates obtained from the resulting 10x10 input x output matrix. See Figures B23–B31 in Appendix B5-I for the full range of x-axis (*i.e.*, 0–365 days/year).

Results for TN do not follow those for SS and TP at Conowingo – the three models show generally similar loadings, except that M<sub>3</sub> is above M<sub>1</sub> and M<sub>2</sub> for the range of 0-6 days/year (Figure 3.5g). Further along the x-axis, M<sub>3</sub> and M<sub>2</sub> become indistinguishable but are both lower than M<sub>1</sub> (Figure B29, c-d). For reservoir input, TN loading generally follows the pattern of M<sub>1</sub> > M<sub>2</sub> > M<sub>3</sub> (Figures 3.5h, B30), suggesting recent TN load decline in the upstream watershed. For TN net deposition, the pattern is similar to that of SS or TP, with M<sub>3</sub> showing less net deposition than the other two models (Figure 3.5i and Figure B31).

Overall, the stationary-model results are consistent with the other indications, suggesting that: (1) reservoir inputs of SS, TP, and TN have generally declined, (2) reservoir outputs of SS and TP have increased, and (3) reservoir net deposition of SS and TP has declined considerably. Such estimates serve to better quantify changes in reservoir trapping capability and can be useful to inform the refinement of Chesapeake Bay Watershed Model for purposes of Chesapeake Bay TMDL allocation (Linker *et al.*, 2013; Shenk and Linker, 2013). Although this retrospective analysis does not speak to the issue of future conditions, it can serve to inform and constrain future analysis and modeling, which is a research topic of great importance to Chesapeake Bay restoration.

### ***3.3.3.2. Changes in Reservoir Trapping Efficiency as a Function of Flow***

As with the standard WRTDS analysis (Section 3.3.2.2), we find it useful to consider the stationary-model results in regard to the flow interval associated with major changes in reservoir function. Thus, we similarly evaluated loadings as a function of discharge for stationary-model estimates and fitted LOWESS curves to the modeled loading *vs.* observed discharge. These plots are provided in Figures B32-B40 of Appendix B5-II.

Readers should appreciate that these statistically modeled loading-discharge relationships are highly uncertain at the highest discharges (*i.e.*, those associated with infrequent weather-related events) due to the scarcity of concentration monitoring data at one or more locations during these discharges and the related fact that available measurements may not accurately represent the proper flow-weighted distribution of concentration. Although full discussion of this issue is beyond our current scope, we nonetheless caution against over-interpretation of model results at such extreme discharges. In our discussion below, we focus on the broader range of discharges where more data are available and with special focus on SS and TP because these constituents are most sensitive to reservoir trapping.

For SS at Conowingo (reservoir output), the  $M_3$  curve is clearly above  $M_1$  and  $M_2$  for discharge above  $\sim 2,500 \text{ m}^3/\text{s}$  but less distinguishable at lower discharge (Figure B32). For reservoir input, however, differences among the three models are observed mainly for discharge below  $\sim 5,000 \text{ m}^3/\text{s}$ , with  $M_3$  loading being the lowest (Figure B33). For reservoir net deposition, the  $M_3$  curve is clearly below  $M_1$  and  $M_2$  for a wide range of discharges ( $> \sim 150 \text{ m}^3/\text{s}$ ) (Figure B34).

Results for TP are similar to those for SS, but with the clear separation of curves extending down to  $\sim 1,700 \text{ m}^3/\text{s}$  for Conowingo (Figure B35). And as with SS, input of TP under  $M_3$  shows smaller loading than under  $M_1$  and  $M_2$  for discharge below  $\sim 5,000 \text{ m}^3/\text{s}$  (Figure B36). Finally, reservoir net deposition of TP is smaller under  $M_3$  than  $M_1$  and  $M_2$  for the entire range of discharges (Figure B37).

Overall, these loading-discharge relationships confirm the previous finding (Section 3.3.2.3) that diminished reservoir trapping of SS and TP has occurred under a wide range

of flow conditions, including flows well below the literature-reported scour threshold (Gross *et al.*, 1978). Therefore, even if future Conowingo discharge remains largely below this threshold, there will likely be continued decreased net trapping and increased loadings for SS and TP at Conowingo Dam. Moreover, these stationary-model results confirm that the observed changes could have occurred because of diminished reservoir trapping efficiency and there is no need to invoke climatic factors such as increased streamflow variability to explain the changes.

### ***3.3.3.3. Cumulative Loading for Selected Representative Wet, Average, and Dry Years***

The developed stationary models can also be used to explore estimated cumulative increases in storage (*i.e.*, cumulative net deposition) for different types of hydrological conditions, that is, “wet,” “average,” and “dry” years. Toward that end, we have selected 2003, 2007, and 2001, as representative years for wet, average, and dry conditions, respectively. These years had average flows of 1718, 1009, and 667 m<sup>3</sup>/s, and represent the 4<sup>th</sup>, 13<sup>th</sup>, and 27<sup>th</sup> highest-flow years, respectively, based on ranking of annual Conowingo streamflow between 1987 and 2013. We have used the actual discharge data and the three developed models to estimate net deposition under each condition. These results are provided in Figure 3.6 and Figures B41-B43 of Appendix B5-III.

For SS net deposition under the wet year scenario (2003), M<sub>3</sub> shows less cumulative net deposition than the other two models (Figure 3.6a). Moreover, the upper confidence limit of M<sub>3</sub> (upper red dashed line) is always below the lower confidence limits of M<sub>2</sub> and M<sub>1</sub> (lower blue and green dashed lines), suggesting a statistically significant difference of reservoir performance with the M<sub>3</sub> model. Similar patterns were also observed with the average year (2007) and dry year (2001) scenarios (Figures 3.6b-3.6c).

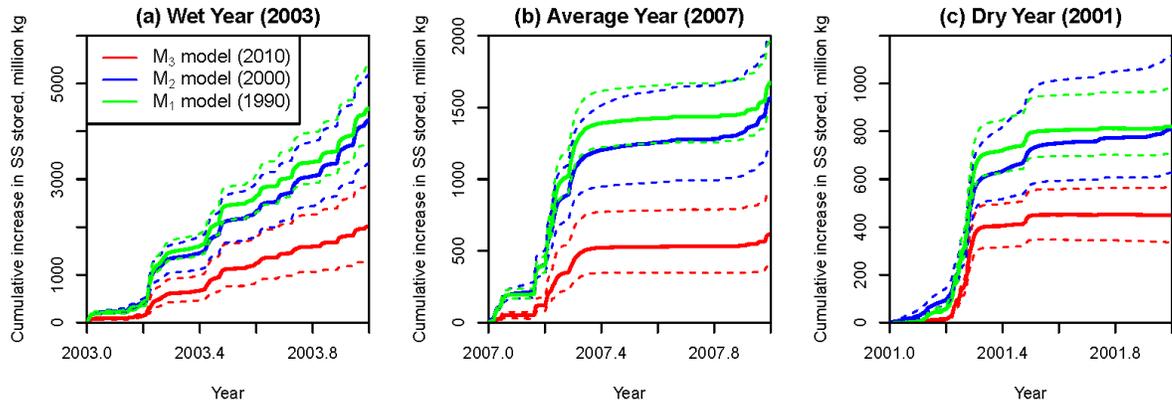


Figure 3.6. Modeled cumulative reservoir storage over the course of three selected wet, average, and dry calendar years (*i.e.*, 2003, 2007, and 2001) with respect to suspended sediment (SS) loads based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. See Figure B41 in Appendix B5-III for results on SS output and input loadings.

Not surprisingly, absolute values of cumulative storage increase with wetness of the year, reflecting larger incoming loads for deposition during wet years. (Note the difference of *y-axis* scale among the three panels in Figure 3.6.) Similar patterns were also observed for TP net deposition modeling (Figure B42). Overall, these year-specific comparisons support our major conclusion of decreased net deposition under different hydrological conditions.

Finally, a note about extreme events -- although one may be tempted to similarly reconstruct cumulative storage for the truly wettest years (*e.g.*, 2011, 2004, or 1996 in this record), these years are associated with “extreme flow events” (*i.e.*, TS Lee, Hurricane Ivan, and a major upstream ice-jam release, respectively). For discharges during these extreme flow events, concentration data are relatively sparse and perhaps

not fully representative of the range associated with the full history of the events' hydrographs. Statistical modeling has especially high uncertainty in such cases and we therefore avoided these years for the purpose of Figure 3.6.

### **3.4. Management Implications**

This paper presents three types of analyses on decadal-scale changes in sediment and nutrient loadings across Conowingo Reservoir on Susquehanna River --  $C-Q$  relationships inferred from observed data, loading estimates from standard WRTDS models, and loading estimates from stationary WRTDS models. All three analyses consistently show that average annual net deposition rates of SS and TP have declined in recent years under a wide range of flow conditions that include flows well below the literature-reported scour threshold, with correspondingly increased loads delivered from Susquehanna River to Chesapeake Bay (for any given flow) relative to previous decades. Future progress in Chesapeake Bay restoration will depend on accurate predictions of how inputs of SS, TP, and TN to the reservoirs will be modulated by processes taking place in the reservoirs. Management actions in the Susquehanna River basin will need to be adjusted to reflect the future role that sediment accumulation and remobilization behind Conowingo Dam will have on the delivery of SS, TP, and TN to the Bay. Our analysis of the evolution of the system to date can help constrain and inform the development and application of improved predictive models of reservoir performance, and particularly the incorporation of such models in the ongoing upgrade of the Chesapeake Bay Partnership's Watershed Model (Linker *et al.*, 2013; Shenk and Linker, 2013). Our results and methods are also applicable to other reservoir systems that may be similarly approaching a state of dynamic equilibrium with respect to sediment storage.

### 3.5. Supporting Information

Supporting information to this chapter is provided in Appendix B.

### 3.6. Acknowledgements

This work was supported by Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090), Maryland Water Resources Research Center (2015MD329B), National Science Foundation (CBET-1360415), and U.S. Geological Survey Chesapeake Bay Ecosystem Program. We thank Wendy McPherson for providing the photo of Conowingo Dam. We thank Joel Blomquist (USGS), Ken Staver (University of Maryland), and three anonymous reviewers for their constructive comments.

### 3.7. Literature Cited

- Belval, D. L. and L. A. Sprague, 1999. Monitoring nutrients in the major rivers draining to Chesapeake Bay. U.S. Geological Survey Water-Resources Investigations Report 99-4238, p. 8. [http://va.water.usgs.gov/online\\_pubs/WRIR/99-4238/wrir\\_99\\_4238\\_text.pdf](http://va.water.usgs.gov/online_pubs/WRIR/99-4238/wrir_99_4238_text.pdf).
- Brakebill, J. W., S. W. Ator and G. E. Schwarz, 2010. Sources of suspended-sediment flux in streams of the Chesapeake Bay Watershed: A regional application of the SPARROW model. *J. Am. Water Resour. Assoc.* 46:757-776, DOI: 10.1111/j.1752-1688.2010.00450.x.
- Burt, T. P., F. Worrall, N. J. K. Howden and M. G. Anderson, 2015. Shifts in discharge-concentration relationships as a small catchment recovers from severe drought. *Hydrol. Process.* 29:498-507, DOI: 10.1002/hyp.10169.

- Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.
- Evans, C. and T. D. Davies, 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. *Water Resour. Res.* 34:129-137, DOI: 10.1029/97WR01881.
- Friedl, G. and A. Wüest, 2002. Disrupting biogeochemical cycles—Consequences of damming. *Aquat. Sci.* 64:55-65, DOI: 10.1007/s00027-002-8054-0.
- Gellis, A. C., C. R. Hupp, M. J. Pavich, J. M. Landwehr, W. S. L. Banks, B. E. Hubbard, M. J. Langeland, J. C. Ritchie and J. M. Reuter, 2008. Sources, Transport, and Storage of Sediment at Selected Sites in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2008-5186, Reston, VA, p. 95. <http://pubs.usgs.gov/sir/2008/5186/>.
- Gomez and Sullivan Engineers, 2012. 2011 Conowingo Pond Bathymetric Survey, Appendix F of Final Study Report, Sediment Introduction and Transport Study, Conowingo Hydroelectric Project. p. 129.
- Gross, M. G., M. Karweit, W. B. Cronin and J. R. Schubel, 1978. Suspended sediment discharge of the Susquehanna River to northern Chesapeake Bay, 1966 to 1976. *Estuaries* 1:106-110, <http://www.springerlink.com/index/r4796375t48061x3.pdf>.
- Hagy, J. D., W. R. Boynton, C. W. Keefe and K. V. Wood, 2004. Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658, DOI: 10.1007/bf02907650.
- Hainly, R. A., L. A. Reed, H. N. J. Flippo and G. J. Barton, 1995. Deposition and

simulation of sediment transport in the Lower Susquehanna River reservoir system. U.S. Geological Survey U.S. Geological Survey Water-Resources Investigations Report 95-4122, Lemoyne, PA, p. 39.

<http://pubs.er.usgs.gov/publication/wri954122>.

Helsel, D. R. and R. M. Hirsch, 2002. Statistical Methods in Water Resources. *U.S. Geological Survey Techniques of Water-Resources Investigations Book 4, Chapter A3*. U.S. Geological Survey, Reston, VA, p. 522.

<http://pubs.usgs.gov/twri/twri4a3/>.

Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.

Hirsch, R. M., 2014. Large Biases in Regression-Based Constituent Flux Estimates: Causes and Diagnostic Tools. *J. Am. Water Resour. Assoc.* 50:1401-1424, DOI: 10.1111/jawr.12195.

Hirsch, R. M., R. B. Alexander and R. A. Smith, 1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resour. Res.* 27:803-813, DOI: 10.1029/91WR00259.

Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.

Hirsch, R. M. and L. De Cicco, 2015. User guide to Exploration and Graphics for RivEr

Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). U.S. Geological Survey Techniques and Methods Book 4, Chapter A10, Reston, VA, p. 93. <http://dx.doi.org/10.3133/tm4A10>.

Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.

Horowitz, A. J., V. C. Stephens, K. A. Elrick and J. J. Smith, 2012. Concentrations and annual fluxes of sediment-associated chemical constituents from conterminous US coastal rivers using bed sediment data. *Hydrol. Process.* 26:1090-1114, DOI: 10.1002/hyp.8437.

House, W. A. and M. S. Warwick, 1998. Hysteresis of the solute concentration/discharge relationship in rivers during storms. *Water Res.* 32:2279-2290, DOI: 10.1016/S0043-1354(97)00473-9.

Jossette, G., B. Leporcq, N. Sanchez and Philippon, 1999. Biogeochemical mass-balances (C, N, P, SI) in three large reservoirs of the Seine basin (France). *Biogeochemistry* 47:119-146, DOI: 10.1007/bf00994919.

Karl, T. R. and R. W. Knight, 1998. Secular Trends of Precipitation Amount, Frequency, and Intensity in the United States. *Bull. Am. Meteorol. Soc.* 79:231-241, DOI: 10.1175/1520-0477(1998)079<0231:STOPAF>2.0.CO;2.

Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C.

- Stevenson, 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* 303:1-29, DOI: 10.3354/meps303001.
- Kemp, W. M., P. Sampou, J. Caffrey, M. Mayer, K. Henriksen and W. R. Boynton, 1990. Ammonium recycling versus denitrification in Chesapeake Bay sediments. *Limnol. Oceanogr.* 35:1545-1563, <http://www.jstor.org/stable/10.2307/2837739>.
- Langland, M. J., 2009. Bathymetry and sediment-storage capacity change in three reservoirs on the Lower Susquehanna River, 1996-2008. U.S. Geological Survey Scientific Investigations Report 2009-5110, Reston, VA, p. 21. <http://pubs.usgs.gov/sir/2009/5110/>.
- Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18. <http://dx.doi.org/10.3133/ofr20141235>.
- Langland, M. J. and R. A. Hainly, 1997. Changes in bottom-surface elevations in three reservoirs on the lower Susquehanna River, Pennsylvania and Maryland, following the January 1996 flood - implications for nutrient and sediment loads to Chesapeake Bay. U.S. Geological Survey Water-Resources Investigations Report 97-4138, Lemoyne, PA, p. 34. <http://pa.water.usgs.gov/reports/wrir97-4138.pdf>.
- Lefrançois, J., C. Grimaldi, C. Gascuel-Oudoux and N. Gilliet, 2007. Suspended sediment and discharge relationships to identify bank degradation as a main sediment source on small agricultural catchments. *Hydrol. Process.* 21:2923-2933, DOI: 10.1002/hyp.6509.

- Linker, L. C., R. A. Batiuk, G. W. Shenk and C. F. Cerco, 2013. Development of the Chesapeake Bay Watershed Total Maximum Daily Load Allocation. *J. Am. Water Resour. Assoc.* 49:986-1006, DOI: 10.1111/jawr.12105.
- Massoudieh, A., A. Gellis, W. S. Banks and M. E. Wieczorek, 2013. Suspended sediment source apportionment in Chesapeake Bay watershed using Bayesian chemical mass balance receptor modeling. *Hydrol. Process.* 27:3363-3374, DOI: 10.1002/hyp.9429.
- Merritts, D., R. Walter, M. Rahnis, J. Hartranft, S. Cox, A. Gellis, N. Potter, W. Hilgartner, M. Langland, L. Manion, C. Lippincott, S. Siddiqui, Z. Rehman, C. Scheid, L. Kratz, A. Shilling, M. Jenschke, K. Datin, E. Cranmer, A. Reed, D. Matuszewski, M. Voli, E. Ohlson, A. Neugebauer, A. Ahamed, C. Neal, A. Winter and S. Becker, 2011. Anthropocene streams and base-level controls from historic dams in the unglaciated mid-Atlantic region, USA. *Philosophical Transactions of the Royal Society A* 369:976-1009, DOI: 10.1098/rsta.2010.0335.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Murphy, R. R., W. M. Kemp and W. P. Ball, 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* 34:1293-1309, DOI: 10.1007/s12237-011-9413-7.
- Qian, S. S., 2010. *Environmental and Ecological Statistics with R*. Boca Raton, Florida, Chapman and Hall/CRC Press, ISBN 9781420062069

- Reed, L. A. and S. A. Hoffman, 1997. Sediment Deposition in Lake Clarke, Lake Aldred, and Conowingo Reservoir, Pennsylvania and Maryland, 1910-93. US Geological Survey Water-Resources Investigations Report 96-4048, Lemoyne, PA, p. 14.  
<http://pubs.usgs.gov/wri/1996/4048/report.pdf>.
- Richardson, M. C., 2012. Contributions of streamflow variability, concentration-discharge shifts and forested wetlands to terrestrial-aquatic solute export in Precambrian Shield headwater catchments. *Ecohydrology* 5:596-612, DOI: 10.1002/eco.244.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- Sprague, L. A., M. J. Langeland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. U.S. Geological Survey Water-Resources Investigations Report 00-4218, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).
- Susquehanna River Basin Commission, 2014. Sediment and nutrient assessment program.  
<http://www.srbc.net/programs/cbp/nutrientprogram.htm>.
- Testa, J. M., D. C. Brady, D. M. Di Toro, W. R. Boynton, J. C. Cornwell and W. M. Kemp, 2013. Sediment flux modeling: Simulating nitrogen, phosphorus, and silica cycles. *Estuar. Coast. Shelf Sci.* 131:245-263, DOI: 10.1016/j.ecss.2013.06.014.
- The Lower Susquehanna River Watershed Assessment Team, 2014. Lower Susquehanna

River Watershed Assessment, MD and PA: Phase I Draft Report. p. 185.

<http://mddnr.chesapeakebay.net/LSRWA/report.cfm>.

U.S. Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment. Annapolis, MD.

<http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html>.

U.S. Geological Survey, 2014. Surface-water data for the nation.

Walter, R. C. and D. J. Merritts, 2008. Natural streams and the legacy of water-powered mills. *Science* 319:299-304, DOI: 10.1126/science.1151716.

Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI: 10.1016/j.scitotenv.2013.02.012.

Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.

## Chapter 4. Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns<sup>11</sup>

### Abstract

The export of nitrogen (N), phosphorus (P), and suspended sediment (SS) is a long-standing management concern for the Chesapeake Bay watershed, USA. Here I present a comprehensive evaluation of nutrient and sediment loads over the last three decades at multiple locations in the Susquehanna River basin (SRB), Chesapeake's largest tributary watershed. Sediment and nutrient riverine loadings, including both dissolved and particulate fractions, have generally declined at all sites upstream of Conowingo Dam (non-tidal SRB outlet). Period-of-record declines in riverine yield are generally smaller than those in source input, suggesting the possibility of legacy contributions. Consistent with other watershed studies, these results reinforce the importance of considering lag time between the implementation of management actions and achievement of river quality improvement. Whereas flow-normalized loadings for particulate species have increased recently below Conowingo Reservoir, those for upstream sites have declined, thus substantiating conclusions from prior studies about decreased reservoir trapping

---

<sup>11</sup> This chapter (Abstract through Section 4.8) has been published as: Zhang, Q.; Ball, W. P.; Moyer, D. L., Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns. *Sci. Total Environ.* **2016**, 563-564: 1016-1029, doi: [10.1016/j.scitotenv.2016.03.104](https://doi.org/10.1016/j.scitotenv.2016.03.104). Co-author Ball was involved in hypothesis development, study design, results interpretation and editing. Co-author Moyer was involved in results interpretation and editing. Copyright 2016 Elsevier. Reproduced/modified by permission of Elsevier. All figures, tables, and data were created by Qian Zhang unless otherwise indicated. Section, table, and figure numbers have been modified. References have also been re-formatted for consistency among chapters. Some minor substantive changes from the published manuscript have been made and are identified with footnotes.

efficiency. In regard to streamflow effects, statistically significant log-linear relationships between annual streamflow and annual constituent load suggest the dominance of hydrological control on the inter-annual variability of constituent export. Concentration-discharge relationships revealed general chemostasis and mobilization effects for dissolved and particulate species, respectively, both suggesting transport-limitation conditions. In addition to affecting annual export rates, streamflow has also modulated the relative importance of dissolved and particulate fractions, as reflected by its negative correlations with dissolved P/total P, dissolved N/total N, particulate P/SS, and total N/total P ratios. For land-use effects, period-of-record median annual yields of N, P, and SS all correlate positively with the area fraction of non-forested land but negatively with that of forested land under all hydrological conditions. Overall, this work has informed understanding with respect to four major factors affecting constituent export (*i.e.*, source input, reservoir modulation, streamflow, and land use) and demonstrated the value of long-term river monitoring.

#### **4.1. Introduction**

Chesapeake Bay has experienced persistent summer hypoxia conditions that have been attributed to a combination of excessive nutrient and sediment inputs from its watershed (Kemp *et al.*, 2005; Shenk and Linker, 2013) and density-driven vertical stratification (Pritchard and Schubel, 2001; Murphy *et al.*, 2011). Reduction of nitrogen (N), phosphorus (P), and sediment loadings has therefore been a long-standing focus of Chesapeake watershed management, which has been reinforced recently with the promulgation of Chesapeake Bay Total Maximum Daily Loads (TMDLs) (U.S.

Environmental Protection Agency, 2010) and state-wide efforts to establish watershed implementation plans (Linker *et al.*, 2013a; Shenk and Linker, 2013).

Among the tributaries to Chesapeake Bay, Susquehanna River is the largest (Hagy *et al.*, 2004; Murphy *et al.*, 2011; Zhang *et al.*, 2015) (Figure 4.1). It is one of nine major tributaries that collectively account for over 90% of river input above the head of tides and which have a total non-tidal drainage area that accounts for ~78% of the bay's total watershed area (Langland *et al.*, 1995; Belval and Sprague, 1999). Of this total non-tidal drainage, the Susquehanna River contributed 62% of streamflow, 65% of total N (TN), 46% of total P (TP), and 41% of suspended sediment (SS) between 1979 and 2012 (Zhang *et al.*, 2015). The relatively lower fractional contributions of TP and SS reflect historical trapping by the Lower Susquehanna River Reservoir System (LSRRS). This system, however, has become less capable of retaining sediment and particulate nutrients in recent years as it approaches sediment storage capacity (Hirsch, 2012; Zhang *et al.*, 2013; Zhang *et al.*, 2016).

For rivers with concentration-discharge monitoring data, water-quality analyses have often focused on how concentration varies with not only time, discharge, and season, but also with changes in source inputs (*e.g.*, fertilizer, manure, point sources, and atmospheric deposition), system function (*e.g.*, reservoirs), land use, and hydro-climatic factors (Sprague *et al.*, 2000; Harris, 2001; Sobota *et al.*, 2009; Howarth *et al.*, 2012). For the Chesapeake watershed, a multi-party collaboration is underway within the Chesapeake Bay Program partnership to seek explanations for water-quality changes in Chesapeake tributaries (Keisman *et al.*, 2015). In this context, our work is intended as a cursory examination of various types of data available for the Susquehanna River basin

(SRB). Specifically, a recently developed statistical method for estimating daily loads was combined with some relatively simple and traditional mass-balance and concentration-discharge ( $C-Q$ ) approaches to (a) accurately estimate riverine concentration and loading based on sparse monitoring data; (b) evaluate riverine loading trends by better accommodating inter-annual streamflow variability; (c) examine relationships between estimated concentration and streamflow for categorizing export patterns; and (d) analyze factors affecting constituent export, *e.g.*, source inputs, reservoir modulation, streamflow, and land use. This work demonstrates that the traditional approaches, despite some important shortcomings, can nonetheless be useful toward understanding some of the most important patterns and controls of constituent export. While more sophisticated approaches have become available for evaluating riverine export (*e.g.*, Chen *et al.*, 2014; Green *et al.*, 2014; Ai *et al.*, 2015; Hale *et al.*, 2015) and others are under development (even among our own team), we have considered these methods as beyond the current scope, and reserve their application for future work.

In the above context, we have undertaken a comprehensive evaluation of (1) temporal trends of nutrient and sediment loadings at seven long-term monitoring locations and (2) spatial variations of nutrient and sediment budgets of major sub-basins in the SRB.

Specifically, we have focused on addressing four motivating questions:

- Q<sub>1</sub>: What have been the general patterns of temporal trends in riverine nutrient/sediment loadings? In particular, have trends been consistent (a) across all seven monitoring locations and (b) between dissolved and particulate species?
- Q<sub>2</sub>: What have been the general trends in watershed source inputs and how have their magnitudes compared with those in riverine loadings?

Q3: Which sub-basins have been net sources (or storages) of loadings and what has been the role of streamflow on constituent export?

Q4: How do sub-basins compare in regard to constituent yield (*i.e.*, loading per area) and how do differences relate to those in land use distribution?

This work offers a unique opportunity to understand these aspects for a large watershed. First, long-term river monitoring data are available at multiple locations across this watershed, which allowed combined temporal and spatial analyses. In addition, the well-documented data on watershed source inputs fostered a quantitative comparison of changes in source input and changes in riverine yield. Moreover, the contrasting land use distributions of the sub-basins facilitated an evaluation of land-use effect on export. Finally, we were also able to compare the upstream sub-basins with re-analysis of the previously studied downstream reservoir system (*i.e.*, the LSRRS) to better highlight and confirm temporal aspects of the LSRRS's modulation of loadings. Overall, this work should help inform the management of Chesapeake Bay's largest tributary and also foster comparisons with rivers in other geographical regions (within the Chesapeake watershed and beyond) for better understanding patterns and controls of constituent export.

## **4.2. Methods**

### **4.2.1. Study Area and Data**

The non-tidal SRB covers portions of New York, Pennsylvania, and Maryland, USA. (Figure 4.1). The basin's outlet at Conowingo Dam has been monitored by the U.S. Geological Survey (USGS) since the late 1970s. This site is about 10 miles from the river

mouth and receives 99% of the streamflow from the entire SRB (Belval and Sprague, 1999). Upstream and in Pennsylvania, six sites have been monitored by the Susquehanna River Basin Commission (SRBC) since the mid-1980s. Among them, Towanda, Danville, and Marietta monitor the mainstem of Susquehanna River, whereas Lewisburg, Newport, and Conestoga monitor the West Branch Susquehanna River, Juniata River, and Conestoga River, respectively, all of which are tributaries to the Susquehanna (Figure 4.1). Details of these sites are summarized in Table 4.1.

At each site, daily streamflow data were compiled from the USGS National Water Information System (NWIS) Web Interface (U.S. Geological Survey, 2014a). Water-quality concentration data were compiled from the USGS NWIS for Conowingo and from the SRBC website (Susquehanna River Basin Commission, 2014) for the six upstream sites. These data included concentration measurements for SS, TP, TN, dissolved phosphorus (DP), and dissolved nitrogen (DN). Temporal coverages of these water-quality samples are summarized in Table C1 (Appendix C). The average number of sampled days ranges between 25.5 and 40.4 per year. The samples at each site were collected across the full range of hydrological conditions in each year and comprised at least one sample in each month of the year and 8 targeted samples during times of stormflow (*i.e.*, periods of elevated discharge).

#### **4.2.2. Statistical Method for Loading Estimation: WRTDS**

To provide reasonable estimates of daily concentrations – as needed for estimation of daily and annual loads – it is necessary to augment these relatively sparse water-quality concentration data through statistical treatments. Typically, daily concentrations are estimated on the basis of correlations of concentration over three parameters: time,

## Susquehanna River Basin

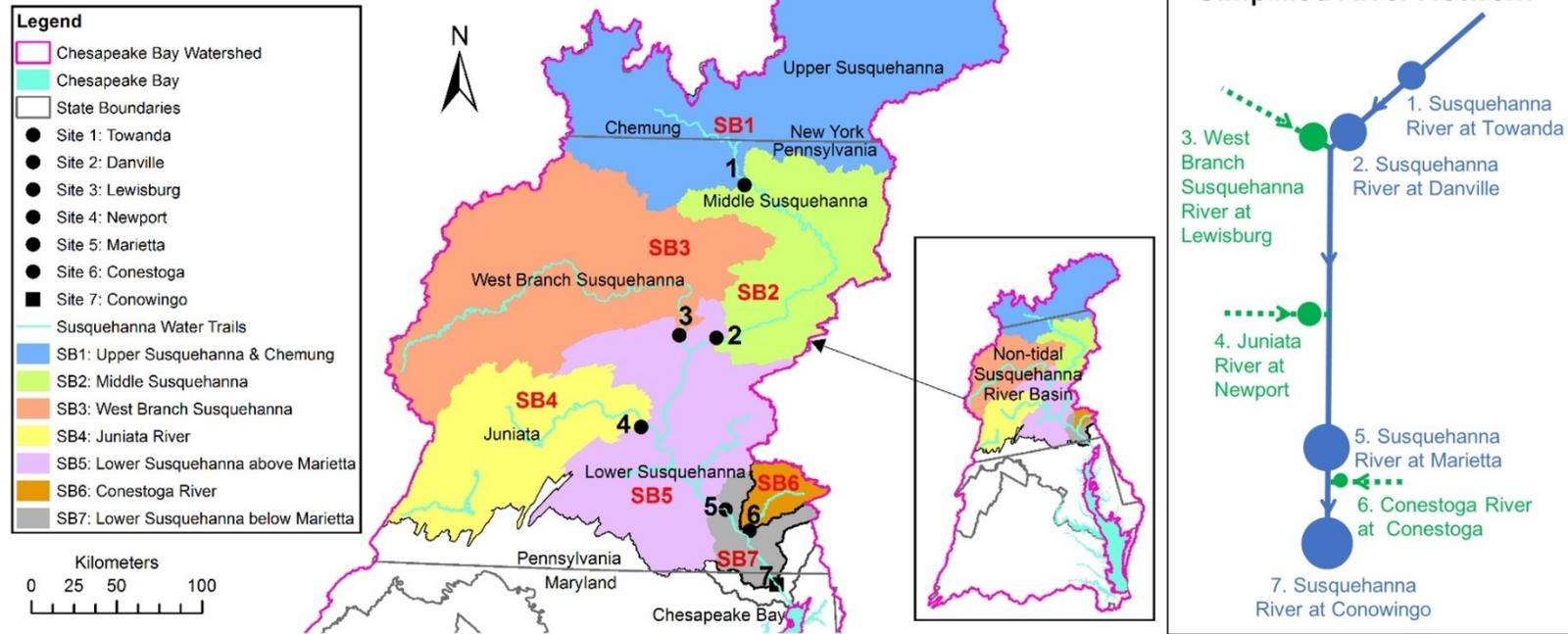


Figure 4.1. Map of the Susquehanna River basin (SRB), showing the seven long-term monitoring sites (No. 1-7) and the seven sub-basins (SB1-SB7). Conowingo (#7) is the non-tidal SRB's outlet. Inset shows the SRB's location within the Chesapeake Bay watershed. The diagram of "Simplified River Network" shows four sites on the river mainstem and three sites on the tributaries to Susquehanna River. See Table 4.1 for details of the sites and sub-basins.

Table 4.1. Details of the long-term monitoring sites and sub-basins in the Susquehanna River basin. <sup>a</sup>

Monitoring Sites							
Index in Figure 4.1	USGS Site Number	Site Name (Short Name)	Drainage Area, km <sup>2</sup>	Upstream Land Use (percent)			
				Urban	Agricultural	Forested	Other
1	01531500	Susquehanna River at Towanda, PA (Towanda)	20,194	4	35	60	1
2	01540500	Susquehanna River at Danville, PA (Danville)	29,060	5	33	60	2
3	01553500	West Branch Susquehanna River at Lewisburg, PA (Lewisburg)	17,734	2	15	81	2
4	01567000	Juniata River at Newport, PA (Newport)	8,687	2	28	69	1
5	01576000	Susquehanna River at Marietta, PA (Marietta)	67,314	4	30	64	2
6	01576754	Conestoga River at Conestoga, PA (Conestoga)	1,217	14	60	23	3
7	01578310	Susquehanna River near Conowingo, MD (Conowingo)	70,189	2	29	67	2
Sub-basins							
Index in Figure 4.1	Sub-basin Name	Calculation of Net Export	Drainage Area, km <sup>2</sup>	Upstream Land Use (percent)			
				Urban	Agricultural	Forested	Other
SB1	Upper Susquehanna River & Chemung River	Towanda (site #1)	20,194	4.0	35.0	60.0	1.0
SB2	Middle Susquehanna River	Danville (#2) – Towanda (#1)	8,866	7.3	28.4	60.0	4.3
SB3	West Branch Susquehanna River	Lewisburg (#3)	17,734	2.0	15.0	81.0	2.0
SB4	Juniata River	Newport (#4)	8,687	2.0	28.0	69.0	1.0
SB5	Lower Susquehanna River above Marietta	Marietta (#5) – Newport (#4) – Lewisburg (#3) – Danville (#2)	11,834	6.0	46.6	44.7	2.7
SB6	Conestoga River	Conestoga (#6)	1,217	14.0	60.0	23.0	3.0
SB7	Lower Susquehanna River below Marietta	Conowingo (#7) – Marietta (#6) – Conestoga (#5)	1,658	NA <sup>b</sup>	NA	NA	NA

<sup>a</sup> modified from Table 3 and Table 8 in Sprague *et al.* (2000)

<sup>b</sup> not available

season, and discharge (*e.g.*, Cohn *et al.*, 1989). For this work, we applied a method recently adopted by the USGS called “Weighted Regressions on Time, Discharge, and Season (WRTDS)” (Hirsch *et al.*, 2010). This approach has been shown to offer better performance than prior regression-based methods because it does not rely on those methods’ problematic assumptions about the homoscedasticity of model errors, constancy of seasonal trends in concentration, or constancy of the concentration-discharge relationship (Hirsch *et al.*, 2010; Moyer *et al.*, 2012; Chanut *et al.*, 2016).

In general, WRTDS can produce two types of concentration and loading estimates, which are called “*true-condition*” and “*flow-normalized*” estimates, respectively. True-condition estimates are model-based approximations of the real history of riverine concentration or loading and are relevant to understanding actual downstream impacts. By contrast, the flow-normalization method uses the full history of flows on the given calendar date to effectively remove the effects of inter-annual streamflow variability. It should therefore better reflect the effects of changes in source inputs and system function (Hirsch *et al.*, 2010). Because this method considers flow data from the entire record, it requires more computational effort than the true-condition estimates (Hirsch and De Cicco, 2015).

At each site, WRTDS was run using the *EGRET* (Exploration and Graphics for RivEr Trends) version 2.2.0 (Hirsch and De Cicco, 2015) to produce both the true-condition and flow-normalized concentration and loading estimates for each day in the record for each water-quality species (*i.e.*, SS, TP, TN, DP, and DN). For all WRTDS runs, we used the default settings specified by the user guide (Hirsch and De Cicco, 2015) – see details in the online Supplementary Material. The daily loading estimates were averaged to obtain

annual loading estimates for each calendar year between the start year (various among sites; see Table C1) and 2013. In addition, particulate phosphorus (PP) and particulate nitrogen (PN) loadings were inferred by subtracting DP and DN from TP and TN loadings, respectively. Annual yields for each site were calculated by dividing the annual loading estimates by their respective drainage areas. Long-term median loadings and yields for each site are provided in Table C2. Finally, annual true-condition loadings for each site were divided by their respective annual discharges to estimate annual flow-weighted concentrations ( $C_{\text{Annual-FW}}$ ).

For each WRTDS run, residual plots were generated to evaluate model performance (not shown). These plots showed that unaccounted residuals from WRTDS generally have no structural relationship with time, discharge, or season. All derived estimates from this work, along with the river monitoring data, are stored at the publicly accessible Johns Hopkins University Data Archive (Zhang and Ball, 2016). Additionally, loads for TN, TP, and SS for all seven sites can be downloaded from the USGS-designated website (U.S. Geological Survey, 2014b).

#### **4.2.3. Trend and Mass-Balance Analyses**

To address the four questions posed in Section 4.1, we conducted two major types of analysis based on WRTDS estimates: “*trend analysis*” and “*mass-balance analysis*”. Questions Q<sub>1</sub> and Q<sub>2</sub> were aimed toward better understanding when and where riverine loadings have changed. For such “*trend analysis*,” we focused on the synthesis of WRTDS *flow-normalized* estimates. To better understand these riverine trends, we compiled and analyzed watershed source input data made available to us by the Chesapeake Bay Program Office (CBPO) (Shenk and Linker, 2013). These data are used

as input to the Chesapeake Bay Watershed Phase 5.3.2 model – see <http://ches.communitymodeling.org/models/CBPhase5/index.php> for details. The most relevant of these data for our purposes are: atmospheric deposition data (estimated using developed regression models – see Grimm and Lynch (2005)) and Linker *et al.* (2013a)) for details); fertilizer and manure application data (estimated using agricultural census data); and data for point-source contributions (including significant/non-significant dischargers, industrial flows, and combined sewer overflows). For each of these four major categories (and several minor others), the CBPO has estimated monthly source inputs from each drainage basin between 1984 and 2011.

Questions Q<sub>3</sub> and Q<sub>4</sub> were aimed toward better understanding the relative contributions of loadings from the Susquehanna sub-basins and the effects of streamflow and land use on constituent export. For these questions, “*mass-balance analysis*” was conducted using WRTDS *true-condition* estimates. This type of analysis is particularly suitable to the non-tidal SRB because of the well-positioned locations of the monitoring sites (Figure 4.1). Specifically, we divided the non-tidal SRB into seven sub-basins (SBs), namely, Upper Susquehanna River plus Chemung River (SB1), Middle Susquehanna River (SB2), West Branch Susquehanna River (SB3), Juniata River (SB4), Lower Susquehanna River above Marietta (SB5), Conestoga River (SB6), and Lower Susquehanna below Marietta (SB7) (Figure 4.1, Table 4.1). Note that SB7 covers the LSRRS and its vicinity. The seven sub-basins range between 1,217 and 20,194 km<sup>2</sup> in drainage area. For each sub-basin, riverine input constituent load is the flux entering its river reach, including the flux passing the monitoring site at the upstream limit of the reach and the tributary flux entering that reach, if monitored. Output load is the flux

passing the monitoring site at the downstream limit of the reach. The output load was subtracted from riverine input load to determine whether each sub-basin was a net source (*i.e.*, riverine output > riverine input) or net storage (*i.e.*, riverine output < riverine input) (Table 4.1). This analysis assumes that WRTDS load estimation can approximately reproduce actual mass-balance relations across sites, which is an expected but not mathematically certain condition. Research is underway to better understand uncertainties and imprecisions of such estimates – irrespective of these concerns, however, the mass-balance results presented herein should shed some useful insights on relative loading contributions among sub-basins.

### **4.3. Results**

#### **4.3.1. Temporal Trends in Flow-normalized Riverine Loadings**

To compare loading trends across all seven long-term sites for particulate and dissolved constituents, we summarized flow-normalized (FN) modeled loadings for SS, TP, DP, PP, TN, DN, and PN in Figure 4.2. By integrating out the effects of inter-annual variability in streamflow, these FN loadings (Figure 4.2b-h) show much smoother trends than true-condition loadings or streamflow (Figure 4.2a). FN-modeled loadings of SS show overall downward trends at all sites except Conowingo (Figure 4.2b). Among these sites, Conestoga had the highest early-period SS yield but also showed the strongest decline in FN loading. By contrast, Conowingo shows a clear rise since around 2000.

FN-modeled trends of TP, DP, and PP are shown in Figure 4.2c-e, respectively. At all sites except Conowingo, TP has shown general overall declines over time, but with some short periods of stable or slightly rising loads (Figure 4.2c). These TP declines are

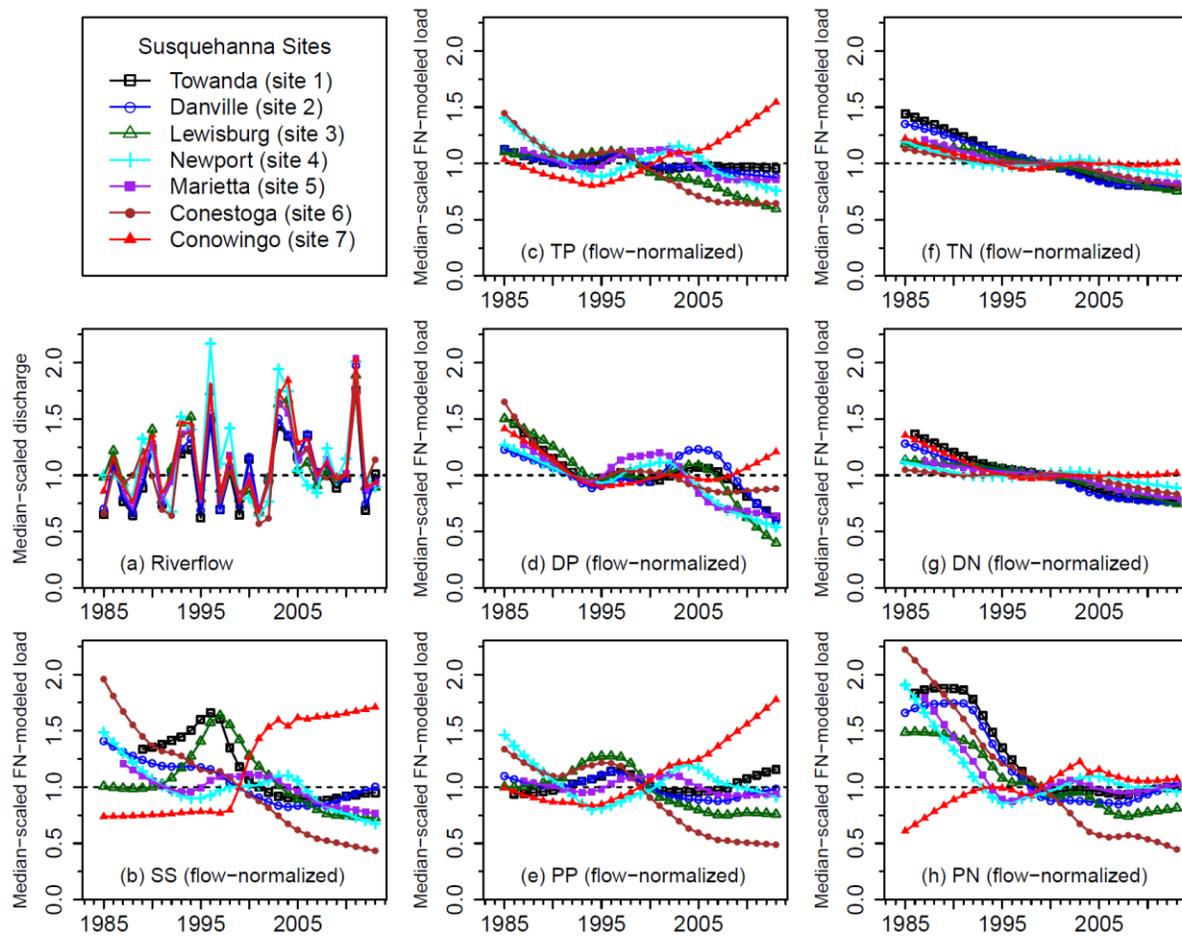


Figure 4.2. Reconstructed time series of (a) annual discharge and annual *flow-normalized* (FN) loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN at the seven Susquehanna sites. To aid comparison, all *y-axis* values have been scaled by respective long-term annual medians (see Table C2).

accompanied by declining FN-modeled PP loading since the late 1990s (Figure 4.2e) and by declining FN-modeled DP loading in both earlier (1986-1993) and more recent (2005-2013) periods (Figure 4.2d). In comparison, FN-modeled PP trends (Figure 4.2e) have closely followed those of SS (Figure 4.2b), with high correlations at each site (linear correlation coefficients: 0.47-0.92; median: 0.83), which reflects the critical role of sediment in PP transport. As in the case of SS, Conestoga had the highest early-period yields of TP and PP (see Table C4) and showed the strongest declines in FN-modeled TP and PP loadings. By contrast, FN loadings of TP and PP at Conowingo exhibit clear rises in recent decades.

FN-modeled trends of TN, DN, and PN are shown in Figure 4.2f-h, respectively. TN loadings show steady declines at all sites except Conowingo (Figure 4.2f) and these declines have been primarily driven by declines in DN loadings (Figure 4.2g). Among these sites, Towanda results show the strongest fractional decline in FN loadings of TN and DN, whereas the Conestoga estimates show the strongest absolute decline in TN yield (Table C4). For PN, FN estimates of loading show declines at all sites except Conowingo, with Conestoga showing the strongest decline (Figure 4.2h). In contrast with the upstream sites, Conowingo results show steady (but slight) rises in DN and TN loadings in recent years and a much stronger rise in PN loading throughout the study period.

#### **4.3.2. Comparison of Changes in Riverine Yield and Source Input**

To further evaluate riverine loading trends in the context of management, it is useful to consider the contemporary histories of watershed source input. Toward this end, input data for major source categories between 1984 and 2011 are plotted together with our

riverine estimates in Figure 4.3 for TP and Figure 4.4 for TN. Period-of-record averages of annual riverine conditions (*i.e.*, river flow, concentration, and yield) and annual source inputs for the period of 1987-2011 are summarized in Table C3, which is the longest period that has data at all sites. In addition, the period-of-record change in source input yield ( $\Delta_{\text{Input}}$ ) for each site was quantified, *i.e.*,  $\Delta_{\text{Input}} = 2011 \text{ yield} - 1987 \text{ yield}$ . Similarly, the period-of-record change ( $\Delta$ ) was quantified for each individual source input. These changes, along with the initial (1987) and final (2011) conditions, are summarized in Table C4 and discussed below.

For TP, total source inputs have declined generally at all sites except Newport (Figure 4.3, Table C4). Among the other six sites, the decline for Conestoga ( $\Delta_{\text{Input}} = -304 \text{ kg km}^{-2}$ ) is much greater than those for the other sites (-29 to  $-122 \text{ kg km}^{-2}$ ; Table C4).

Individual sources showed negative  $\Delta$  for 17 of 21 source-site combinations (Table C4), with only the following exceptions: fertilizer at Conestoga ( $+23 \text{ kg km}^{-2}$ ) and manure at Newport ( $+83 \text{ kg km}^{-2}$ ), Marietta ( $+10 \text{ kg km}^{-2}$ ), and Conowingo ( $+3 \text{ kg km}^{-2}$ ). Of the various source categories, declines in estimated manure input contributed the most to overall declines in estimated total input at Conestoga, Danville, and Towanda, whereas declines in fertilizer contributed the most to the overall declines in total input at Conowingo, Marietta, and Lewisburg. The only positive  $\Delta_{\text{Input}}$  (total input) occurs at Newport ( $+51 \text{ kg km}^{-2}$ ), which is entirely attributable to manure rise.

For TN, total source inputs have declined consistently at all sites (Figure 4.4, Table C4). Danville has the largest decline ( $-1,146 \text{ kg km}^{-2}$ ) and Newport has the smallest ( $-469 \text{ kg km}^{-2}$ ) (Table C4). For individual sources, 25 of 28 source-site combinations have negative  $\Delta$  (Table C4), with positive  $\Delta$  only for manure at Newport ( $+266 \text{ kg km}^{-2}$ ) and

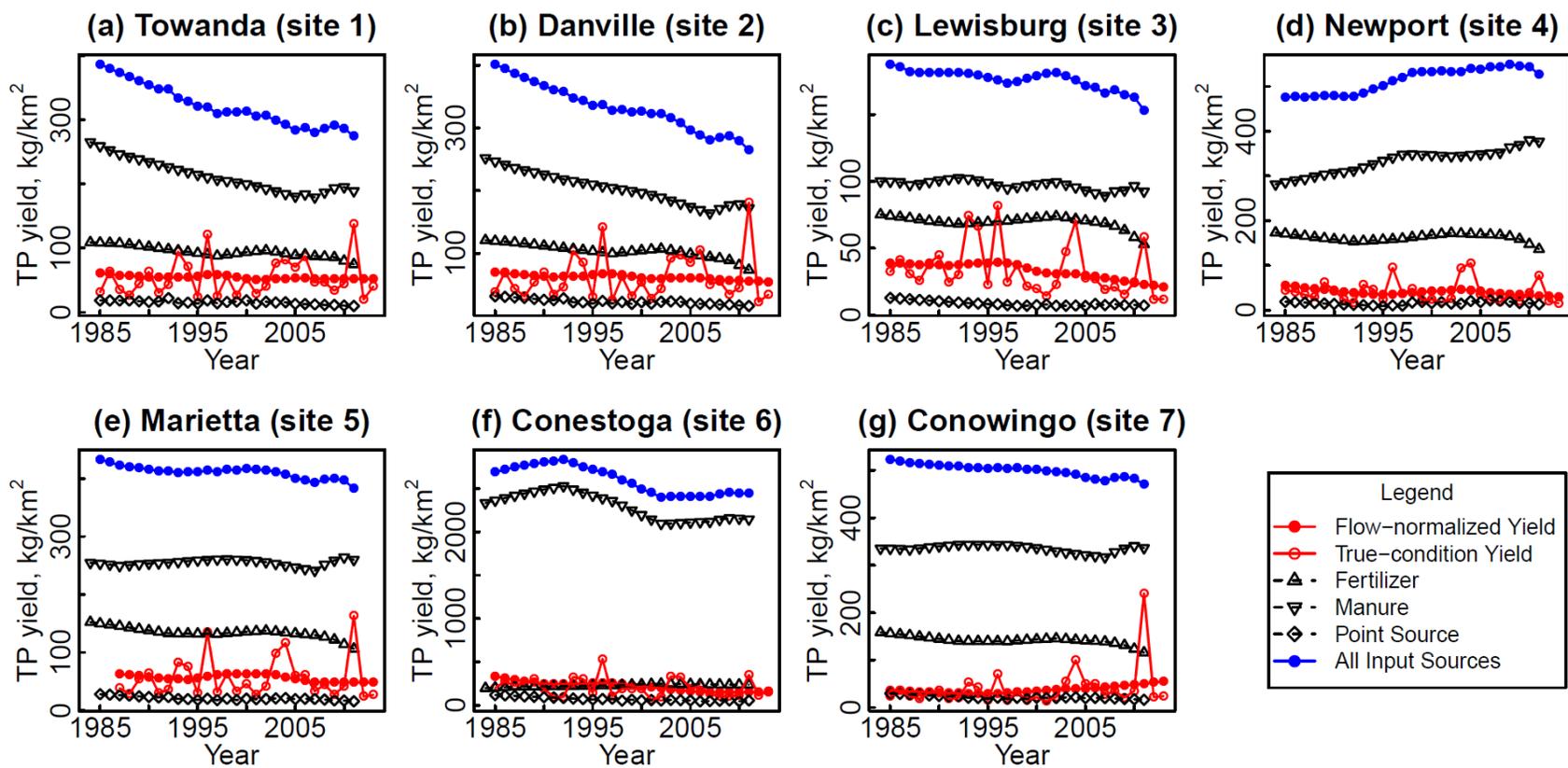


Figure 4.3. Reconstructed time series of WRTDS-estimated TP riverine yield (flow-normalized and true-condition estimates) and yields from major source inputs (fertilizer, manure, point source, and sum of all sources) for the seven Susquehanna sites: (a) Towanda, (b) Danville, (c) Lewisburg, (d) Newport, (e) Marietta, (f) Conestoga, and (g) Conowingo. Note that the *y-axis* scale varies with plot.

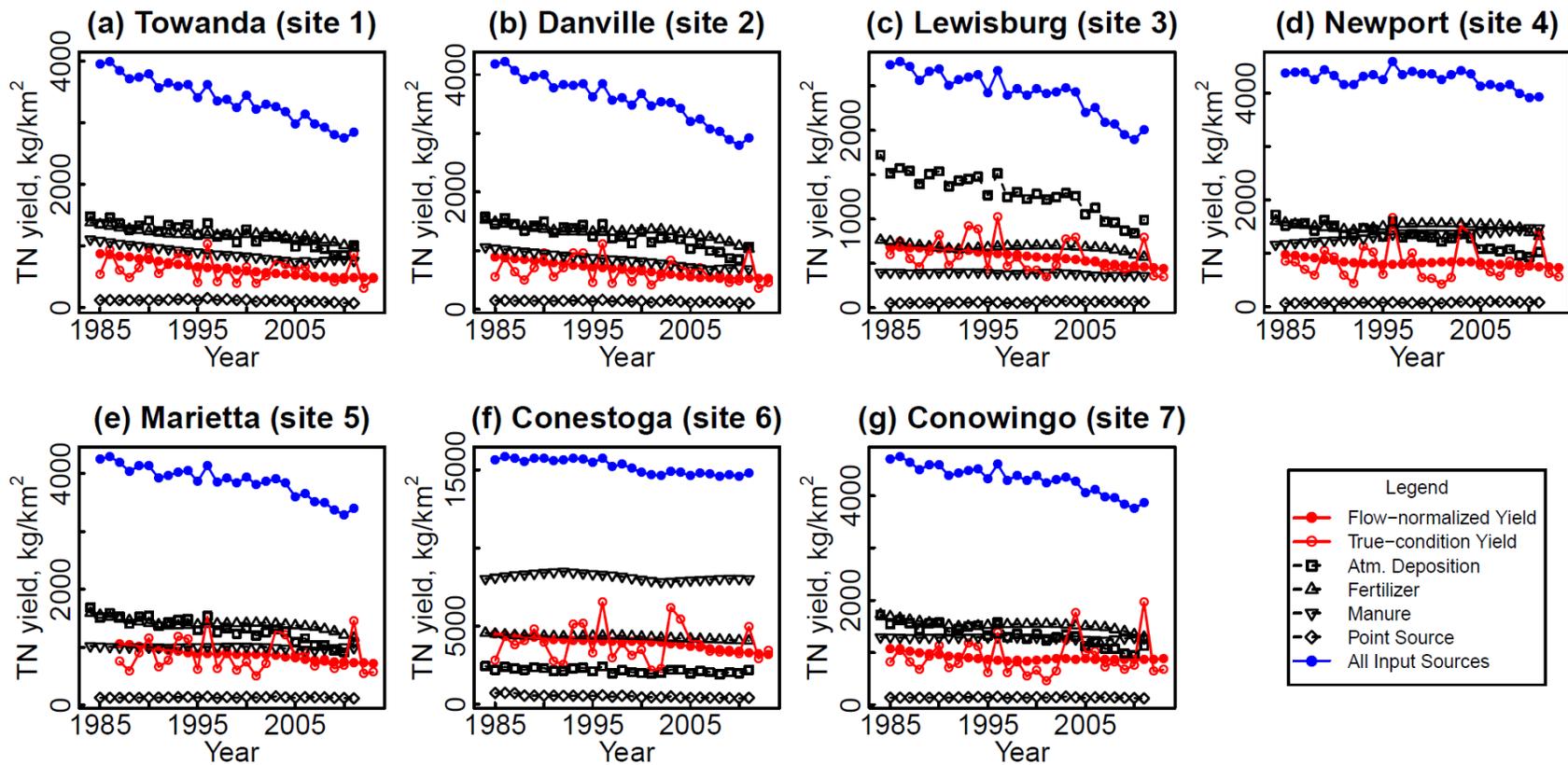


Figure 4.4. Reconstructed time series of WRTDS-estimated TN riverine yield (flow-normalized and true-condition estimates) and yields from major source inputs (atmospheric deposition, fertilizer, manure, point source, and sum of all sources) for the seven Susquehanna sites: (a) Towanda, (b) Danville, (c) Lewisburg, (d) Newport, (e) Marietta, (f) Conestoga, and (g) Conowingo. Note that the *y-axis* scale varies with plot.

for point source at Newport (+13 kg km<sup>-2</sup>) and Lewisburg (+11 kg km<sup>-2</sup>). Atmospheric deposition showed consistent declines at all sites and with similar  $\Delta$  (-356 to -552 kg km<sup>-2</sup>) for all sites except Conestoga, where  $\Delta$  was much smaller (-87 kg km<sup>-2</sup>). Fertilizer showed consistently negative  $\Delta$  at all sites, while manure and point source showed negative  $\Delta$  at most sites, with the few exceptions noted above. Of these, the positive  $\Delta$  in manure at Newport (+266 kg km<sup>-2</sup>) is the most substantial and has counteracted a substantial portion of the negative  $\Delta$  in atmospheric deposition (-552 kg km<sup>-2</sup>). Among the individual sources, declines in estimated atmospheric deposition have had the greatest contribution to overall declines in estimated total input at all sites except Conestoga (Table C4).

For purposes of comparison to the  $\Delta_{\text{Input}}$  values, we calculated the period-of-record (1987-2011) changes in flow-normalized riverine yield ( $\Delta_{\text{FN-Yield}}$ ). The  $\Delta_{\text{FN-Yield}}$  values for TP and TN are negative for all cases except TP at Conowingo (Table C4). For both TP and TN, Conestoga had the strongest decline among all sites (-146 kg P km<sup>-2</sup> and -1,078 kg N km<sup>-2</sup>). The  $\Delta_{\text{FN-Yield}}$  values are shown against the  $\Delta_{\text{Input}}$  values in Figure 4.5, from which it is evident that  $\Delta_{\text{FN-Yield}}$  values are consistently lower. Moreover, the non-Conestoga sites show no relation for TP and a low-slope relation for TN.<sup>12</sup> (Note that the Conestoga site is clearly from a different population than the other sites.) For quantitative comparison, their ratios, *i.e.*,  $\Delta_{\text{FN-Yield}}/\Delta_{\text{Input}}$ , were calculated as a simple description of the fraction of source reduction that has been realized in the riverine yield decline. For TP, these ratios are 0.52, 0.48, 0.37, 0.09, and 0.06 for Lewisburg, Conestoga, Marietta,

---

<sup>12</sup> Relative to our published paper (Zhang *et al.*, 2016), the underlined text is a clarification of the intended interpretations.

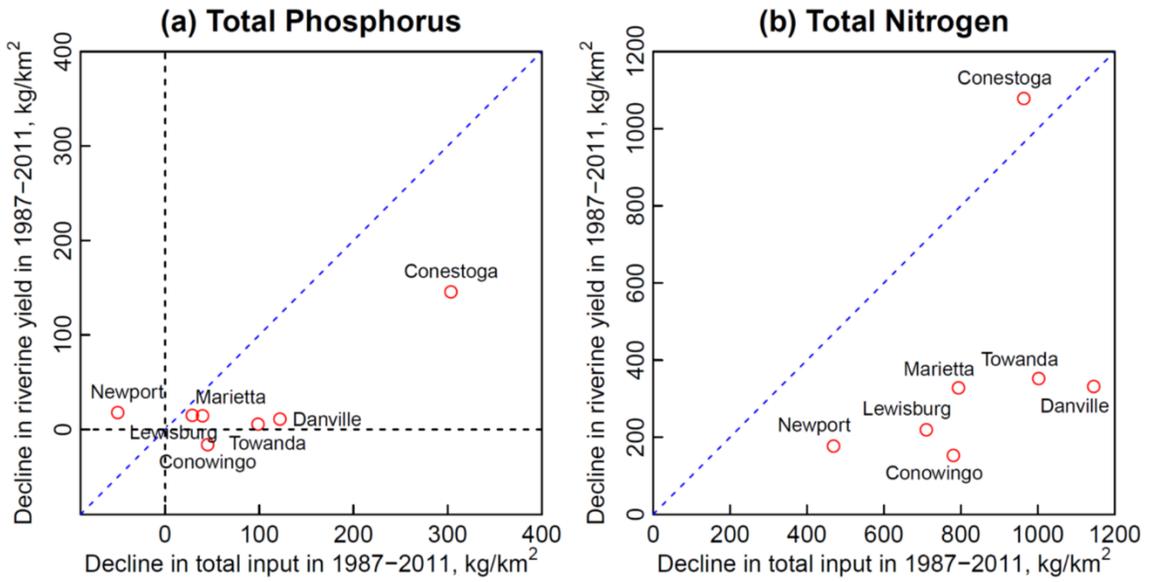


Figure 4.5. Comparison between declines in total source input and declines in flow-normalized riverine yield for the seven Susquehanna sites for the period of 1987-2011. See Table C3 for details. The blue diagonal line indicates the 1:1 reference line. (Note that the Conestoga site is clearly from a different population than the other sites. These non-Conestoga sites show no relation for TP and a low-slope relation for TN.)

Danville, and Towanda, respectively. Notably, the ratio is negative for both Conowingo (-0.35) and Newport (-0.35). For TN, the ratios are 1.12, 0.41, 0.38, 0.35, 0.31, 0.29, and 0.20 for Conestoga, Marietta, Newport, Towanda, Lewisburg, Danville, and Conowingo, respectively. In general, these ratios are  $< 1.0$  (13 of 14 cases).

#### 4.3.3. Mass Balances of Sub-basins and Effects of Streamflow on Export

To examine whether all Susquehanna sub-basins have been net loading sources (*i.e.*, riverine output  $>$  riverine input), we calculated true-condition annual net loading for each constituent between 1985 and 2013 for each sub-basin (Figure 4.6). The results show that all six upstream sub-basins (*i.e.*, SB1-SB6) have been net exporters for nutrients and sediment in almost every year throughout the last three decades, including dry years. In

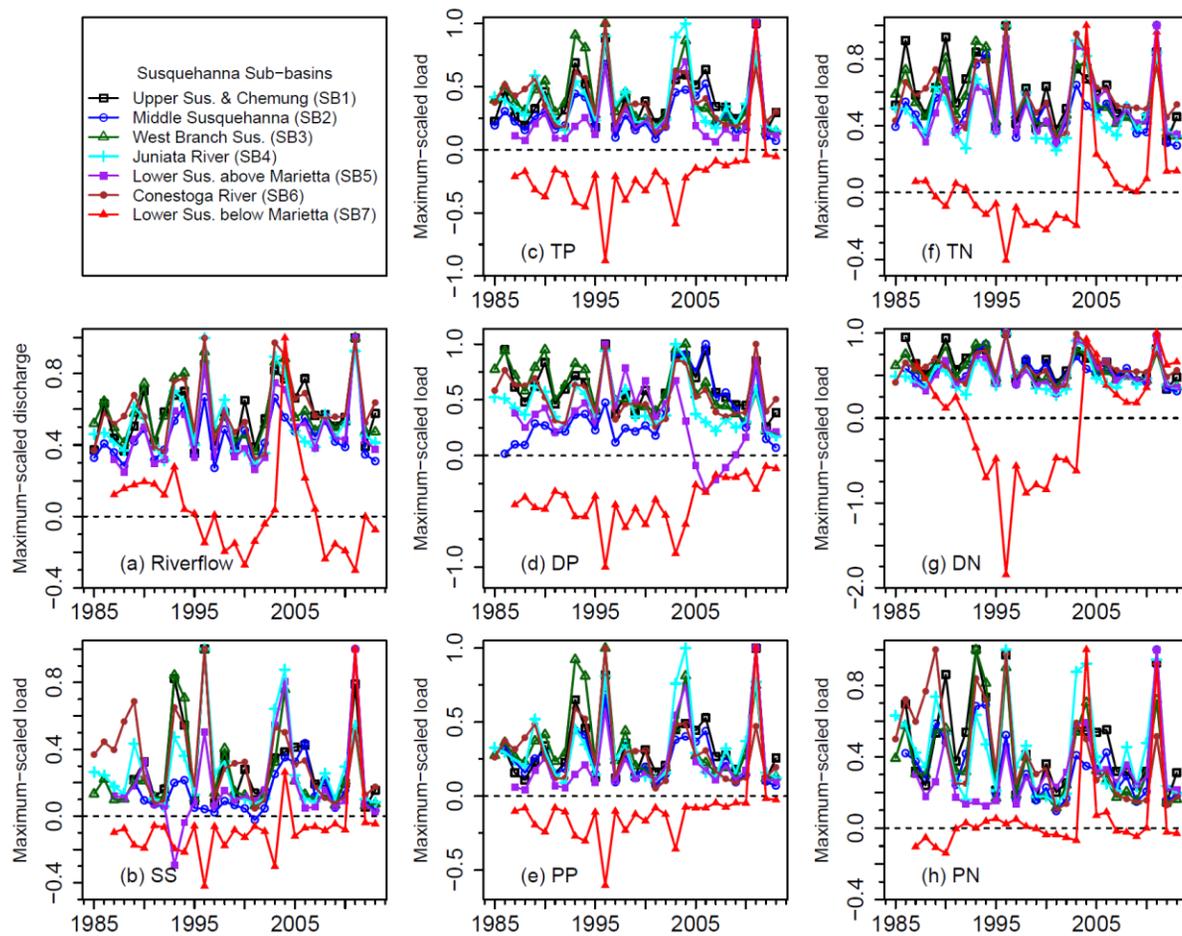


Figure 4.6. Reconstructed time series of (a) annual discharge and annual *true-condition* net contributions of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN by the seven Susquehanna sub-basins. To aid comparison, *all y-axis* values have been scaled by respective long-term annual maxima.

contrast, the most downstream sub-basin, SB7 (Lower Susquehanna River below Marietta), shows negative values that reflect net accumulations of N, P, and SS, as expected for this reservoir-dominated sub-basin (Figure 4.6). Rises in the SB7 plots, however, are observed in recent decades, especially for particulate species, suggesting that SB7 may soon become a net neutral or net positive source of nutrients and sediment to the downstream reach.

Peaks for streamflow and all constituents have occurred concurrently in SB1-SB6, with all plots showing a striking similarity in the timing of significant export (Figure 4.6). To further examine the dominance of hydrological control on constituent export, we analyzed the relationships between annual loading ( $L_{\text{Annual}}$ ) and annual discharge ( $Q_{\text{Annual}}$ ). Despite considerable scatter with some site-species combinations, strong linear  $\log[L_{\text{Annual}}] \sim \log[Q_{\text{Annual}}]$  relationships (p-value < 0.01) are observed for all species at all sites (Figure C1). Thus  $Q_{\text{Annual}}$  alone is a strong predictor of  $L_{\text{Annual}}$ . Within this context, and given the definition of load ( $L = QC$  or  $\log[L] = \log[Q] + \log[C]$ ), a slope of 1.0 would be expected for conditions of constant concentration, and deviations from this value are indicative of the nature of  $C$ - $Q$  relationships. An alternative approach to investigate such effects is to directly examine concentration data, as is done in Figure C2, where we have plotted annual flow-weighted concentration ( $C_{\text{Annual-FW}}$ , calculated as  $L_{\text{Annual}}/Q_{\text{Annual}}$ ) against area-normalized annual discharge,  $(Q/A)_{\text{Annual}}$ . Note that this annual averaging helps mitigate some of the issues associated with the fact that  $C$ - $Q$  relationships can vary with time and season and also depend on time of sampling within a hydrograph (*e.g.*, during rising and falling limbs), which can be an especially important problem for high-discharge events that are only sparsely sampled. Nonetheless, we have

identified years with extreme-discharge events in Figure C2 as a means of qualitatively looking for outliers. Evidently, these specific years (1996, 2004, 2011) fall within the general trend. Overall, the approximately linear slope of  $\log[C_{\text{Annual-FW}}] \sim \log[(Q/A)_{\text{Annual}}]$  can coarsely reveal whether export patterns follow dilution (slope  $< 0$ ), chemostasis (slope  $\sim 0$ ), or mobilization (slope  $> 0$ ) (Godsey *et al.*, 2009; Stallard and Murphy, 2014). For our sites, the results show general chemostasis effects for dissolved and dissolved-dominated species (*i.e.*, DN, TN, and DP) but mobilization effects for particulate and particulate-dominated species (*i.e.*, PN, TP, PP, and SS) (Figure 4.7; C2).

Considering the above distinction between particulate and dissolved constituents, streamflow may have played an important role in modulating the relative importance of dissolved and particulate fractions. To verify such effects, we have plotted ratios of annual DP to annual TP loads (DP/TP), annual DN to annual TN loads (DN/TN), and

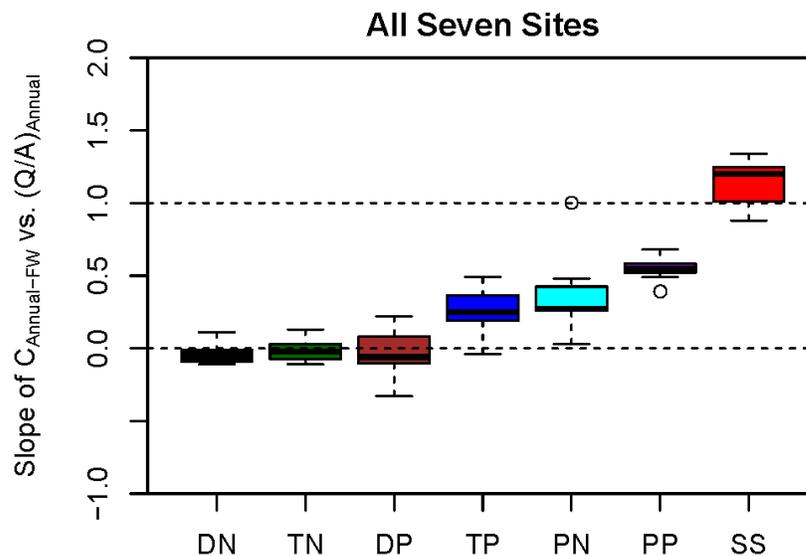


Figure 4.7. Fitted linear slopes for relations between annual flow-weighted concentration ( $C_{\text{Annual-FW}}$ ) and area-normalized annual discharge ( $[Q/A]_{\text{Annual}}$ ) on log-log scale at the seven Susquehanna sites for each water-quality constituent. Detailed data and slope fits are presented in Figure C2.

annual PP to annual SS loads (PP/SS) against  $(Q/A)_{\text{Annual}}$  for each site and observed general negative correlations (Figure 4.8a-c). Moreover, as a means of considering nutrient loadings in the context of algal growth, we plotted TN/TP molar ratio against area-normalized discharge on both annual (Figure 4.8d) and daily (Figure 4.8e) scales. (The daily ratios are considered more representative of instantaneous ratios than the annual ratios.) Following the convention of Qian *et al.* (2000)), we classified the TN/TP molar ratios into three nominal categories with respect to possible nutrient limitation: (1) P-limitation ( $\text{TN/TP} > 20$ ), (2) co-limitation by both N and P ( $10 \leq \text{TN/TP} \leq 20$ ), and (3) N-limitation ( $\text{TN/TP} < 10$ ). We emphasize that these categories are nominal only – although based on the classic Redfield ratio (Redfield, 1958), the cut-off values do not reflect any actual knowledge about limitations in the given systems. Rather, our interest is in the comparative ratios and trends. In these regards, both plots show generally lower TN/TP ratios during high discharge at the Susquehanna sites, including differences in regard to the nominally limiting nutrient. Notably, the daily ratio follows a clear spatial gradient, with values increasing from upstream to downstream sites and with Conowingo consistently having the highest ratio (Figure 4.8f).

#### **4.3.4. Relative Contributions by Sub-basins and Effects of Land Use on Export**

To compare the relative contributions of sub-basins to the total annual loadings at Conowingo (*i.e.*, delivered loading from the non-tidal SRB to Chesapeake Bay) between 1985 and 2013, we quantified the fractional contributions (FCs) of each sub-basin for each species – see Figure 4.9. (Note that FCs of these seven sub-basins always sum up to one in each year and that the negative FCs of SB7 [Lower Susquehanna River below Marietta] correspond to net storage in the reservoir system.) For streamflow, SB1 (Upper

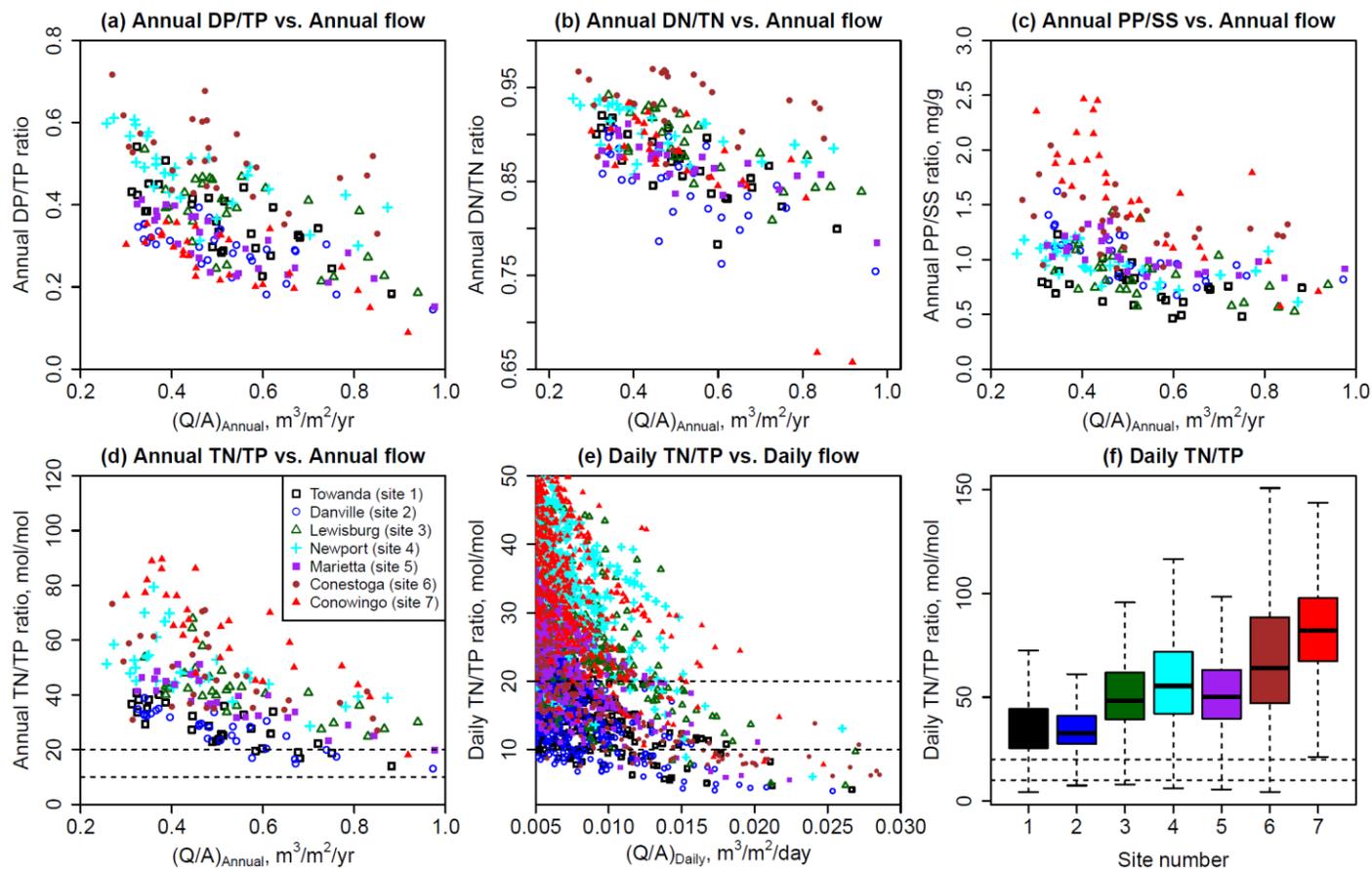


Figure 4.8. Relations between area-normalized annual discharge ( $[Q/A]_{\text{Annual}}$ ) and annual ratios of (a) DP/TP, (b) DN/TN, (c) PP/SS, and (d) TN/TP at the seven Susquehanna sites. Plot (e) shows relations between area-normalized daily discharge ( $[Q/A]_{\text{Daily}}$ ) and daily TN/TP ratio for each site. Plot (f) summarizes the daily TN/TP ratio at each site with boxplots. The region between dashed lines in plots (d)-(f) represent the nominally co-limitation condition by both N and P (*i.e.*,  $10 \leq \text{TN/TP} \leq 20$ ).

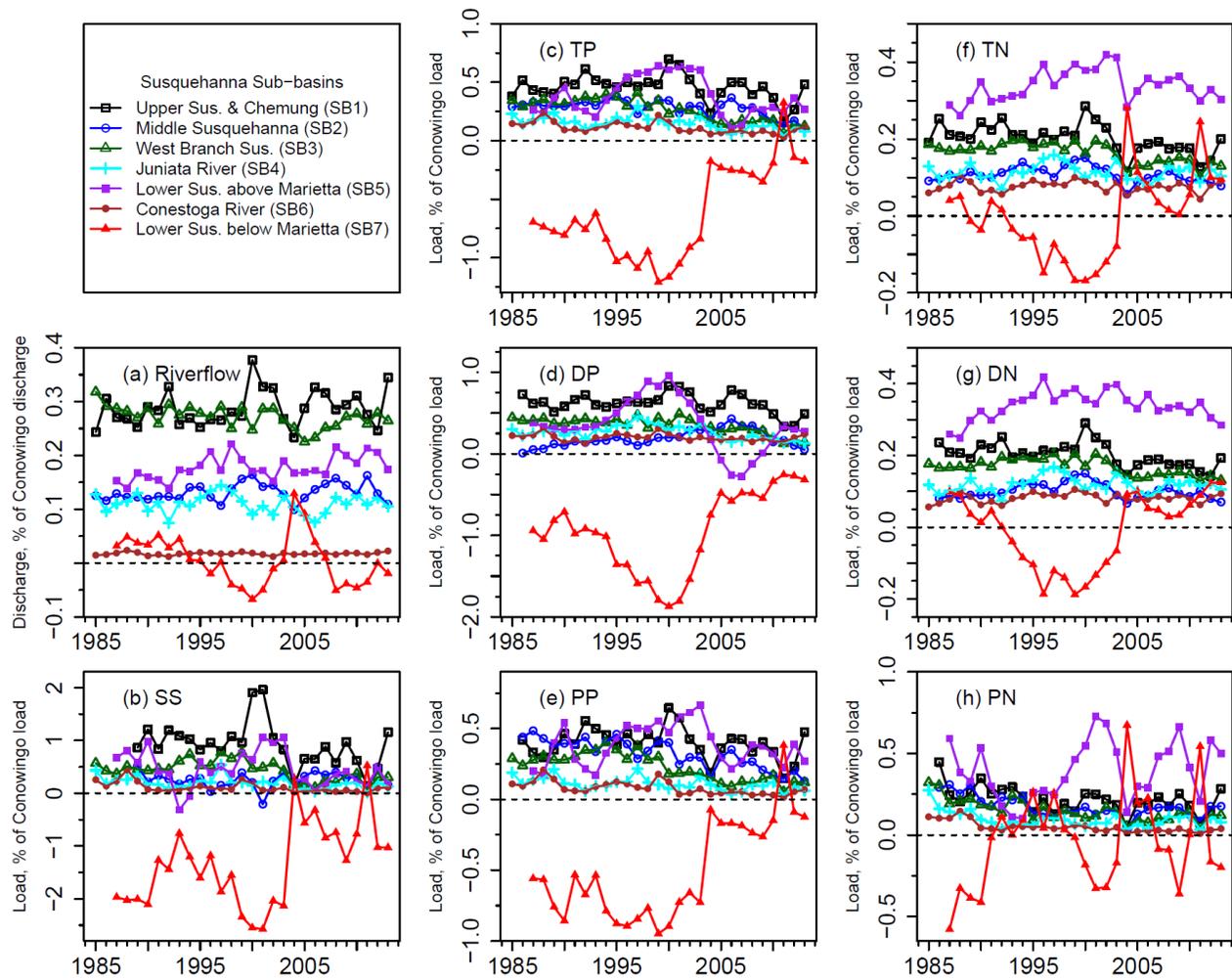


Figure 4.9. Fractional contributions (FC) of each sub-basin to (a) annual discharge and annual *true-condition* loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN at Conowingo (river fall-line). Note that FCs of all sub-basins sum up to one in each year.

Susquehanna River & Chemung River) and SB3 (West Branch Susquehanna River) have the highest FCs throughout the last three decades as they have the largest drainage areas (Table 4.1). Consistent with the large area, SB1 also has the highest FCs for SS, TP, DP, and PP. SB3, however, has generally low FCs for all constituents. Another major deviation between rankings of streamflow FC and constituent FC is observed with SB5 (Lower Susquehanna River above Marietta). This sub-basin is about 60% of SB1 in drainage area (Table 4.1), but it has the highest FCs for all N species (*i.e.*, TN, DN, and PN) and the second highest FCs for all P species (*i.e.*, TP, DP, and PP) that are only slightly lower than those of SB1. In comparison to SB5, SB7 (Lower Susquehanna River below Marietta) is also located in the Lower Susquehanna area and dominated by agricultural land; however, this reservoir-dominated sub-basin does not export constituents in a similar way. Instead, various species are at least partially retained. To further quantify the relationships between constituent export and land use, we plotted the period-of-record medians of  $(Q/A)_{\text{Annual}}$  and annual constituent yield against area fractions of major land uses, namely, non-forested (*i.e.*, agricultural, urban, and others) and forested (Figure 4.10). (SB7 was excluded from this analysis to remove the complication of reservoir effects.) Simple linear regressions were developed between log-transformed median annual yield and area fractions of land uses. Due to data limitation (number of sub-basins = 6), our linear models involve only one explanatory variable and thus cannot account for interactions between “forested” and “non-forested” lands or for any additional variability that is associated with different categories of “non-forested” lands. Nonetheless, this simple approach can provide some qualitative insights on land-use effects. In general, period-of-record median of  $(Q/A)_{\text{Annual}}$  is almost invariant between

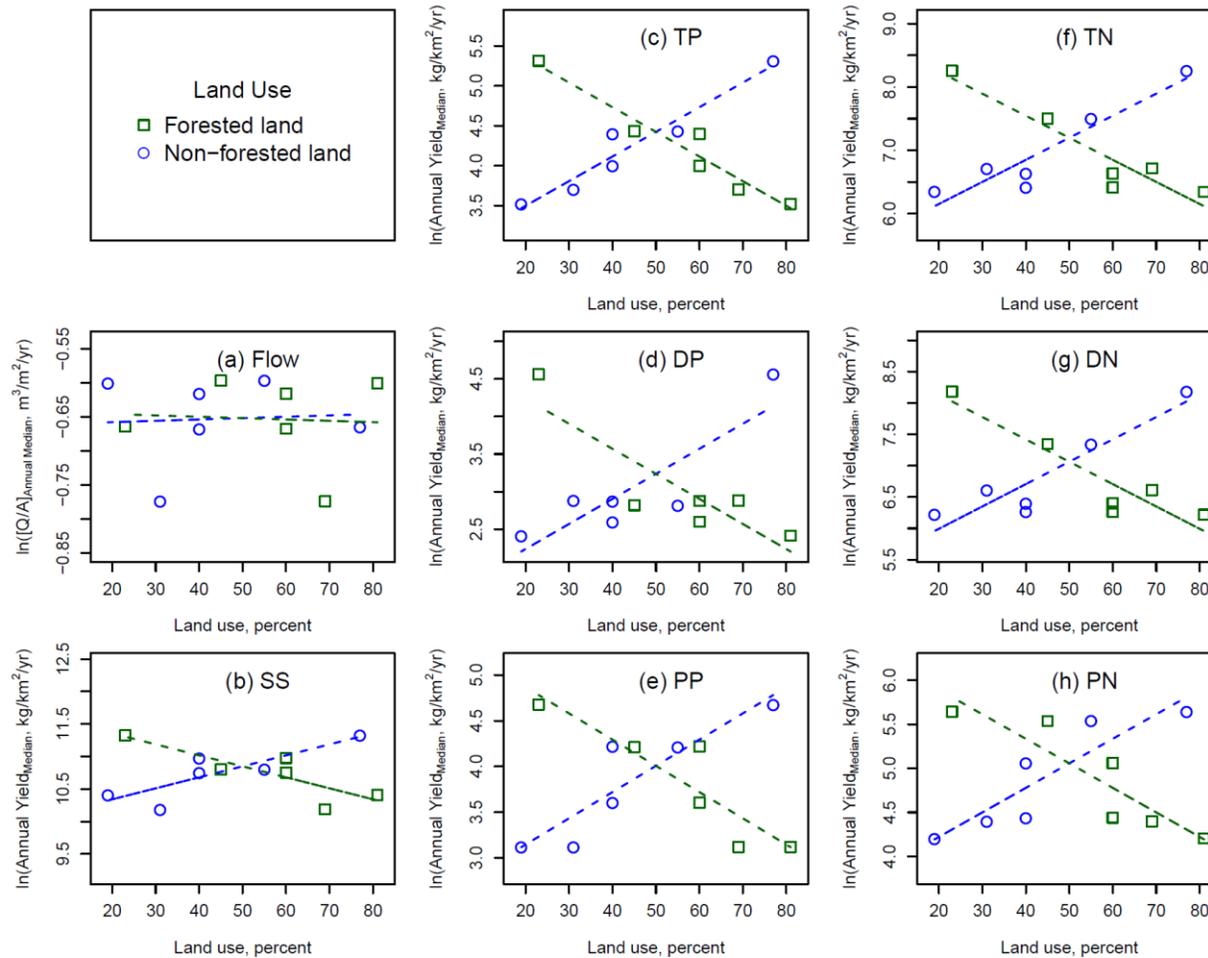


Figure 4.10. Relations between area fractions of two types of land use (*i.e.*, forested and non-forested) and log-transformed period-of-record medians of (a) area-normalized annual discharge ( $[Q/A]_{\text{Annual Median}}$ ) and annual *true-condition* loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN in the seven sub-basins. Each point represents one sub-basin. Dashed lines are linear fits.

land use types. In contrast, median annual yields of N, P, and SS are strongly affected by land uses: the median annual yields correlate positively with the area fraction of non-forested land but negatively with that of forested land.

To evaluate land-use effects under different hydrological conditions, we categorized each year in the period of 1987-2013 to three flow classes, *i.e.*, wet, average, and dry years. These were determined according to the ranking of annual-average streamflow: the highest 30%, the lowest 30%, and the middle 40% are classified as “wet,” “dry,” and “average,” respectively. Correlation analyses presented above (Figure 4.10) were then separately conducted on subsets of data corresponding to each flow class. Results for non-forested land show positive effects on log-transformed median yields for all species under all three flow classes, with generally similar slopes but different intercepts – wet years always have much larger intercepts than average or dry years (Figure C3). For forested land, the slopes between log-transformed median yields and area fraction are consistently negative under all three flow classes, and intercepts are again larger in wet years than average or dry years (Figure C4).

## **4.4. Discussion**

### **4.4.1. Temporal Trends in Flow-normalized Riverine Loadings**

FN-modeled loadings show general declines for all species (*i.e.*, dissolved, particulate, and total) at all Susquehanna sites upstream of Conowingo Reservoir in the last three decades (Section 4.3.1; Figure 4.2). The general consistency in timing and magnitude across sites indicates that changes have been relatively uniform spatially, presumably reflective of basin-wide effects from management controls. Although it is

difficult to establish causation (which is beyond the scope of this work), we list below some major management actions that have possibly affected the observed trends. For SS (Figure 4.2b), the general declines may reflect improvements in land management practices with respect to control of sediment sources and transport. For P, the declines in DP (Figure 4.2d) at least partially benefited from the implementation of a P-detergent ban in Pennsylvania since 1990 (Litke, 1999) and nutrient removal technology upgrade at WWTPs since the 1980s (Chesapeake Executive Council, 1988; Chesapeake Bay Program, 1998). For example, point source input at Marietta is calculated to have declined by 172,000 kg between 1995 and 2011, which is 51% of the estimated decline in DP riverine loading over the same period (336,000 kg). In comparison, the PP declines (Figure 4.2e) are more reflective of nonpoint source controls, including at least fertilizer and manure reductions and P-based nutrient management (Weld *et al.*, 2002). For TN (and DN), the declines (Figure 4.2f-g) are likely related to historical controls on point sources (particularly WWTP upgrade) and nonpoint sources (*e.g.*, fertilizer and manure applications) (Chesapeake Executive Council, 1988; Chesapeake Bay Program, 1998) and measures taken in response to the Clean Air Act and associated reductions in nitrogen oxide emissions from coal-fired power plants and automobiles (Eshleman *et al.*, 2013; Linker *et al.*, 2013b).

In contrast with the upstream sites, Conowingo showed clear rises in FN loadings of SS, TP, PP, and PN in recent years (Figure 4.2), re-affirming the trends documented previously (Hirsch, 2012; Zhang *et al.*, 2013). Complementary to these FN-modeled trends, true-condition estimates for SB7 (Lower Susquehanna River below Marietta) show decreased net annual storage in recent decades, especially for particulate

constituents (Section 4.3.3; Figure 4.6). These results collectively suggest declining trapping performance by the LSRRS (mainly Conowingo Reservoir) and possibly associated effects on biogeochemical transformations (*e.g.*, mineralization, biotic uptake, burial in sediments, denitrification) during the (presumably declining) residence time in the reservoir. As sediment accumulates in the reservoir, cross-sectional area becomes less available for flow, thereby increasing the average horizontal flow velocity, decreasing the vertical depth from water surface to sediment bed, and increasing the relative importance of wind-induced turbulence. In this regard, our parallel work focusing on the reservoir reach has used several different approaches to demonstrate that decreased reservoir trapping has occurred under a wide range of flow conditions – see Zhang *et al.* (2016). To further understand these processes and the associated effects on reservoir modulation of upstream inputs, which is of growing concern to watershed managers (Friedrichs *et al.*, 2014; The Lower Susquehanna River Watershed Assessment Team, 2014), continued monitoring and research is indispensable. Toward this end, one major research project is already underway (University of Maryland Center for Environmental Science, 2016).

#### **4.4.2. Comparison of Changes in Riverine Yield and Source Input**

Our evaluation of the source input changes (Section 4.3.2) provides additional evidence that is useful for explaining the riverine trends. For both TP (Figure 4.3) and TN (Figure 4.4), total source input and the major individual sources have declined in the drainage basins of most sites. For individual sources, the largest declines are generally associated with manure or fertilizer for TP and atmospheric deposition for TN. Two notable anomalies are observed, however. One anomaly is manure input at Marietta, which has risen for both TN and TP. These rises may be explained by an increase in

estimated animal numbers (by  $\sim 1.6\% \text{ yr}^{-1}$ ) in the Juniata River basin, as estimated from data provided by the CBPO (Yactayo, 2015). The other anomaly is atmospheric deposition at Conestoga, for which the decline was much smaller than the other sites. This spatial difference can be related to increases in estimated ammonia deposition associated with more intense agricultural activities in the Conestoga basin (Shenk, 2015).

The  $\Delta_{\text{FN-Yield}}/\Delta_{\text{Input}}$  ratios provide a simple quantitative measure of the fraction of source reduction that has been realized in the riverine yield decline (Figure 4.5). The generally positive ratios (12 of 14 cases) indicate that riverine yield has indeed declined in response to source reductions in different parts of the SRB. Several anomalies are noted. For TP, both Conowingo and Newport show negative ratios. For Conowingo, the negative ratio reflects a positive change in TP yield despite a negative change in source input, owing to declines in TP retention within Conowingo Reservoir (see Section 4.4.1). For Newport, the negative ratio reflects a negative change in TP yield despite a positive change in source input that is more difficult to explain – at this site, the result may imply enhanced nutrient processing in the Juniata River basin. For TN, Conowingo has the smallest ratio, reflecting decreased retention of the PN fraction within the reservoir. By contrast, Conestoga has the highest ratio (1.12) among all sites, reflecting an overall decline in riverine yield even greater than that in source input. Note that, however, Conestoga is a small basin and thus the quality of source input data may not be as high as for the larger basins. Considering all sites except Conowingo and Conestoga,  $\Delta_{\text{FN-Yield}}/\Delta_{\text{Input}}$  ratio has a coefficient of variation of 2.4 for TP but only 0.14 for TN, which implies that, in the absence of information from other types of data or process-based modeling, TN is much more predictable from source input than is TP.

A consistent pattern is that  $\Delta_{\text{FN-Yield}}/\Delta_{\text{Input}}$  ratios are generally less than 1.0 (13 of 14 cases), reflecting the fact that riverine outputs at the Susquehanna sites have remained relatively constant despite strong changes in source inputs. We note that similar patterns have been documented for other watersheds within the Chesapeake Bay, Mississippi River, and Lake Erie basins, and that prior authors have speculated that such results could reflect continuing contributions from legacy sources (Meals *et al.*, 2009; Basu *et al.*, 2010; Jarvie *et al.*, 2013; Sharpley *et al.*, 2013). We speculate that our results may similarly reflect such sources. For the Chesapeake region, the legacy stores primarily comprise groundwater for N (Bachman *et al.*, 1998; Sanford and Pope, 2013), surface soils and river sediments for P (Ator *et al.*, 2011; Sharpley *et al.*, 2013), and stream corridors and reservoir beds for sediment (Gellis *et al.*, 2008; Walter and Merritts, 2008). These legacy stores originated primarily from agricultural fertilizer applications (Brush, 2009; Meals *et al.*, 2009; Sharpley *et al.*, 2013) and historical land clearances (Gellis *et al.*, 2008; Langland, 2015) and can be released during high flow and erosional events.

Definitive confirmation of such “legacy effects” for our sites may not be possible without complete understanding of the sources and sinks (*e.g.*, plant uptake, denitrification) and of the processes controlling constituent accumulation and release. (Included in such concerns, for example, would be the uncertainties and accuracies of the currently available data for source input and possible over-statement of the assumed efficiency of implemented best management practices.) Nonetheless, some authors have come to other conclusions about the causes of similar trends at other locations. In particular, Basu *et al.* (2010) observed generally low inter-annual variability in reconstructed time series of  $C_{\text{Annual-FW}}$  at their study sites (within the Mississippi-

Atchafalaya River and Baltic Sea basins) and attributed this observation to legacy sources created by long histories of anthropogenic inputs that greatly exceeded removal mechanisms. Although also not confirmed directly through mass-balance calculations, others (*e.g.*, Gellis *et al.*, 2008; Jarvie *et al.*, 2013; Sharpley *et al.*, 2013) have posited similar hypotheses, and other studies have provided supporting data. For example, legacy sources have been reported to contribute as much as over half of riverine TP and TN fluxes in some Chinese agricultural watersheds (Chen *et al.*, 2014; Chen *et al.*, 2015). Additionally, a major fraction (median = 48%) of riverine TN load at 36 Chesapeake sites was estimated to come from baseflow contributions (Bachman *et al.*, 1998). For our seven sites, results show temporal invariance in  $C_{\text{Annual-FW}}$  at all sites for dissolved species (DN, TN, and DP) and at all sites except Conestoga and Lewisburg for particulate species (PN, TP, PP, and SS) (Figure C5). These patterns are consistent with a similar conclusion of so-called “biogeochemical stationarity”<sup>13</sup> (Basu *et al.*, 2010; Thompson *et al.*, 2011) and is similarly speculated to reflect the effects of legacy sources. The management implication is that short-term water-quality improvement in Susquehanna River should not be expected to follow a “one-to-one” correspondence with reduction of contemporary source inputs (*e.g.*, fertilizer and manure) and that larger long-term gains may follow only after the depletion of legacy sources. (In this regard, the relatively strong decline in  $C_{\text{Annual-FW}}$  for particulate species at Conestoga and Lewisburg likely reflects some combination of strong decline in source inputs and depletion of legacy sources.) Overall, our results reinforce the importance of considering lag time between the implementation

---

<sup>13</sup> Basu *et al.* (2010) have used the term “biogeochemical stationarity” for the phenomenon of low temporal variability in nutrient concentration, which was observed across a range of managed watersheds.

of management actions and achievement of river quality improvement. Such lag times may be on the order of years to decades for N and P (Jarvie *et al.*, 2013; Sanford and Pope, 2013; Sharpley *et al.*, 2013) and much longer for upland sediment management practices in watersheds with large transport-length scales (Pizzuto *et al.*, 2014).

#### **4.4.3. Mass Balances of Sub-basins and Effects of Streamflow on Export**

Mass-balance analysis of true-condition estimates reveals a striking similarity among all six sub-basins upstream of Conowingo Reservoir with respect to the timing of significant exports (Section 4.3.3; Figure 4.6). This suggests similar conditions of rainfall and material processing in SB1-SB6 and implies that  $Q_{\text{Annual}}$  is the principal factor controlling  $L_{\text{Annual}}$ , with relatively less influence from other factors (*e.g.*, seasonally-varying biogeochemical processes). The statistically significant linear slopes for  $\log[L_{\text{Annual}}] \sim \log[Q_{\text{Annual}}]$  confirm the dominance of hydrological control on the inter-annual variability of constituent exports (Figure C1), which suggests generally transport-limitation conditions, as has also been similarly observed with other watersheds (*e.g.*, Howarth *et al.*, 2006; Alvarez-Cobelas *et al.*, 2008; Alvarez-Cobelas *et al.*, 2009; Sobota *et al.*, 2009; Basu *et al.*, 2010; Howarth *et al.*, 2012).

Despite the above commonality of hydrological control, dissolved and particulate species exhibit markedly distinctive export behaviors based on the fitted linear slopes of  $\log[C_{\text{Annual-FW}}] \sim \log[(Q/A)_{\text{Annual}}]$ . Specifically, dissolved species are dominated by chemostasis effects (slope  $\sim 0$ ), whereas particulate species are dominated by mobilization effects (slope  $> 0$ ) (Figure 4.7; C2). A likely explanation is that dissolved species are dominated by processes of subsurface transport, storage, and mixing that are relatively homogeneous over a range of spatial and temporal scales (Gall *et al.*, 2013;

Kirchner and Neal, 2013; Harman, 2015), whereas particulate species are dominated by surface transport that are more susceptible to episodic exports. In both cases, the general lack of dilution patterns indicates that none of these constituents has been supply-limited, implying sufficient storage of excess constituent mass in these sub-basins. This finding is consistent with the “legacy sources” hypothesis discussed in Section 4.4.2.

In addition to affecting the annual export rates of dissolved and particulate constituents, streamflow has also played an important role in modulating the relative importance of dissolved and particulate fractions (Figure 4.8). This analysis was limited to annual estimates for describing general patterns and constraining seasonal effects. The observed negative correlations between  $(Q/A)_{\text{Annual}}$  and annual DP/TP and annual DN/TN ratios (Figure 4.8a-b) likely reflect surface mobilization of particulate (inorganic and organic) fractions during high discharges (Sharpley *et al.*, 1999; Pionke *et al.*, 2000). In this context, DN has always been the dominant fraction (>70%) of TN, whereas DP has been only a minor fraction of TP except during very low flows. The negative correlation between  $(Q/A)_{\text{Annual}}$  and annual PP/SS (Figure 4.8c) is expected given that (1) transported sediments contain a higher fraction of fine-sized particles during low flows and (2) finer sediments have higher specific surface areas for P absorption (Horowitz *et al.*, 2012; Zhang *et al.*, 2015). The negative correlation between streamflow and TN/TP molar ratio (Figure 4.8d-e) indicates that high-flow flushing affects P to a larger degree than N. In regard to the nominally limiting nutrient (Figure 4.8d-e), the lower TN/TP ratios during high discharges is an issue of interest that is consistent with the different export mechanisms for the two species. The spatial trend of daily TN/TP, which increases from upstream to downstream sites (Figure 4.8f), imply incrementally more net export of TN

than TP as one moves from upstream to downstream reaches of Susquehanna River over the study period. The implication is that the TN-to-TP mass ratio for contributions from the downstream river reaches and surrounding watersheds are greater than those from the upstream counterparts. Two possible scenarios, for example, could include (1) greater P contributions from steeper terrain in upstream watersheds or (2) greater N contributions from downstream agricultural lands. These are speculations only, however, and other scenarios are also possible. More definitive understanding would require additional data collection and study.

#### **4.4.4. Relative Contributions by Sub-basins and Effects of Land Use on Export**

The relative contributions of each sub-basin to total non-tidal SRB load are consistent with expectations based on relative drainage area and dominant land use (Section 4.3.4; Figure 4.9). For SB3 (West Branch Susquehanna River), its relatively large drainage area but low FCs for all types of constituents reflect the facts that SB3 has the highest fraction of forested area (81%) and that forested land should have relatively lower source inputs and higher assimilation capacity than non-forested land. By contrast, SB5 (Lower Susquehanna River above Marietta) has a relatively small area (60% of SB1) but high FCs for constituents, reflecting its larger fraction of agricultural area (47% compared to 35% in SB1) and smaller fraction of forested area (45% compared to 60% in SB1). This disproportionately larger contribution by SB5 is consistent with previous findings (Ator *et al.*, 2011) and deserves management considerations. In contrast with SB5, the patterns of SB7 (Lower Susquehanna River below Marietta) effectively demonstrate that nutrient/sediment export has been significantly modulated by major human modulation of the landscape, *i.e.*, river damming. Particularly, various N and P species can be at least

partially retained within the reservoir through particle sedimentation and algal uptake, followed by processes of bacterial degradation, denitrification, and burial in sediments (Jossette *et al.*, 1999; Friedl and Wüest, 2002). Moreover, the SB7 results also illustrate that such reservoir modulation (retention and release) has varied considerably as it approaches sediment storage capacity.

The follow-up analysis (Figure 4.10) has developed regression models between log-transformed period-of-record median yield and area fractions of land uses. The results show that period-of-record median yields of N, P, and SS all correlate positively with the area fraction of non-forested (*i.e.*, human-disturbed) but negatively with that of forested land (Figure 4.10). Moreover, these land-use effects are observed under all three flow classes, but with consistently larger intercepts during wet years (Figures C3-C4). This latter aspect may relate to both (a) increased mobilization of surface and sub-surface constituents and (b) decreased biogeochemical assimilation (*e.g.*, less denitrification or biotic uptake) that could result from shorter transit times during high-flow conditions (Howarth *et al.*, 2006; Alvarez-Cobelas *et al.*, 2009). These findings on land-use effects are particularly relevant to management of the SRB and also consistent with published findings elsewhere (*e.g.*, Jordan *et al.*, 1997b; Harris, 2001; Sobota *et al.*, 2009; Worrall *et al.*, 2012).

#### **4.5. Conclusions**

This paper provides a comprehensive evaluation of nutrient and sediment exports from multiple locations in the Susquehanna River basin over the last three decades. Our work has demonstrated the value of long-term data and the utility of “traditional” approaches of trend analysis, mass-balance calculation, and examination of  $C$ - $Q$  relations

for understanding riverine export. This synthesis of temporal and spatial patterns has provided information on four major factors affecting constituent export, namely, source input, reservoir modulation, streamflow, and land use, as summarized below:

- (1) “*Source input*”: Nutrient and sediment riverine loadings have generally declined at sites in Susquehanna River upstream of Conowingo Reservoir. These declines, particularly those of TN, seem to have followed source input reductions in the concurrent period of 1987-2011.<sup>14</sup> For both TN and TP, however, the generally  $<1.0$   $\Delta_{FN-Yield}/\Delta_{Input}$  ratios and the general temporal invariance of  $C_{Annual-FW}$  at these sites suggest the possibility of legacy contributions, as proposed by other investigators in prior major watershed studies. These results reinforce the importance of considering lag time between the implementation of management actions and achievement of river quality improvement.
- (2) “*Reservoir modulation*”: The contrast of mass-balance results in sub-basin SB7 with multiple upstream sub-basins effectively demonstrates how a major reservoir system (the LSRRS) has caused this sub-basin to behave far differently than any of the upstream reaches. As previously discussed in prior papers, the data indicate substantial retention of particulate species within the LSRRS, but with retention rates decreasing over time as the reservoir approaches sediment storage capacity. Consequently, flow-normalized loadings for particulate species have increased recently below Conowingo Reservoir, despite general declines at upstream sites.
- (3) “*Streamflow*”: Statistically significant linear  $\log[L_{Annual}] \sim \log[Q_{Annual}]$  relationships at

---

<sup>14</sup> Relative to our published paper (Zhang *et al.*, 2016), the underlined text is a more accurate statement of the intended concept.

the monitoring sites suggest the dominance of hydrological control on the inter-annual variability of constituent exports. The associated  $\log[C_{\text{Annual-FW}}] \sim \log[(Q/A)_{\text{Annual}}]$  patterns generally show chemostasis effects for dissolved species and mobilization effects for particulate species, both implying transport-limited (as opposed to source-limited) conditions. In addition to affecting annual export rates, streamflow has also affected the relative importance of dissolved and particulate fractions, as reflected by the negative correlations between  $(Q/A)_{\text{Annual}}$  and DP/TP, DN/TN, PP/SS, and TN/TP ratios.

- (4) “*Land use*”: The relative contributions of the sub-basins are consistent with expectations based on relative drainage area and dominant land use. Period-of-record median annual yields of N, P, and SS all correlate positively with the area fraction of non-forested land but negatively with that of forested land, and these patterns are observed under all hydrological classes.

These findings with respect to factors affecting riverine export are consistent with prior studies on a broad range of watersheds. These results have direct bearing toward better management of this large watershed and the attainment of Chesapeake Bay TMDLs. Moreover, our approaches are transferable to other Chesapeake tributaries and to rivers in other geographical regions. Last but not least, this work effectively illustrates how science-based management can benefit from maintaining open-access to high quality long-term monitoring data at multiple locations in watersheds.

#### **4.6. Supporting Information**

Supporting information to this chapter is provided in Appendix C. All derived data from this work, along with the raw river monitoring data, are stored at the publicly accessible Johns

Hopkins University Data Archive via <http://dx.doi.org/10.7281/T1QN64NW> (Zhang and Ball, 2016).

#### **4.7. Acknowledgements**

This work was supported by the Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090), Maryland Water Resources Research Center (2015MD329B), and National Science Foundation (CBET-1360415). We thank Gary Shenk and four anonymous reviewers for their constructive comments. We thank Gary Shenk, Guido Yactayo, and Gopal Bhatt (Chesapeake Bay Program Office) for providing the source input data. We acknowledge the U.S. Geological Survey and Susquehanna River Basin Commission for providing access to the river monitoring data.

#### **4.8. Literature Cited**

- Ai, L., Z. H. Shi, W. Yin and X. Huang, 2015. Spatial and seasonal patterns in stream water contamination across mountainous watersheds: Linkage with landscape characteristics. *Journal of Hydrology* 523:398-408, DOI: 10.1016/j.jhydrol.2015.01.082.
- Alvarez-Cobelas, M., D. G. Angeler and S. Sánchez-Carrillo, 2008. Export of nitrogen from catchments: A worldwide analysis. *Environ. Pollut.* 156:261-269, DOI: 10.1016/j.envpol.2008.02.016.
- Alvarez-Cobelas, M., S. Sánchez-Carrillo, D. G. Angeler and R. Sánchez-Andrés, 2009. Phosphorus export from catchments: a global view. *J. N. Am. Benthol. Soc.* 28:805-820, DOI: 10.1899/09-073.1.
- Ator, S. W., J. W. Brakebill and J. D. Blomquist, 2011. Sources, fate, and transport of

- nitrogen and phosphorus in the Chesapeake Bay watershed: An empirical model. U.S. Geological Survey Scientific Investigations Report 2011-5167, Reston, VA, p. 27. <http://pubs.usgs.gov/sir/2011/5167/>.
- Bachman, L. J., B. Lindsey, J. Brakebill and D. S. Powars, 1998. Ground-water discharge and base-flow nitrate loads of nontidal streams, and their relation to a hydrogeomorphic classification of the Chesapeake Bay Watershed, middle Atlantic coast. US Geological Survey Water-Resources Investigations Report 98-4059, Baltimore, MD, p. 71. <http://pubs.usgs.gov/wri/wri98-4059/>.
- Basu, N. B., G. Destouni, J. W. Jawitz, S. E. Thompson, N. V. Loukinova, A. Darracq, S. Zanardo, M. Yaeger, M. Sivapalan, A. Rinaldo and P. S. C. Rao, 2010. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophys. Res. Lett.* 37:L23404, DOI: 10.1029/2010gl045168.
- Belval, D. L. and L. A. Sprague, 1999. Monitoring nutrients in the major rivers draining to Chesapeake Bay. U.S. Geological Survey Water-Resources Investigations Report 99-4238, p. 8. [http://va.water.usgs.gov/online\\_pubs/WRIR/99-4238/wrir\\_99\\_4238\\_text.pdf](http://va.water.usgs.gov/online_pubs/WRIR/99-4238/wrir_99_4238_text.pdf).
- Brush, G. S., 2009. Historical Land Use, Nitrogen, and Coastal Eutrophication: A Paleocological Perspective. *Estuaries Coasts* 32:18-28, DOI: 10.1007/s12237-008-9106-z.
- Chanat, J. G., D. L. Moyer, J. D. Blomquist, K. E. Hyer and M. J. Langland, 2016. Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year

2012. U.S. Geological Survey U.S. Geological Survey Scientific Investigations Report 2015-5133, Reston, VA, p. 76. <http://dx.doi.org/10.3133/sir20155133>.
- Chen, D., M. Hu and R. A. Dahlgren, 2014. A dynamic watershed model for determining the effects of transient storage on nitrogen export to rivers. *Water Resour. Res.* 50:7714-7730, DOI: 10.1002/2014WR015852.
- Chen, D., M. Hu, Y. Guo and R. A. Dahlgren, 2015. Influence of legacy phosphorus, land use, and climate change on anthropogenic phosphorus inputs and riverine export dynamics. *Biogeochemistry* 123:99-116, DOI: 10.1007/s10533-014-0055-2.
- Chesapeake Bay Program, 1998. Chesapeake Bay watershed model application and calculation of nutrient and sediment loadings, Appendix F: Point source loadings. Annapolis, MD, p. 693.  
[http://www.chesapeakebay.net/content/publications/cbp\\_12313.pdf](http://www.chesapeakebay.net/content/publications/cbp_12313.pdf).
- Chesapeake Executive Council, 1988. Baywide nutrient reduction strategy: an agreement commitment report. Annapolis, MD.  
[http://www.chesapeakebay.net/content/publications/cbp\\_12114.pdf](http://www.chesapeakebay.net/content/publications/cbp_12114.pdf).
- Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.
- Eshleman, K. N., R. D. Sabo and K. M. Kline, 2013. Surface Water Quality Is Improving due to Declining Atmospheric N Deposition. *Environ. Sci. Technol.* 47:12193-12200, DOI: 10.1021/es4028748.
- Friedl, G. and A. Wüest, 2002. Disrupting biogeochemical cycles—Consequences of damming. *Aquat. Sci.* 64:55-65, DOI: 10.1007/s00027-002-8054-0.

- Friedrichs, C., T. Dillaha, J. Gray, R. Hirsch, A. Miller, D. Newburn, J. Pizzuto, L. Sanford, J. Testa, G. V. Houtven and P. Wilcock, 2014. Review of the Lower Susquehanna River Watershed Assessment. Chesapeake Bay Program Scientific and Technical Advisory Committee Report No. 14-006, Edgewater, Maryland, p. 40.
- Gall, H. E., J. Park, C. J. Harman, J. W. Jawitz and P. S. C. Rao, 2013. Landscape filtering of hydrologic and biogeochemical responses in managed catchments. *Landscape Ecol.* 28:651-664, DOI: 10.1007/s10980-012-9829-x.
- Gellis, A. C., C. R. Hupp, M. J. Pavich, J. M. Landwehr, W. S. L. Banks, B. E. Hubbard, M. J. Langland, J. C. Ritchie and J. M. Reuter, 2008. Sources, Transport, and Storage of Sediment at Selected Sites in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2008-5186, Reston, VA, p. 95. <http://pubs.usgs.gov/sir/2008/5186/>.
- Godsey, S. E., J. W. Kirchner and D. W. Clow, 2009. Concentration-discharge relationships reflect chemostatic characteristics of US catchments. *Hydrol. Process.* 23:1844-1864, DOI: 10.1002/hyp.7315.
- Green, C. T., B. A. Bekins, S. J. Kalkhoff, R. M. Hirsch, L. Liao and K. K. Barnes, 2014. Decadal surface water quality trends under variable climate, land use, and hydrogeochemical setting in Iowa, USA. *Water Resour. Res.* 50:2425-2443, DOI: <http://doi.wiley.com/10.1002/2013WR014829>.
- Grimm, J. W. and J. A. Lynch, 2005. Improved Daily Precipitation Nitrate and Ammonium Concentration Models for the Chesapeake Bay Watershed. *Environ. Pollut.* 135:445-455, DOI: 10.1016/j.envpol.2004.11.018.

- Hagy, J. D., W. R. Boynton, C. W. Keefe and K. V. Wood, 2004. Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658, DOI: 10.1007/bf02907650.
- Hale, R. L., N. B. Grimm, C. J. Vörösmarty and B. Fekete, 2015. Nitrogen and phosphorus fluxes from watersheds of the Northeast U.S. from 1930-2000: Role of anthropogenic nutrient inputs, infrastructure, and runoff. *Global Biogeochem. Cycles* 29:341-356, DOI: 10.1002/2014GB004909.
- Harman, C. J., 2015. Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed. *Water Resour. Res.* 51:1–30, DOI: 10.1002/2014WR015707.
- Harris, G. P., 2001. Biogeochemistry of nitrogen and phosphorus in Australian catchments, rivers and estuaries: Effects of land use and flow regulation and comparisons with global patterns. *Marine and Freshwater Research* 52:139-149, DOI: 10.1071/MF00031.
- Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.
- Hirsch, R. M. and L. De Cicco, 2015. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). U.S. Geological Survey Techniques and Methods Book 4, Chapter A10, Reston, VA, p. 93. <http://dx.doi.org/10.3133/tm4A10>.

- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Horowitz, A. J., V. C. Stephens, K. A. Elrick and J. J. Smith, 2012. Concentrations and annual fluxes of sediment-associated chemical constituents from conterminous US coastal rivers using bed sediment data. *Hydrol. Process.* 26:1090-1114, DOI: 10.1002/hyp.8437.
- Howarth, R., D. Swaney, G. Billen, J. Garnier, B. Hong, C. Humborg, P. Johnes, C.-M. Mörtz and R. Marino, 2012. Nitrogen fluxes from the landscape are controlled by net anthropogenic nitrogen inputs and by climate. *Front. Ecol. Environ.* 10:37-43, DOI: 10.1890/100178.
- Howarth, R. W., D. P. Swaney, E. W. Boyer, R. Marino, N. Jaworski and C. Goodale, 2006. The influence of climate on average nitrogen export from large watersheds in the Northeastern United States. *Biogeochemistry* 79:163-186, DOI: 10.1007/s10533-006-9010-1.
- Jarvie, H. P., A. N. Sharpley, B. Spears, A. R. Buda, L. May and P. J. A. Kleinman, 2013. Water Quality Remediation Faces Unprecedented Challenges from “Legacy Phosphorus”. *Environ. Sci. Technol.* 47:8997-8998, DOI: 10.1021/es403160a.
- Jordan, T. E., D. L. Correll and D. E. Weller, 1997. Relating nutrient discharges from watersheds to land use and streamflow variability. *Water Resour. Res.* 33:2579-2579, DOI: 10.1029/97wr02005.
- Jossette, G., B. Leporcq, N. Sanchez and Philippon, 1999. Biogeochemical mass-balances

- (C, N, P, SI) in three large reservoirs of the Seine basin (France). *Biogeochemistry* 47:119-146, DOI: 10.1007/bf00994919.
- Keisman, J., J. Blomquist, S. Phillips, G. Shenk and E. Yagow, 2015. Estimating Land Management Effects on Water Quality Status and Trends. *Scientific and Technical Advisory Committee Publication Number 14-009*. Edgewater, Maryland, p. 28.
- Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C. Stevenson, 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* 303:1-29, DOI: 10.3354/meps303001.
- Kirchner, J. W. and C. Neal, 2013. Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc. Natl. Acad. Sci. U. S. A.* 110:12213-12218, DOI: 10.1073/pnas.1304328110.
- Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18. <http://dx.doi.org/10.3133/ofr20141235>.
- Langland, M. J., P. L. Lietman and S. Hoffman, 1995. Synthesis of nutrient and sediment data for watersheds within the Chesapeake Bay drainage basin. U.S. Geological Survey Water-Resources Investigations Report 95-4233, Lemoyne, PA, p. 121. <http://pubs.er.usgs.gov/publication/wri954233>.
- Linker, L. C., R. A. Batiuk, G. W. Shenk and C. F. Cerco, 2013a. Development of the

- Chesapeake Bay Watershed Total Maximum Daily Load Allocation. *J. Am. Water Resour. Assoc.* 49:986-1006, DOI: 10.1111/jawr.12105.
- Linker, L. C., R. Dennis, G. W. Shenk, R. A. Batiuk, J. Grimm and P. Wang, 2013b. Computing Atmospheric Nutrient Loads to the Chesapeake Bay Watershed and Tidal Waters. *J. Am. Water Resour. Assoc.* 49:1025-1041, DOI: 10.1111/jawr.12112.
- Litke, D. W., 1999. Review of phosphorus control measures in the United States and their effects on water quality. U.S. Geological Survey Water-Resources Investigations Report 99-4007, Denver, CO, p. 43. <http://pubs.usgs.gov/wri/wri994007/>.
- Meals, D. W., S. A. Dressing and T. E. Davenport, 2009. Lag time in water quality response to best management practices: a review. *J. Environ. Qual.* 39:85-96, DOI: 10.2134/jeq2009.0108.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Murphy, R. R., W. M. Kemp and W. P. Ball, 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* 34:1293-1309, DOI: 10.1007/s12237-011-9413-7.
- Pionke, H. B., W. J. Gburek and A. N. Sharpley, 2000. Critical source area controls on water quality in an agricultural watershed located in the Chesapeake Basin. *Ecol. Eng.* 14:325-335, DOI: 10.1016/S0925-8574(99)00059-2.
- Pizzuto, J., E. R. Schenk, C. R. Hupp, A. Gellis, G. Noe, E. Williamson, D. L. Karwan,

- M. O'Neal, J. Marquard, R. Aalto and D. Newbold, 2014. Characteristic length scales and time-averaged transport velocities of suspended sediment in the mid-Atlantic Region, USA. *Water Resour. Res.* 50:790-805, DOI: 10.1002/2013WR014485.
- Pritchard, D. W. and J. R. Schubel, 2001. Human influences on physical characteristics of the Chesapeake Bay. *In: Discovering the Chesapeake: the history of an ecosystem*, P. D. Curtin, G. S. Brush and G. W. Fisher (Editors). The Johns Hopkins University Press, Baltimore, MD, pp. 60-82.
- Qian, S. S., M. E. Borsuk and C. A. Stow, 2000. Seasonal and Long-Term Nutrient Trend Decomposition along a Spatial Gradient in the Neuse River Watershed. *Environ. Sci. Technol.* 34:4474-4482, DOI: 10.1021/es000989p.
- Redfield, A. C., 1958. The biological control of chemical factors in the environment. *Am. Sci.* 46:205-221, <http://www.jstor.org/stable/27827150>.
- Sanford, W. E. and J. P. Pope, 2013. Quantifying Groundwater's Role in Delaying Improvements to Chesapeake Bay Water Quality. *Environ. Sci. Technol.* 47:13330-13338, DOI: 10.1021/es401334k.
- Sharpley, A., H. P. Jarvie, A. Buda, L. May, B. Spears and P. Kleinman, 2013. Phosphorus Legacy: Overcoming the Effects of Past Management Practices to Mitigate Future Water Quality Impairment. *J. Environ. Qual.* 42:1308–1326, DOI: 10.2134/jeq2013.03.0098.
- Sharpley, A. N., W. J. Gburek, G. Folmar and H. B. Pionke, 1999. Sources of phosphorus exported from an agricultural watershed in Pennsylvania. *Agric. Water Manage.* 41:77-89, DOI: 10.1016/s0378-3774(99)00018-9.

- Shenk, G. W., 2015. Personal communication.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- Sobota, D. J., J. A. Harrison and R. A. Dahlgren, 2009. Influences of climate, hydrology, and land use on input and export of nitrogen in California watersheds. *Biogeochemistry* 94:43-62, DOI: 10.1007/s10533-009-9307-y.
- Sprague, L. A., M. J. Langland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. U.S. Geological Survey Water-Resources Investigations Report 00-4218, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).
- Stallard, R. F. and S. F. Murphy, 2014. A Unified Assessment of Hydrologic and Biogeochemical Responses in Research Watersheds in Eastern Puerto Rico Using Runoff-Concentration Relations. *Aquat. Geochem.* 20:115-139, DOI: 10.1007/s10498-013-9216-5.
- Susquehanna River Basin Commission, 2014. Sediment and nutrient assessment program.  
<http://www.srbc.net/programs/cbp/nutrientprogram.htm>.
- The Lower Susquehanna River Watershed Assessment Team, 2014. Lower Susquehanna River Watershed Assessment, MD and PA: Phase I Draft Report. p. 185.  
<http://mddnr.chesapeakebay.net/LSRWA/report.cfm>.
- Thompson, S. E., N. B. Basu, J. Lascrain, A. Aubeneau and P. S. C. Rao, 2011. Relative dominance of hydrologic versus biogeochemical factors on solute export across

- impact gradients. *Water Resour. Res.* 47:W00J05, DOI: 10.1029/2010WR009605.
- U.S. Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment. Annapolis, MD.  
<http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html>.
- U.S. Geological Survey, 2014a. Surface-water data for the nation.
- U.S. Geological Survey, 2014b. Water quality loads and trends at nontidal monitoring stations in the Chesapeake Bay watershed. <http://cbrim.er.usgs.gov>.
- University of Maryland Center for Environmental Science, 2016. UMCES Scientists to Study Water Quality Consequences of Susquehanna River Sediments and Nutrients.
- Walter, R. C. and D. J. Merritts, 2008. Natural streams and the legacy of water-powered mills. *Science* 319:299-304, DOI: 10.1126/science.1151716.
- Weld, J. L., R. L. Parsons, D. B. Beegle, A. N. Sharpley, W. J. Gburek and W. R. Clouser, 2002. Evaluation of phosphorus-based nutrient management strategies in Pennsylvania. *J. Soil Water Conserv.* 57:448-454,  
<http://www.jswconline.org/content/57/6/448.short>.
- Worrall, F., H. Davies, T. Burt, N. J. K. Howden, M. J. Whelan, A. Bhogal and A. Lilly, 2012. The flux of dissolved nitrogen from the UK - Evaluating the role of soils and land use. *Sci. Total Environ.* 434:90-100, DOI: 10.1016/j.scitotenv.2012.01.035.
- Yactayo, G., 2015. Personal communication.
- Zhang, Q. and W. P. Ball, 2016. Data associated with Decadal-scale export of nitrogen, phosphorus, and sediment from the Susquehanna River basin, USA: Analysis and

synthesis of temporal and spatial patterns. Johns Hopkins University Data Archive, Baltimore, MD. <http://dx.doi.org/10.7281/T1QN64NW>, DOI: 10.7281/T1QN64NW.

Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI: 10.1016/j.scitotenv.2013.02.012.

Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.

Zhang, Q., R. M. Hirsch and W. P. Ball, 2016. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* 50:1877-1886, DOI: 10.1021/acs.est.5b04073.

*Page intentionally left blank*

## Chapter 5. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season<sup>15</sup>

### Abstract

To assess historical loads of nitrogen (N), phosphorus (P), and suspended sediment (SS) from the nontidal Chesapeake Bay watershed (NTCBW), we analyzed decadal seasonal trends of flow-normalized loads at the fall-line of nine major rivers that account for >90% of NTCBW flow. Evaluations of loads by season revealed N, P, and SS load magnitudes have been highest in Jan-Mar and lowest in Jul-Sep, but the temporal trends have followed similar decadal-scale patterns in all seasons, with notable exceptions. Generally, total N (TN) load has dropped since the late 1980s, but particulate nutrients and SS have risen since the mid-1990s. The majority of these rises were from Susquehanna River and relate to diminished net trapping at Conowingo Reservoir. Substantial rises in SS were also observed, however, in other rivers. Moreover, the summed rise in particulate P load from other rivers is of similar magnitude as from Susquehanna. Dissolved nutrient loads have dropped in the upland (Piedmont-and-above) rivers, but risen in two small rivers in the Coastal Plain affected by lagged groundwater input. In addition, analysis of fractional contributions revealed consistent N trends across

---

<sup>15</sup> This chapter (Abstract through Section 5.10) has been published as: Zhang, Q.; Brady, D. C.; Boynton, W.; Ball, W. P., Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* **2015**, *51*, (6), 1534–1555, doi: [10.1111/1752-1688.12327](https://doi.org/10.1111/1752-1688.12327). Co-authors Brady and Boynton were involved in results interpretation and editing. Co-author Ball was involved in hypothesis development, study design, results interpretation, and editing. Copyright 2015 Wiley. Reproduced/modified by permission of Wiley. All figures, tables, and data were created by Qian Zhang unless otherwise indicated. Section, table, and figure numbers have been modified. References have also been re-formatted for consistency among chapters. Some minor substantive changes from the published manuscript have been made and are identified with footnotes.

the upland watersheds. Finally, TN:TP ratios have declined in most rivers, suggesting the potential for changes in nutrient limitation. Overall, this integrated study of historical data highlights the value of maintaining long-term monitoring at multiple watershed locations.

## **5.1. Introduction**

### **5.1.1. Chesapeake Bay Hypoxia and Watershed Inputs Control**

Chesapeake Bay, North America's largest estuary and a major national economic resource, has experienced persistent and increasing summer hypoxia and anoxia in the mesohaline region as a result of (1) inputs of anthropogenic nutrients and sediment from the watershed (Boynton and Kemp, 2000; Hagy *et al.*, 2004; Kemp *et al.*, 2005; Murphy, 2014; Murphy *et al.*, 2014) and (2) vertical stratification of the water-column that arises from interactions between seaward-moving surface freshwater and landward-moving bottom saltwater (Boicourt, 1992; Pritchard and Schubel, 2001; Murphy *et al.*, 2011). The resulting pycnocline seasonally isolates deep-water zones and prevents dissolved oxygen (DO) replenishment from surface waters.

This work focuses on evaluations of historical loadings of nitrogen (N), phosphorus (P), and suspended sediment (SS) from the nontidal Chesapeake Bay watershed (NTCBW). These inputs originate from a variety of sources, including agricultural lands, stormwater runoff, wastewater treatment plants (WWTPs), and atmospheric deposition (mainly for N). Nutrient inputs stimulate the growth of algae, which subsequently sink to deeper waters and consume DO at rates exceeding DO supply (Boynton and Kemp, 2000). In addition, nutrient and sediment inputs can reduce light penetration and thus inhibit the growth of submerged aquatic vegetation (Brakebill *et al.*, 2010), which has the

beneficial functions of benthic oxygen production (via photosynthesis), water clarification (via particle filtering), and nutrient sequestration (Kemp *et al.*, 2005).

Recently, the endeavor to reduce nutrient and sediment loadings has been reinforced with the promulgation of regulations on Total Maximum Daily Loads (TMDLs) (U.S. Environmental Protection Agency, 2010) and state-wide efforts to establish watershed implementation plans (WIPs) (Linker *et al.*, 2013a; Shenk and Linker, 2013). In this context, better understanding of the nature of historical loading changes from different regions of the watershed will be essential for developing future management strategies.

### **5.1.2. Statistical Methods for Riverine Loading Estimation**

Selection of methods for estimating nutrient and sediment concentrations and loadings based on low-frequency (*e.g.*, monthly) monitoring data has been an important research topic for decades (Cohn *et al.*, 1989). Historically, the U.S. Geological Survey (USGS) has applied a method known as the ESTIMATOR model (Cohn *et al.*, 1989). Using this method, for instance, Langland *et al.* (2007) reported decreasing trends in flow-adjusted annual concentrations of total N (TN), total P (TP), and SS from 1985 to 2006 at about 74, 68, and 32% of 34 sites in the NTCBW, respectively.

More recently, Hirsch *et al.* (2010) have developed a new method called “Weighted Regressions on Time, Discharge, and Season” (WRTDS). Compared to the ESTIMATOR model, WRTDS can (1) better describe temporal changes in both concentration and load, and (2) more effectively remove the influence of random flow variation. Detailed comparison of WRTDS and prior approaches developed and published by USGS investigators (*e.g.*, the five-parameter LOADEST and seven-parameter LOADEST) are addressed by Moyer *et al.* (2012) and Hirsch (2014). Overall,

the improvements of WRTDS derive in part from the fact that WRTDS does not rely on problematic assumptions about homoscedasticity of model errors, constancy of seasonal trends in concentration, or constancy of the concentration-flow relationship (Hirsch *et al.*, 2010; Moyer *et al.*, 2012; Hirsch, 2014). Instead, the dependencies of concentrations on time, discharge, and season are re-evaluated for each day of estimation in WRTDS, and hence WRTDS is more nonparametric and data-driven than its preceding models.

Therefore, we adopted WRTDS for this work.

### **5.1.3. Insights Gained from Prior Studies on Conowingo Reservoir**

For evaluating nutrient and sediment inputs in the Chesapeake watershed, an important factor to consider is the Lower Susquehanna River Reservoir System (LSRRS), which is located on the Bay's largest tributary. This system historically traps about 2, 45, and 70% of annual TN, TP, and SS loads from Susquehanna River to Chesapeake Bay, respectively (Langland and Hainly, 1997). The Conowingo Reservoir is the most downstream reservoir and the only one that reportedly has not reached its sediment storage capacity (Langland, 2009). Recently, Hirsch (2012) detected upward trends in annual TP and SS loads at the Conowingo Dam (the LSRRS outlet), suggesting that the system may not be as efficient as it used to be in trapping upstream inputs. Zhang *et al.* (2013) further evaluated the respective trends of dissolved and particulate species at both the LSRRS inlet and outlet and found that loadings of these species at the system inlet have generally decreased in all seasons, which reflects declining inputs from some combination of agricultural, point, atmospheric, and stormwater sources in the above-reservoir watershed. In comparison, only dissolved nutrients have decreased at the system outlet, with increases in particulate nutrients and SS observed since the mid-1990s. The

latter phenomenon reflects diminished system capacity for trapping sediment and particulate nutrients; however, the nature and magnitude of these Susquehanna trends have not yet been evaluated relative to those in the other major tributaries and the overall NTCBW.

#### **5.1.4. Aims and Contributions of This Work**

We have used the rich dataset available for the NTCBW to perform an integrated and comprehensive comparison of historical data among sub-watersheds to better understand the complex system changes in the watershed. As an extension to the Susquehanna River study of Zhang *et al.* (2013), this work focuses on historical trends of nutrient and sediment loadings from all nine major tributaries of the NTCBW. The data records vary with sites and constituents but include the last twenty to thirty years. This work was aimed at providing the following new contributions:

- (1) Analysis of historical nontidal streamflow and concentration monitoring data for all nine major Chesapeake tributaries, their summed loadings, and the fractional contribution of each tributary to the summed total.
- (2) Separate analysis of multiple forms of N and P. Major species include SS, TP, TN, particulate P (PP), particulate N (PN), dissolved P (DP), and dissolved N (DN). In comparison, most prior studies (*e.g.*, Langland *et al.*, 2007; Hirsch *et al.*, 2010; Langland *et al.*, 2012) focused on only SS, TP, and TN. The recent work of Moyer *et al.* (2012) analyzed two sub-species (*i.e.*, dissolved orthophosphate and nitrate-plus-nitrite); however, total dissolved and particulate fractions were not explicitly examined.
- (3) Analysis of seasonal loadings and trends. Previous studies have focused on annual

trends in concentration or load. Our intent is to isolate the possible impacts of seasonality in streamflow from the variety of other possible factors (*e.g.*, temperature impacts on biogeochemical processes and timing of fertilizer application).

## **5.2. Data and Methods**

### **5.2.1. Study Area: the NTCBW**

Chesapeake Bay extends 300 km from the seaward end at the Virginia capes (Cape Charles and Cape Henry) to the mouth of Susquehanna River at Havre de Grace, Maryland. The Bay estuarine system comprises the Bay proper and more than fifty tributaries (Pritchard and Schubel, 2001). To assess nutrient and sediment trends in these tributaries, the USGS, state agencies, and other collaborators have been monitoring water quality at many sites (Langland *et al.*, 2007). In particular, the USGS River Input Monitoring (RIM) Program has been monitoring sites at the fall-line of nine major tributaries since the 1980s (Figure 5.1; Table 5.1), namely, Susquehanna, Potomac, James, Rappahannock, Appomattox, Pamunkey, Mattaponi, Patuxent, and Choptank Rivers. These rivers account for more than 90% of streamflow from the NTCBW (USGS, 2012b). This represents a substantial fraction of total Bay input but excludes the wastewater discharge, nonpoint source runoff, and groundwater flow that discharge into tidally influenced regions of the Bay and its tributaries. The NTCBW excludes, for example, some major portions of the Coastal Plain (CP) aquifers on the Delmarva Peninsula (Figure 5.1). However, Choptank River on the Eastern Shore is included, which is located entirely in the CP and thus may be indicative of that region. This river

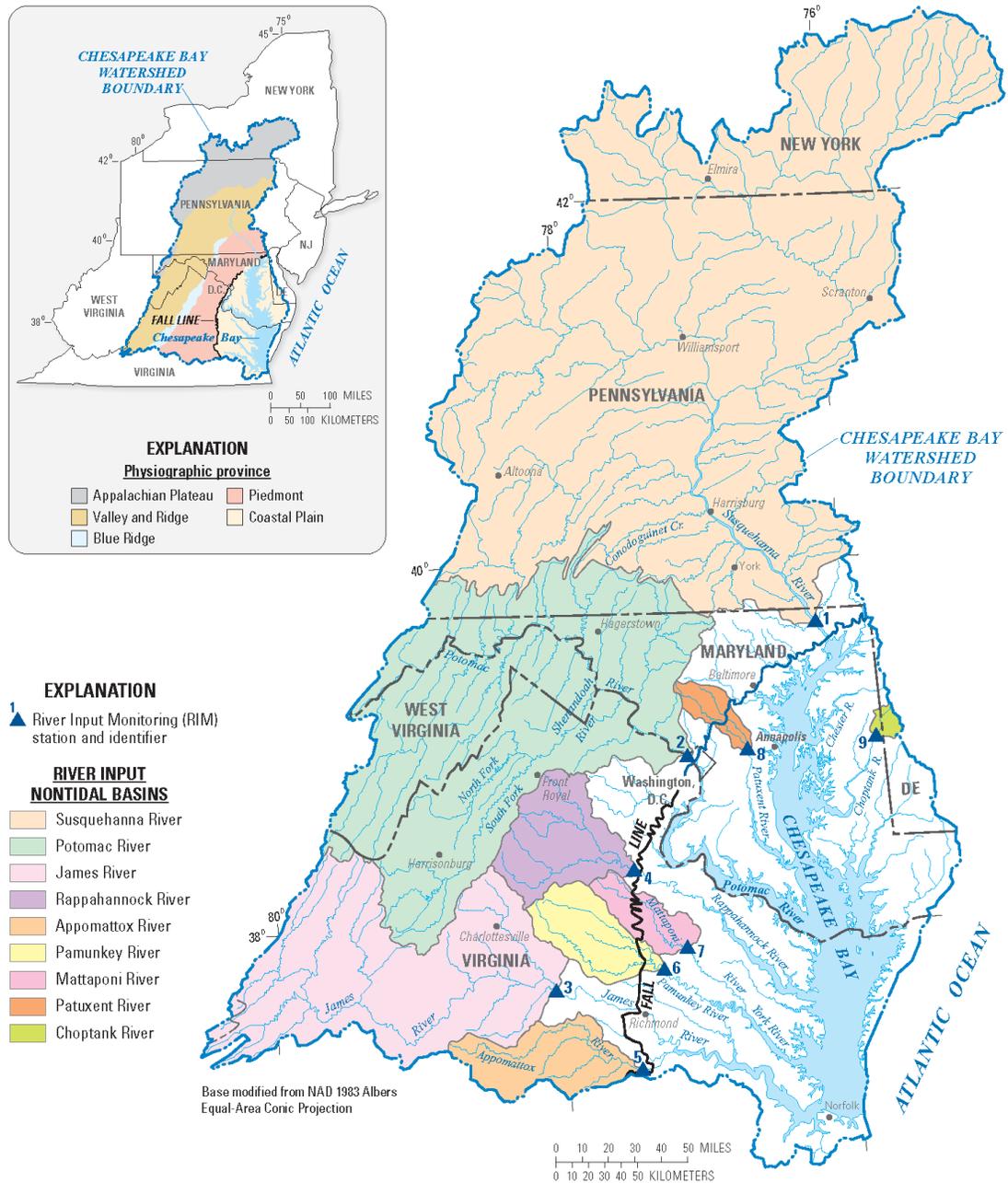


Figure 5.1. Chesapeake Bay watershed and monitoring sites at the fall-line of the nine major tributaries — Susquehanna River (1), Potomac River (2), James River (3), Rappahannock River (4), Appomattox River (5), Pamunkey River (6), Mattaponi River (7), Patuxent River (8), and Choptank River (9). Inset shows the major physiographic provinces. This figure was reproduced from Moyer *et al.* (2012) with permission.

Table 5.1. Details of the monitoring sites.

USGS site number	Site and river name	Upstream physiographic provinces (percent) <sup>a</sup>					Upstream land use (percent) <sup>b</sup>				Mean annual flow statistics (cubic meter per second) <sup>c</sup>			
		CP	PM	BR	VR	AP	Urban	Agricultural	Forested	Other	Min	Median	Mean	Max
01578310	Susquehanna River near Conowingo, MD	0	9	0.9	32	58	2	29	67	2	667	1007	1145	2041
01646580	Potomac River at Chain Bridge, Washington D.C.	0	14	9	70	8	3	35	61	1	181	303	347	787
02035000	James River at Cartersville, VA	0	39	14	46	0	1	16	80	3	91	196	200	387
01668000	Rappahannock River near Fredericksburg, VA	0	70	30	0.1	0	1	36	61	2	18	45	49	110
02041650	Appomattox River at Matoaca, VA	0	100	0	0	0	1	20	72	7	11	31	34	84
01673000	Pamunkey River near Hanover, VA	2	98	0	0	0	1	24	68	7	8	25	28	57
01674500	Mattaponi River near Beulahville, VA	42	58	0	0	0	1	19	69	11	4	15	15	30
01594440	Patuxent River at Bowie, MD	28	72	0	0	0	13	41	38	8	5	10	11	23
01491000	Choptank River near Greensboro, MD	100	0	0	0	0	1	50	29	20	2	4	4	9

Notes: CP, Coastal Plain; PM, Piedmont; BR, Blue Ridge; VR, Valley and Ridge; AP, Appalachian Plateau.

<sup>a</sup> estimated based on physiography data from Fenneman and Johnson (1946).

<sup>b</sup> reproduced from Table 3 in Sprague *et al.* (2000).

<sup>c</sup> calculated based on annual flow data from 1979 to 2012 (USGS, 2013).

basin is largely agricultural land drained mainly through ditches (Sprague *et al.*, 2000). On the Western Shore, only Mattaponi River draws a large portion of its water from CP areas. The other river basins are dominated by four different “upland” (*i.e.*, Piedmont and above) physiographic provinces, as shown in Figure 5.1 and quantified in Table 5.1. The Piedmont, like CP, is dominated by agricultural activity. The Blue Ridge is an area of high relief but also contains some intense farm operations. The Valley and Ridge (VR) covers the western parts of the Susquehanna, Potomac, and James River basins. Further upland, the Appalachian Plateau (AP) covers the northern part of the Susquehanna basin. These last two provinces (VR and AP) are characterized by a combination of forests on mountain slopes and agriculture in the valleys (Shenk and Linker, 2013).

### 5.2.2. Statistical Methods: WRTDS

The WRTDS method is a central tool used in this work. The approach uses daily river flow and less frequent (bi-weekly or monthly) nutrient and sediment concentration data to estimate concentrations and loadings for every day in the record. For each estimation day, the method pre-screens all sampled data to select at least 100 samples that are sufficiently “close” to the estimation day in three dimensions (*i.e.*, time, discharge, and season), and then fits the following equation:

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (5.1)$$

where  $C$  is nutrient or sediment concentration,  $Q$  is daily discharge,  $t$  is time in decimal years,  $\beta_i$  are fitted coefficients, and  $\varepsilon$  is the error term. The adoption of logarithmic forms of  $C$  and  $Q$  in many estimation models (*e.g.*, Cohn *et al.*, 1989; Johnes, 2007; Wang *et al.*, 2015) is due to (1) commonly observed power-law  $C$ - $Q$  relationships and (2) that the residuals are more nearly symmetrical when the dependent variable is  $\ln(C)$  than when it

is  $C$  (e.g., Hirsch *et al.*, 1991; Helsel and Hirsch, 2002). The fitted coefficients in Equation (5.1) are used to estimate the log-concentration on the estimation day with known values of  $t$  and  $Q$ , which is then transformed back to  $C$ . To account for the re-transformation bias, the so-called “smearing factor” has been used in ESTIMATOR and earlier versions of WRTDS -- see Cohn (2005) and Hirsch *et al.* (2010) for details. In the version of WRTDS used in this study, a different approach was adopted to accommodate the presence of censored data -- see the “Censored Data” Section in Moyer *et al.* (2012) for details and justifications. Finally, the estimated daily  $C$  is multiplied by daily discharge to get the daily loading estimate. This method has been fully described elsewhere in context of Chesapeake Bay (Hirsch *et al.*, 2010; Hirsch, 2012; Zhang *et al.*, 2013), Lake Champlain (Medalie *et al.*, 2012), and Mississippi River Basin (Sprague *et al.*, 2011). A flow-chart description of the WRTDS estimation process can be found in the Appendix A of Zhang *et al.* (2013). The WRTDS method has been implemented in the R package called EGRET (Exploration and Graphics for RivEr Trends; version 1.2.5), which is available on the Comprehensive R Archive Network at <http://cran.r-project.org/web/packages/>. A user guide for the package (Hirsch and De Cicco, 2014) is available from the USGS at <http://pubs.usgs.gov/tm/04/a10/>.

In general, WRTDS can produce two types of estimates for both concentration and load – so-called “*true-condition*” and “*flow-normalized*” estimates. The true-condition estimates are useful to understand the real history of water quantity and quality relevant to downstream ecological impact, while the flow-normalized estimates are helpful to evaluate effectiveness of management actions (Hirsch *et al.*, 2010). More specifically, the flow-normalized estimates are designed to effectively remove the sometimes-dramatic

influence of year-to-year random variations in streamflow (but not the seasonal variations in discharge or other non-discharge effects) by integrating streamflow over the long-term hydrological cycle, thus providing much smoother inter-annual variations and making the non-discharge-related trends easier to detect. The WRTDS flow-normalization (FN) method is a USGS-supported and -adopted approach that was developed to better detect trends in riverine loads. In comparison, the previously-used ESTIMATOR-based flow-adjusted concentration (FAC) method is much less appropriate for detecting trends in loads because it assumes that the fractional changes in concentration across dates are the same across all discharges and seasons (Moyer *et al.*, 2012). Moyer *et al.* (2012) have compared the two methods in the context of annual loading trends in the RIM rivers and observed consistent trend results between the two methods in many cases – see their Tables 3-7 for detailed comparisons. For the cases with observed inconsistencies, the authors suggested that FAC trends are too heavily influenced by concentrations during low-to-moderate discharges, whereas FN trends are more heavily influenced by concentrations associated with high discharges. Moreover, FAC is not suitable to disentangle trends among seasons. Given the more pressing need to understand high flux conditions and our emphasis on seasonal loadings, we have chosen to focus on FN for seasonal trend analysis.

Although WRTDS represents the latest generation of methods for water-quality trend analysis, Hirsch *et al.* (2010) listed several aspects of the method that need further improvement. Perhaps the most important of these is a major assumption of the flow-normalization method that the statistics of streamflow are stationary during the study period; however, there is currently no formal procedure to accommodate departures from

this assumption. Given that the analyses conducted here cover only about three decades, and there have been no major changes in water storage and diversion over that time, we believe that this assumption of stationarity in discharge is not problematic, although it could become problematic in the future as climate change and water management change become larger influences on discharge (Hirsch *et al.*, 2010). In addition, there is no published method yet for assigning uncertainty to WRTDS-estimated trends given its original purpose as a tool for exploratory data analysis. However, Hirsch and USGS colleagues are currently developing a prototype “bootstrap” method, based on resampling (with replacement) from the data and re-estimating the model. This critical addition to the WRTDS method will provide one measure of uncertainty that can help clarify if apparent loading trends are real departures from past system behavior due to hopefully identifiable physical processes or whether by simple variations that can be expected to arise by chance from a stationary system (Robert Hirsch, personal communication, 2014). So far, some needed improvements have been accomplished, including addition of capability for analyzing censored data and simultaneous analysis of multiple constituents. This work has taken advantage of these changes in the WRTDS code, which were graciously made available to us by USGS staff. (See Acknowledgements.)

### **5.2.3. Data Compilation and Analyses**

We collected water-quality and daily streamflow data at the nine RIM stations (Figure 5.1) from the USGS National Water Information System Web Interface (U.S. Geological Survey, 2013). The water-quality data included information for several nutrient and sediment constituents. A summary of the periods of record for each constituent is provided in Appendix D1. These record start as early as 1973 and extend to early 2013.

WRTDS was implemented using the statistical package R version 3.1.0 (R Development Core Team, 2014) to produce both the true-condition and flow-normalized estimates for every day in the period of record for five major species, namely, SS (Table D3), DP (Table D4), TP (Table D6), DN (Table D7), and TN (Table D9). For each of the 45 WRTDS runs (*i.e.*, TN, DN, TP, DP, and SS for each of the nine sites), residual plots were generated to evaluate the model performance. More specifically, we plotted residuals (observed log concentration minus estimated log concentration) against (a) estimated concentrations, (b) time, (c) discharges, and (d) months, respectively. Visual inspection of all residual plots indicates that there is a clear lack of trend in these plots, confirming that the unaccounted residuals have no substantial additional structural relationship with time, discharge, or season. A representative subset of the residual plots are provided in Appendix D6 and the complete set is stored at the publicly accessible Johns Hopkins University Data Archive (Zhang and Ball, 2014).

The daily loading estimates in units of kg/day were used to calculate the seasonal averages for each of four quarters of the year that approximately correspond to seasons: January-March (“winter” or S1), April-June (“spring” or S2), July-September (“summer” or S3), and October-December (“fall” or S4). For particulate nutrients, *i.e.*, PP (Table D5) and PN (Table D8), concentration observations were limited at all Maryland sites and thus prevented us from obtaining reliable long-term trends. Alternatively, seasonal loads of PP and PN at these sites were inferred by subtracting DP and DN from TP and TN loads, respectively. These plots are subsequently indicated as “inferred”. For consistency, the same approach has been applied to the Virginia sites, although PP and PN data were more available at those sites. In addition, all Virginia sites lacked DN and PN data prior

to 1995 (Tables D7-D8), thus the summed DN load from the NTCBW prior to 1995 were estimated from results at the four Maryland sites (Susquehanna plus Potomac plus Patuxent plus Choptank). The Maryland sites are known to have accounted for 93.3, 93.4, 91.0, and 92.9% of total loadings from all nine rivers in S1, S2, S3, and S4, respectively, as based on 1996-2012 results. Total loadings prior to 1995 were adjusted accordingly (*i.e.*,  $1/0.933$ ,  $1/0.934$ ,  $1/0.910$ , and  $1/0.929$ ) under the assumption that these ratios are also characteristic of the earlier period. Moreover, there were no SS data at the Virginia sites for 1996-2000 (Table D3) and thus estimates are less accurate around that period and are shown as dashed lines on plots.

The summation of flow-normalized loads from all nine fall-line sites, hereafter referred to as “NTCBW loading”, is taken as an overall measure of trends in the nontidal Bay watershed. We emphasize, however, the NTCBW loading does not include many contributions from some major WWTPs, nonpoint sources, and groundwater discharges that occur below river fall-lines. Although such flows account for only 20% of total estuarine flow, their estimated contributions to nutrient loadings are much greater – 35% for N and 33% for P (Shenk and Linker, 2013). Using the NTCBW input as a common denominator, we have calculated the fractional contributions of streamflow, N, P, and SS to the NTCBW from each of the nine tributaries. This allows an assessment of relative differences in changes of nontidal loading from the nine rivers. Finally, we have calculated the N:P load ratios to identify possible changes in the limiting nutrient to algal growth.

For better comparison among tributaries, all loading estimates (kg/day) were converted to yields ( $\text{kg yr}^{-1} \text{ hectare}^{-1}$ ) using the respective drainage areas (Table 5.2). In

Table 5.2. Summary of long-term median values of annual loads and yields in the nontidal Chesapeake Bay watershed (NTCBW) and nine tributaries. <sup>a</sup>

Watersheds	Drainage area (km <sup>2</sup> )	Long-term median of annual load							Long-term median of annual yield						
		TP	PP	DP	SS	TN	PN	DN	TP	PP	DP	SS	TN	PN	DN
		(kg P/day)			(kg/day)		(kg N/day)		(kg P/yr/hectare)			(kg/yr/hectare)		(kg N/yr/hectare)	
NTCBW	129,585	1.7 x10 <sup>4</sup>	1.3 x10 <sup>4</sup>	3.9 x10 <sup>3</sup>	1.3 x10 <sup>7</sup>	2.5 x10 <sup>5</sup>	4.5 x10 <sup>4</sup>	2.1 x10 <sup>5</sup>	0.49	0.36	0.11	360	7.2	1.3	5.8
NTCBW-SUS <sup>b</sup>	59,396	9.8 x10 <sup>3</sup>	7.5 x10 <sup>3</sup>	2.4 x10 <sup>3</sup>	7.8 x10 <sup>6</sup>	9.0 x10 <sup>4</sup>	2.6 x10 <sup>4</sup>	6.3 x10 <sup>4</sup>	0.60	0.46	0.14	479	5.5	1.6	3.9
Susquehanna	70,189	7.2 x10 <sup>3</sup>	5.3 x10 <sup>3</sup>	1.6 x10 <sup>3</sup>	4.3 x10 <sup>6</sup>	1.6 x10 <sup>5</sup>	2.0 x10 <sup>4</sup>	1.4 x10 <sup>5</sup>	0.38	0.28	0.08	223	8.5	1.0	7.5
Potomac	30,044	4.7 x10 <sup>3</sup>	3.7 x10 <sup>3</sup>	1.1 x10 <sup>3</sup>	4.1 x10 <sup>6</sup>	6.4 x10 <sup>4</sup>	1.5 x10 <sup>4</sup>	4.8 x10 <sup>4</sup>	0.58	0.45	0.14	496	7.8	1.8	5.9
James	16,213	3.2 x10 <sup>3</sup>	2.5 x10 <sup>3</sup>	8.4 x10 <sup>2</sup>	2.3 x10 <sup>6</sup>	1.5 x10 <sup>4</sup>	7.0 x10 <sup>3</sup>	7.4 x10 <sup>3</sup>	0.73	0.55	0.19	519	3.3	1.6	1.7
Rappahannock	4,144	8.1 x10 <sup>2</sup>	6.8 x10 <sup>2</sup>	1.3 x10 <sup>2</sup>	8.3 x10 <sup>5</sup>	5.6 x10 <sup>3</sup>	2.1 x10 <sup>3</sup>	3.5 x10 <sup>3</sup>	0.71	0.60	0.11	732	5.0	1.8	3.1
Appomattox	3,471	1.8 x10 <sup>2</sup>	1.2 x10 <sup>2</sup>	6.2 x10 <sup>1</sup>	5.2 x10 <sup>4</sup>	1.9 x10 <sup>3</sup>	5.8 x10 <sup>2</sup>	1.3 x10 <sup>3</sup>	0.19	0.13	0.07	55	2.0	0.61	1.4
Pamunkey	2,774	1.9 x10 <sup>2</sup>	1.3 x10 <sup>2</sup>	6.9 x10 <sup>1</sup>	1.4 x10 <sup>5</sup>	1.9 x10 <sup>3</sup>	6.1 x10 <sup>2</sup>	1.3 x10 <sup>3</sup>	0.25	0.17	0.09	183	2.5	0.81	1.7
Mattaponi	1,557	6.9 x10 <sup>1</sup>	3.8 x10 <sup>1</sup>	3.3 x10 <sup>1</sup>	2.0 x10 <sup>4</sup>	7.7 x10 <sup>2</sup>	1.5 x10 <sup>2</sup>	6.3 x10 <sup>2</sup>	0.16	0.09	0.08	48	1.8	0.36	1.5
Patuxent	901	1.4 x10 <sup>2</sup>	1.1 x10 <sup>2</sup>	4.0 x10 <sup>1</sup>	6.7 x10 <sup>4</sup>	2.0 x10 <sup>3</sup>	2.4 x10 <sup>2</sup>	1.7 x10 <sup>3</sup>	0.58	0.46	0.16	269	7.9	0.98	6.8
Choptank	293	4.0 x10 <sup>1</sup>	2.6 x10 <sup>1</sup>	1.4 x10 <sup>1</sup>	7.1 x10 <sup>3</sup>	6.0 x10 <sup>2</sup>	6.5 x10 <sup>1</sup>	5.4 x10 <sup>2</sup>	0.49	0.32	0.17	89	7.5	0.81	6.7

Notes: TP, total phosphorus; PP, particulate phosphorus; DP, dissolved phosphorus; SS, suspended sediment; TN, total nitrogen; PN, particulate nitrogen; DN, dissolved nitrogen.

<sup>a</sup> Study period varies with species; see Appendix D1 for data coverage of each species.

<sup>b</sup> “NTCBW-SUS” represents the non-Susquehanna area of the NTCBW, *i.e.*, the southern eight tributaries.

addition, for consistent graphical representation, we have scaled all estimates by the respective long-term *medians* of *annual* averages. These medians are summarized in Table 5.2. Readers interested in the estimates of absolute loadings or yields should interpret plot readings in conjunction with these medians.

All derived data from this work, along with the raw river monitoring data, are stored at the publicly accessible Johns Hopkins University Data Archive (Zhang and Ball, 2014).

### **5.3. Seasonal Trends in the NTCBW**

#### **5.3.1. Difficulty in Evaluating Trends with True-condition Estimates**

As an illustration of the difficulty in evaluating long-term trends from true-condition estimates, we reconstructed 34-year trends (1979-2012) of true-condition seasonal TP (Figure 5.2a) and TN (Figure 5.2b) loads from the NTCBW. As expected, the true-condition loads show a high degree of inter-annual variability driven by streamflow, making it very difficult to discern the non-flow-related trends.

The true-condition loads are useful as best approximations of the actual history of watershed inputs, including the important signature of storm events. In this regard, the range of TP load (Figure 5.2a) is more than double that of TN load (Figure 5.2b), reflecting that P is mainly transported with sediments in particulate form (Pionke *et al.*, 2000; Pärn *et al.*, 2012) and sediments are more strongly affected by discharge than dissolved constituents, presumably because of the effects of rainfall and runoff on surface flushing and erosion of soils, stream banks, and stream beds (Brakebill *et al.*, 2010). Additionally, the enormous spike of TP load in Jul-Sep 2011 was likely due to a

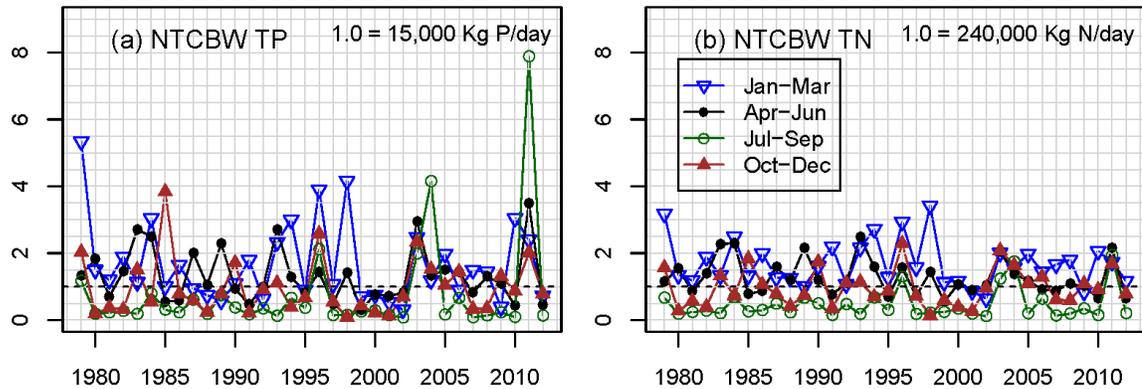


Figure 5.2. True-condition seasonal estimates of (a) total phosphorus (TP) and (b) total nitrogen (TN) loads from the nontidal Chesapeake Bay watershed (NTCBW). Y-axis values are loadings scaled by the respective long-term medians.

combination of both watershed and reservoir effects that were initiated by the consecutive events of Hurricane Irene and Tropical Storm Lee (see section above on “Insights gained from prior studies on Conowingo Reservoir”). In contrast to P, N is mainly transported to rivers in dissolved form via sub-surface flows (Pionke *et al.*, 2000; Pärn *et al.*, 2012), and thus considerably less sensitive to changes in sediment storage or release.

### 5.3.2. Flow-normalized Trends in the NTCBW

The flow-normalized TP load from NTCBW showed a similar “fall-and-then-rise” trend in all seasons (Figure 5.3a). The effect is clearly related to particulate phosphorus – PP showed the same “fall-and-then-rise” trend (Figure 5.3b), whereas DP showed downward trends in all seasons (Figure 5.3c). Both DP and PP contributed to the TP decline until the mid-1990s, but PP alone contributed to the TP rise thereafter.

The flow-normalized SS load from NTCBW also showed a general trend of “fall-and-then-rise” in all seasons (Figure 5.3d). SS load has generally decreased until the 1990s and has since increased rapidly.

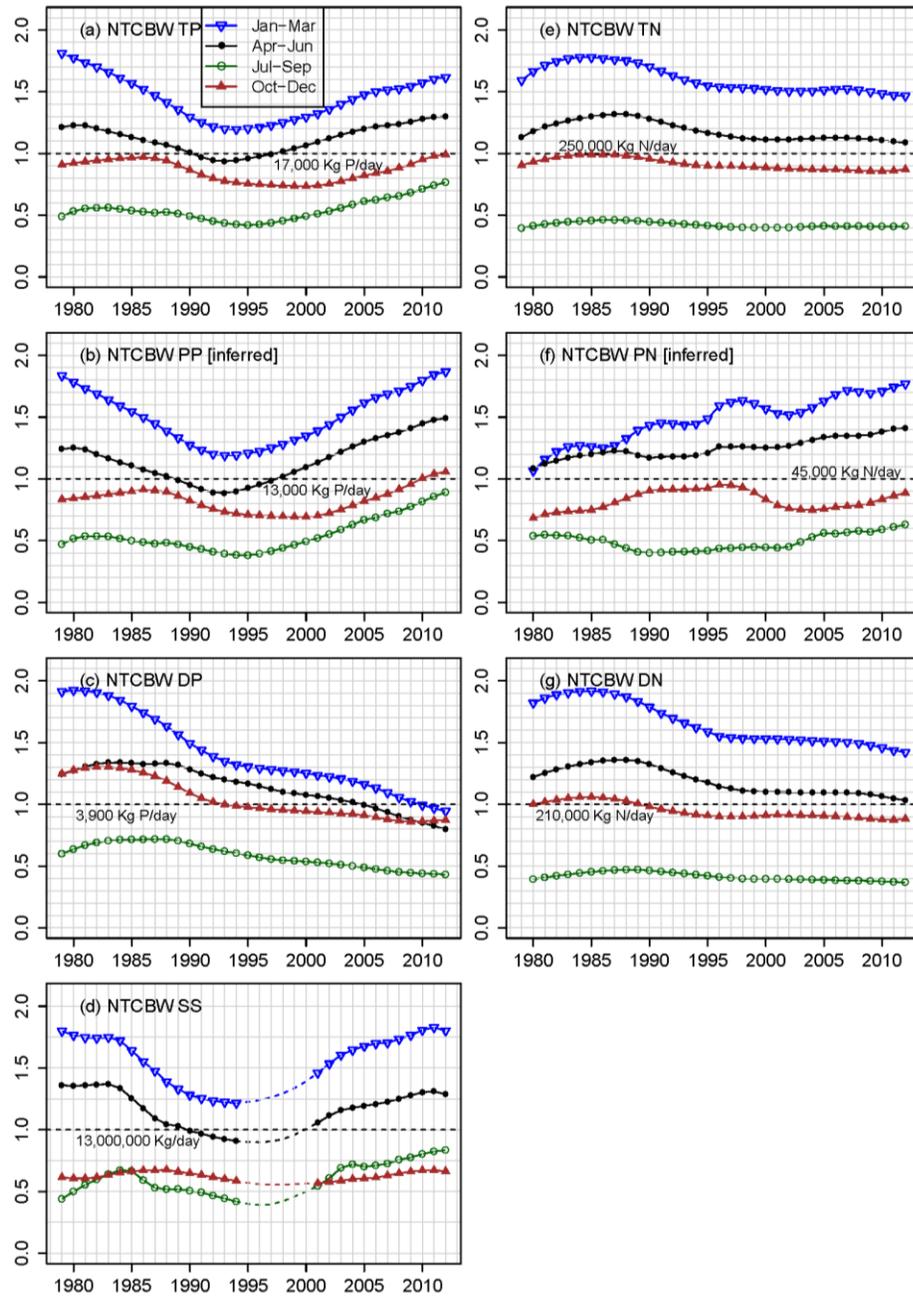


Figure 5.3. Flow-normalized seasonal estimates of (a) total phosphorus (TP), (b) particulate phosphorus (PP), (c) dissolved phosphorus (PP), (d) suspended sediment (SS), (e) total nitrogen (TN), (f) particulate nitrogen (PN), and (g) dissolved nitrogen (DN) loads from the nontidal Chesapeake Bay watershed (NTCBW). Y-axis values are loadings scaled by the respective long-term medians (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

The flow-normalized TN load from NTCBW also showed long-term trends that are similar among all seasons, but generally opposite to those of SS and TP (*i.e.*, “rise-and-then-fall”), with the peak load occurring in the late 1980s (Figure 5.3e). The effect is clearly related to dissolved nitrogen – DN showed the same “rise and then fall” trend across all seasons (Figure 5.3g), whereas PN showed generally rising trends in three of four seasons (S1, S2, and S3) (Figure 5.3f).

In summary, the above analyses highlight the remarkable dissimilarity between trends for dissolved and particulate species. The general decline in dissolved species (*i.e.*, DN and DP) in all seasons suggests a positive effect of management actions that include the banning of P in detergents, WWTP upgrades, and N emission controls on power plants and automotive vehicles, all of which are known to have affected N and P loads from the NTCBW during the last twenty to thirty years. A smaller decline in DN relative to DP was observed, however, and may possibly relate to either less effective N amelioration or longer residence time of DN in the system, with correspondingly longer lags in response (Böhlke and Denver, 1995; Sanford and Pope, 2013). In regard to the lags, apparent ages of groundwater in most parts of the Chesapeake watershed have been reported to be 20 years or more, depending on hydrological condition, geochemistry, and land use (Focazio *et al.*, 1997).

In contrast to the dissolved species, loadings of the particulate species (*i.e.*, SS, PP, and PN) from NTCBW have been generally rising since the mid-1990s, probably owing to a combination of (1) land-related practices (*e.g.*, land clearance, farm fertilization, and urbanization) that enhance land imperviousness and erosion (Gellis *et al.*, 2008; Brakebill *et al.*, 2010), (2) increasing erosion of river bed and bank sediments (Brakebill *et al.*,

2010; Massoudieh *et al.*, 2013), (3) removal of small mill dams (Walter and Merritts, 2008; Merritts *et al.*, 2011), and (4) diminished trapping capacity of major reservoirs (Hirsch, 2012; Zhang *et al.*, 2013), which is known to be important for the fall-line loadings from Susquehanna River. Such rising trends in particulate species are of concern to management for at least two reasons. First, these constituents can reduce light penetration and inhibit growth of submerged aquatic vegetation. Second, although PP and PN are not immediately available for algal consumption, a portion of these constituents can be subsequently decomposed into soluble and bioavailable forms that can sustain algae growth (Kemp and Boynton, 1984; Brady *et al.*, 2013; Testa *et al.*, 2013). Such regeneration from particulate nutrients can be promoted under conditions of high temperature (Kemp and Boynton, 1984) and low oxygen concentration (Boynton *et al.*, 1995), which are characteristics of Chesapeake Bay in summer. Therefore, future research should examine the factors affecting these rising trends.

### **5.3.3. Dominance by Susquehanna River**

NTCBW trends closely follow those of Susquehanna River at the fall-line site (Figure D2). This similarity is not surprising, given that Susquehanna is the largest single contributor in the NTCBW in terms of river flow (62%), TN load (65%), TP load (46%), and SS load (41%), as based on true-condition estimates during the period of 1979-2012. Considering flow-normalized estimates, median annual contributions during this period were roughly similar to those based on true-condition estimates: 65, 43, and 37% for TN, TP, and SS, respectively. The flow-normalized estimate of annual contribution is lower than the true-condition estimate for SS only, which likely reflects nonstationarity of SS retention performance in Conowingo Reservoir. Considering both types of estimates, the

lower fractional contributions of TP and SS relative to TN and flow are attributable to the historical retention of these constituents by Conowingo Reservoir (see section above on “Insights gained from prior studies on Conowingo Reservoir”). In this regard, rising trends of SS, PP, and PN have been recently occurring due to diminished trapping capacity of the reservoir (Hirsch, 2012; Zhang *et al.*, 2013).

Because Susquehanna River dominates trend behavior of the NTCBW, we conducted a separate analysis of the other eight tributaries by summing their loads for various species. The results (Figure 5.4) confirm that the above-noted contrast between dissolved and particulate species has also been generally reflected in the summed loadings of these other eight tributaries.

Further comparison can be obtained by quantifying the 10-year average rates of rise in particulate species (for 2002-2012) for the NTCBW and comparing these with those of Susquehanna River alone. Rates of rise in SS, PP, TP (dominated by PP), and PN for the NTCBW were  $2.43 \times 10^5$ , 460, 393, and 792 kg/day/year, respectively. Corresponding rates of rise in Susquehanna were  $2.23 \times 10^5$ , 255, 269, and 736 kg/day/year, respectively. Thus, during this 10-year period, Susquehanna accounted for 92% of SS rise and 93% of PN rise in the NTCBW, but only 56% of PP and 68% of TP rises. We conclude, therefore, that SS and PN rises in the NTCBW have been driven mostly by Susquehanna. PP rise has been driven by contributions from all tributaries, with Susquehanna contributing to the rise in roughly the same proportion as its contribution to flow. The implication is that the “extra” (reservoir-driven) increases in recent SS from Susquehanna have not caused its proportion of PP loading to increase substantially, relative to the other eight tributaries. This issue can be more directly examined with the PP/SS mass ratio.

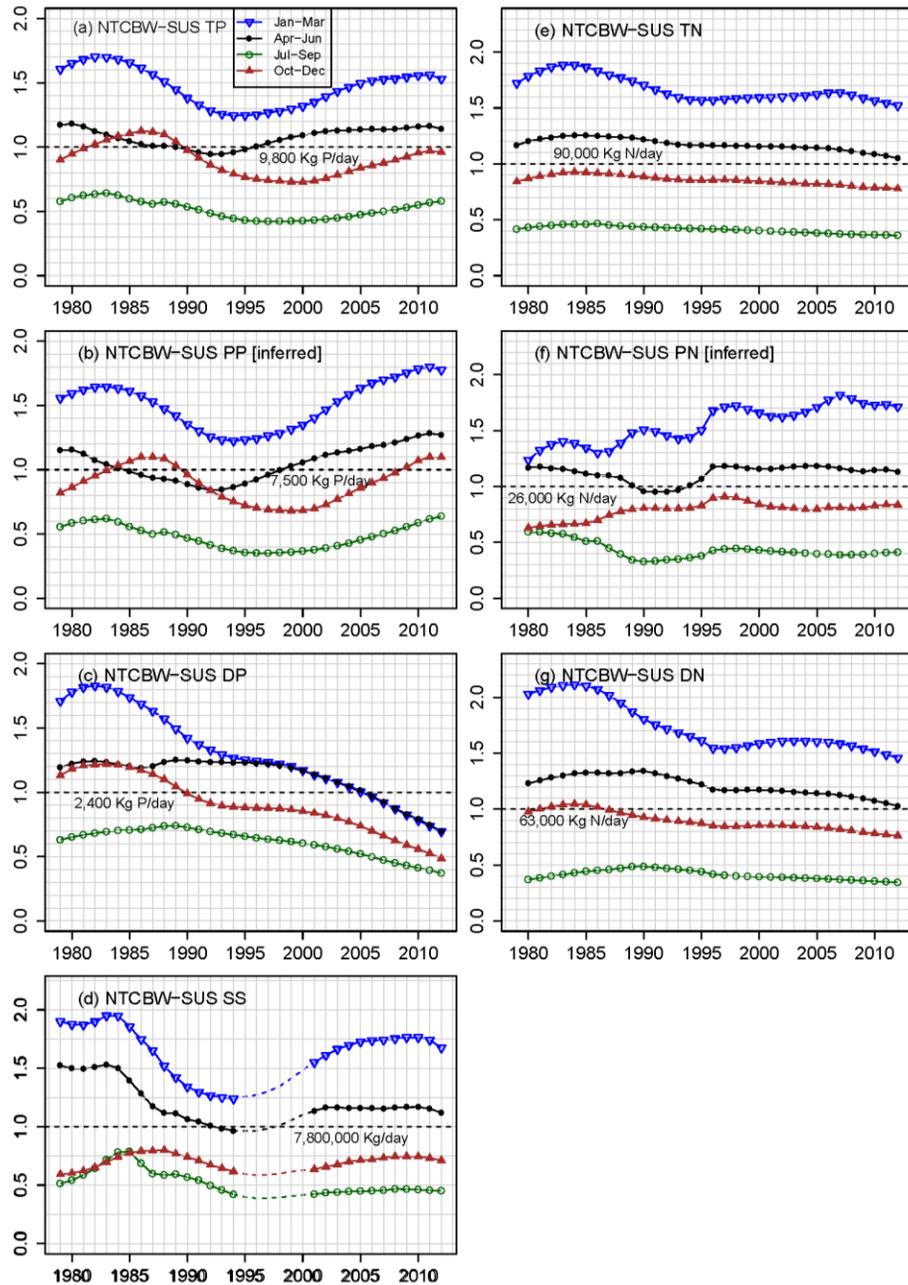


Figure 5.4. Flow-normalized seasonal estimates of (a) total phosphorus (TP), (b) particulate phosphorus (PP), (c) dissolved phosphorus (PP), (d) suspended sediment (SS), (e) total nitrogen (TN), (f) particulate nitrogen (PN), and (g) dissolved nitrogen (DN) loads from all nine rivers but the Susquehanna (*i.e.*, “nontidal Chesapeake Bay watershed – Susquehanna” [“NTCBW-SUS”]). Y-axis values are loadings scaled by the respective long-term medians (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

Interestingly, the ratio has been generally smaller in high-flows than in low-flows for Susquehanna River (data not shown), implying that the sediment discharged during high-flow events is relatively low in PP. In addition, similar effects of streamflow on the PP/SS ratio has also been observed for the non-Susquehanna part of the NTCBW. These results are consistent with other findings reported in the literature. For example, Hainly *et al.* (1995) documented higher P concentrations in deposited sediments in Susquehanna River near the Conowingo Dam than in its upstream channel. During high-flow events, the relatively low-P sediments in the upstream channel are more likely mobilized and then discharged through the Conowingo Dam, resulting in lower overall PP/SS ratios. In addition, Horowitz *et al.* (2012) documented that P concentrations in bed sediments in the RIM rivers are generally lower than those in suspended sediments. During high-flow events, the relatively low-P bed sediments are more likely mobilized and mixed with the high-P suspended sediment, resulting in lower overall PP/SS ratios. In addition, more of the otherwise “settleable” and typically lower-P, large-sized sediments are likely to escape trapping during storm events. We contend that further research is needed, including more sediment monitoring at various flow conditions, in order to more rigorously test these hypotheses and distinguish between the alternatives of enhanced scour or reduced trapping in Conowingo Reservoir.

In addition to the above analysis of particulate species increase in loading during 2002-2012, we have also separately calculated NTCBW and Susquehanna rates of decline in dissolved species for 1985-1995 (the period with most pronounced decline). 10-year average rates of decline in DP, DN, and TN (dominated by DN) for the NTCBW were -107, -3476, and -3114 kg/day/year, respectively. Corresponding average rates of

decline in Susquehanna were -58, -2266, and -1997 kg/day/year, respectively. Thus, during this 10-year period, Susquehanna accounted for 55, 65, and 64% of the decline in DP, DN, and TN in the NTCBW, respectively, all of which were close to its fractional contribution to NTCBW flow discharge.

In summary, most contributions of Susquehanna (PP, DP, DN, TN; 55% to 68%) were in rough agreement with its proportion of flow (62%), while contribution of SS rise during the 2002 to 2012 period was significantly higher (92%). Overall, the lack of progress in particulate species reduction, together with the decline of dissolved species, appears to be a phenomenon that is occurring in many different parts of the Bay watershed. In the two sections below, we examine more closely the individual seasonal loading trends for all nine major tributaries, namely, Susquehanna (Figure D2), Potomac (Figure D3), James (Figure D4), Rappahannock (Figure D5), Appomattox (Figure D6), Pamunkey (Figure D7), Mattaponi (Figure D8), Patuxent (Figure D9), and Choptank (Figure D10).

## **5.4. Comparison of Loads among Seasons**

### **5.4.1. Comparison of Load Magnitudes among Seasons**

The intra-annual variation in load magnitudes varies among tributaries, but most follow a similar trend: winter (S1) > spring (S2) > fall (S4) > summer (S3). For TN in particular, this ranking applies to the NTCBW and all nine tributaries. For all other constituents, it applies to at least half of the rivers. In general, this ranking has been largely controlled by the relative magnitude of streamflow in each season, which usually exhibits the same ranking. A related piece of evidence is the apparent synchronization in

monthly estimates (Figure 5.5). These data show that the variables of interest – river discharge, TN load, TP load, and SS load – are generally highest in March or April and lowest in July or August in all rivers. In addition to flow, an important factor affecting nutrient concentration is the extent of biological removal in the watershed, primarily through plant uptake (*e.g.*, crops, algae) and microbial modification (*e.g.*, denitrification) (Mulholland and Hill, 1997). Although difficult to quantify, these processes are generally more active during warmer months (*e.g.*, Jul-Sep), resulting in lower nutrient concentration.

Although the above ranking relates to the majority of tributaries, some notable exceptions were observed. First, for some river-constituent combinations, fall (S4) was found to have a lower load than summer (S3), thus being the lowest among all seasons. A key example of this is the SS load in Susquehanna (Figure D2d) (and therefore NTCBW). This likely reflects effects of storm events that usually occur in summer (S3). Interestingly, this deviation from the norm was not observed for nutrients (Figure D2), implying that nutrients are on average less affected by storm events than sediment. This is consistent with the previously noted difference in SS and PP with regard to “extra” (reservoir-driven) loadings from Susquehanna (see section above on “Dominance by Susquehanna River”) – *i.e.*, Susquehanna fractional contributions are higher for SS (92%) than PP (68%).

A second exception was in cases where fall (S4) load was similar to or above spring (S2) load. Examples include DN load in Rappahannock (Figure D5g), Appomattox (Figure D6g), and Pamunkey (Figure D7g). In these cases, the concurrent (1995-2012) flow ratios (S2 to S4 average flow, or  $Q_{S2}/Q_{S4}$ ) were respectively 1.17, 1.04, and 1.03, all

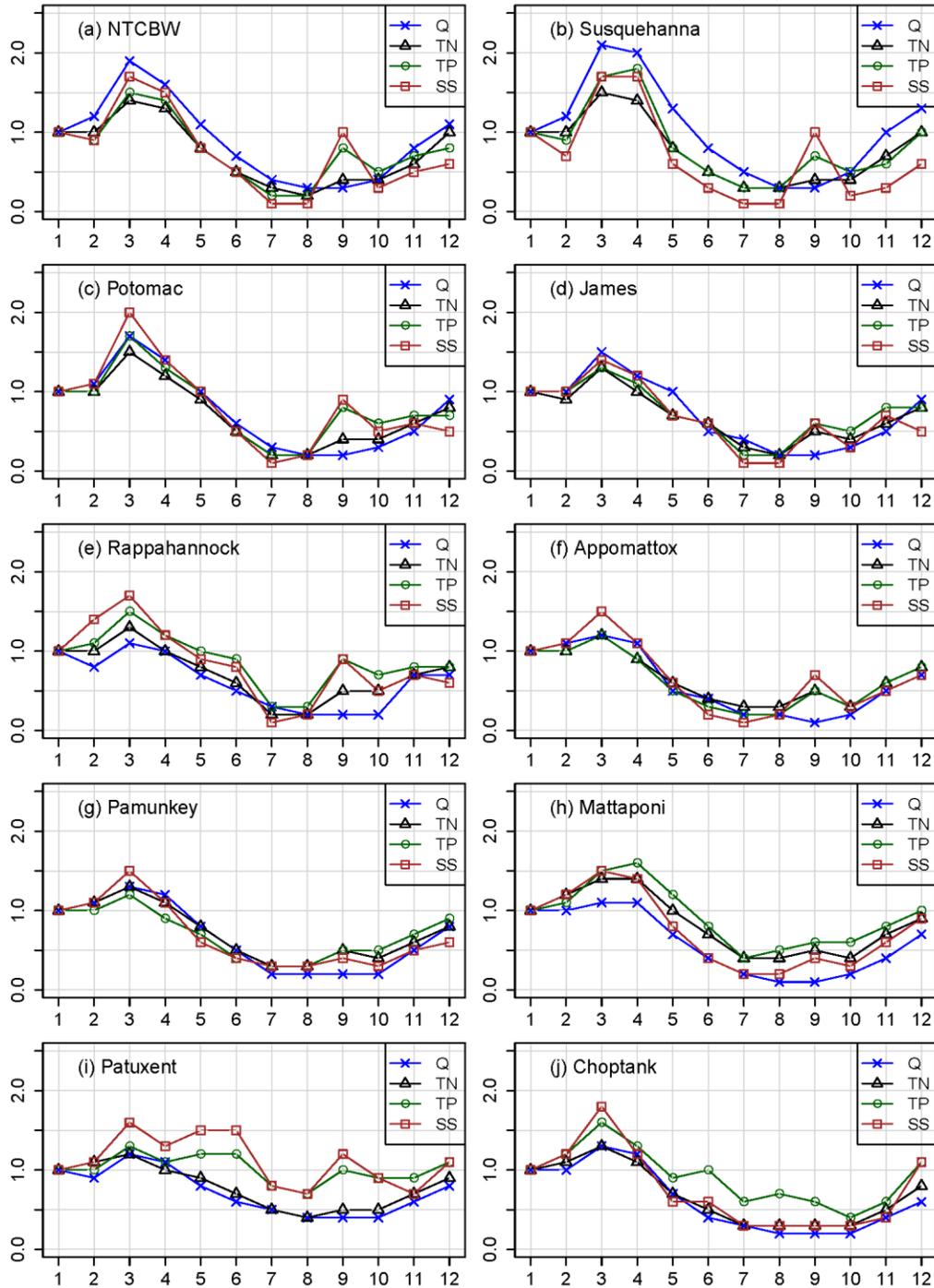


Figure 5.5. Intra-annual variations of river flow, total nitrogen (TN) load, total phosphorus (TP) load, and suspended sediment (SS) load in the nontidal Chesapeake Bay watershed (NTCBW) and the nine individual tributaries, as based on long-term (1970s-2012) median values. X-axis represents calendar months. Y-axis values have been scaled by the respective values in January.

close to one. In addition, three of the rivers (Appomattox, Susquehanna, and Pamunkey) exhibited complex “cross-over” patterns – *i.e.*, with loadings rising in one season while falling in another, and to an extent sufficient to cause a mid-term switch in the relative ranking (Figures D6a-c, D2c, D7b). For Appomattox and Susquehanna, the changes in relative loading magnitudes roughly follow changes in the flow ratio  $Q_{S2}/Q_{S4}$ . Details are complex as described in Table D10 (Appendix D3), but are suggestive of some possibly important intra-annual trends in loadings in these rivers. Among these “cross-overs”; however, neither S2 nor S4 has ever been the season of the year with the highest P load; rather, S1 has always been the season with the highest load. In addition, we have separately investigated whether these “cross-overs” are dependent on seasonal definitions and have found that (1) Appomattox P (TP, PP, and DP) loads have dropped more significantly in March to May than in the other months, and (2) Susquehanna DP load has dropped more significantly in February to June than in the other months (Table D10; Figures D12-D15). Concurrent with these “cross-overs,” the S2 flow in both rivers has become less dominant than S4 flow, resulting in loads that are more similar between the two seasons. These flow trends between S2 and S4 are perhaps related to previously observed trends of increased rain relative to snow and depressed June flow due to greater evapotranspiration and less groundwater flow from spring snowmelt (Hayhoe *et al.*, 2007; Najjar *et al.*, 2010). Moreover, as flow became more similar in the two seasons, some nonflow factors (*e.g.*, biogeochemical processes) may have exerted more effect on the relative ranking of the loading magnitudes. To the extent that P influences algal growth in Appomattox and Susquehanna, these trends in S2 and S4 loadings would tend to favor fewer spring and more fall algal blooms in recent years. The flow ratio  $Q_{S2}/Q_{S4}$

for Pamunkey River has also decreased in the more recent period and in fact became less than one after the cross-over year. For this river however, the switch in PP load ranking is opposite to the other two rivers, with S2 load being higher than S4 load after the cross-over year (Table D10; Figure D16). As with the other intra-annual changes, these changes may or may not be indicative of long-term trends but merit further investigation with respect to causative factors.

Finally, there were some cases where all seasons had similar loads, for example TP, PP, and DP in Patuxent River (Figure D9a-c). The nontidal portion of this watershed is point-source dominated in terms of P, and thus reasonably exhibited little intra-annual variation. In contrast, DN, PN, and TN followed the common ranking noted above, probably because they are more influenced by flow-related nonpoint sources and atmospheric deposition. These latter sources collectively represent 38% of N input, as compared with 19% for P input (Ator *et al.*, 2011).

#### **5.4.2. Comparison of Load Trends among seasons**

Decadal trends in loadings have been generally similar among all seasons, implying that changes in watershed function and land use have had similar impacts on loadings at all times of the year.

One exception to this pattern was with respect to P load in Potomac River. For this river, both TP and PP in fall (S4; Figure D3a-b) showed similar decadal-scale fall-rise trends as in the other seasons, but with an apparent 3- to 5-year lag in timing, along the decadal time axis. Another exception was SS load in Susquehanna, where the SS trend in summer (S3; Figure D2d) was remarkably different, owing to Jul-Sep storm events and reservoir effects previously noted.

Other exceptions in seasonal trends were mainly associated with PN load. However, we emphasize that trends of this species must be interpreted with caution, because PN was estimated as the difference between two much larger numbers (*i.e.*, TN and DN). Nonetheless, a particularly interesting case is Choptank (Figure D10f), where PN load in winter (S1) has increased for decades, even as PN has decreased in all other seasons. This trend, although uncertain, would be consistent with a hypothesis of intensified and/or earlier fertilizer applications in winter.

## **5.5. Comparison of Loads among Tributaries**

In this section, we compare the long-term trends in seasonally averaged nutrient and sediment loads among the nine tributaries. Although the smaller rivers do not exert large impacts on the NTCBW loading, trends observed in these rivers are of course locally important, both within the rivers themselves and within the Bay near their mouths, and with commensurate management implications relative to TMDL development. In terms of loading per unit watershed area, some of the smaller tributaries, *e.g.*, Patuxent and Choptank, show median annual N, P, and SS yields that are commensurate or even higher than those from the larger tributaries (Table 5.2; Figure D11).

### **5.5.1. Particulate (SS, PP, PN) and Particulate-dominated (TP) Species**

Similar to the previously noted trends of NTCBW, most tributaries have shown a long-term “fall-and-then-rise” trend in SS load (Figures 5.6-5.7). Notable rises were observed since the 1990s or 2000s in the four largest rivers (Susquehanna, Potomac, James, and Rappahannock; Figure 5.6) and three small rivers (Appomattox, Pamunkey, and Patuxent; Figure 5.7).

Similarly, most tributaries have shown a long-term “fall-and-then-rise” trend in TP load (Figures 5.6-5.7). Notable rises were observed since the 1990s or 2000s in the four largest rivers (Susquehanna, Potomac, James, and Rappahannock; Figure 5.6) and two small rivers (Pamunkey and Choptank; Figure 5.7). The effect is clearly related to PP, which strongly increased in Susquehanna (Figure D2b), Potomac (Figure D3b), James (Figure D4b), Rappahannock (Figure D5b), Pamunkey (Figure D7b), and Choptank (Figure D10b). The other particulate species, PN has also generally increased in Susquehanna (Figure D2f), Potomac (Figure D3f), James (Figure D4f), Appomattox (Figure D6f), and Pamunkey (Figure D7f).

Factors affecting these trend reversals are largely unclear. As previously noted, the Susquehanna trends have been dominated by reservoir effects. However, for the other rivers, no concrete evidence is available to support a similar argument. Overall, these trend reversals for SS, PP, and the PP-dominated TP are matters of considerable concern for management and deserving of further research. Moreover, a recent nation-wide study has suggested that the P problem might not be limited to the NTCBW: the percentage of stream length in good condition for P has decreased from 52.8% in 2004 to 34.2% in 2008-2009 in the Eastern Highlands, where the majority of NTCBW is located (USEPA, 2013).

Some exceptions from this general pattern were noted. The most important case is for Patuxent River, where both TP (Figure 5.7; Figure D9a) and PP (Figure D9b) loads have declined dramatically in recent decades. This is at least partially related to P-detergent ban and biological nutrient removal in WWTP (Boynton *et al.*, 2008). Another important example is for Choptank River (Figure D10d), where SS has declined in all seasons since

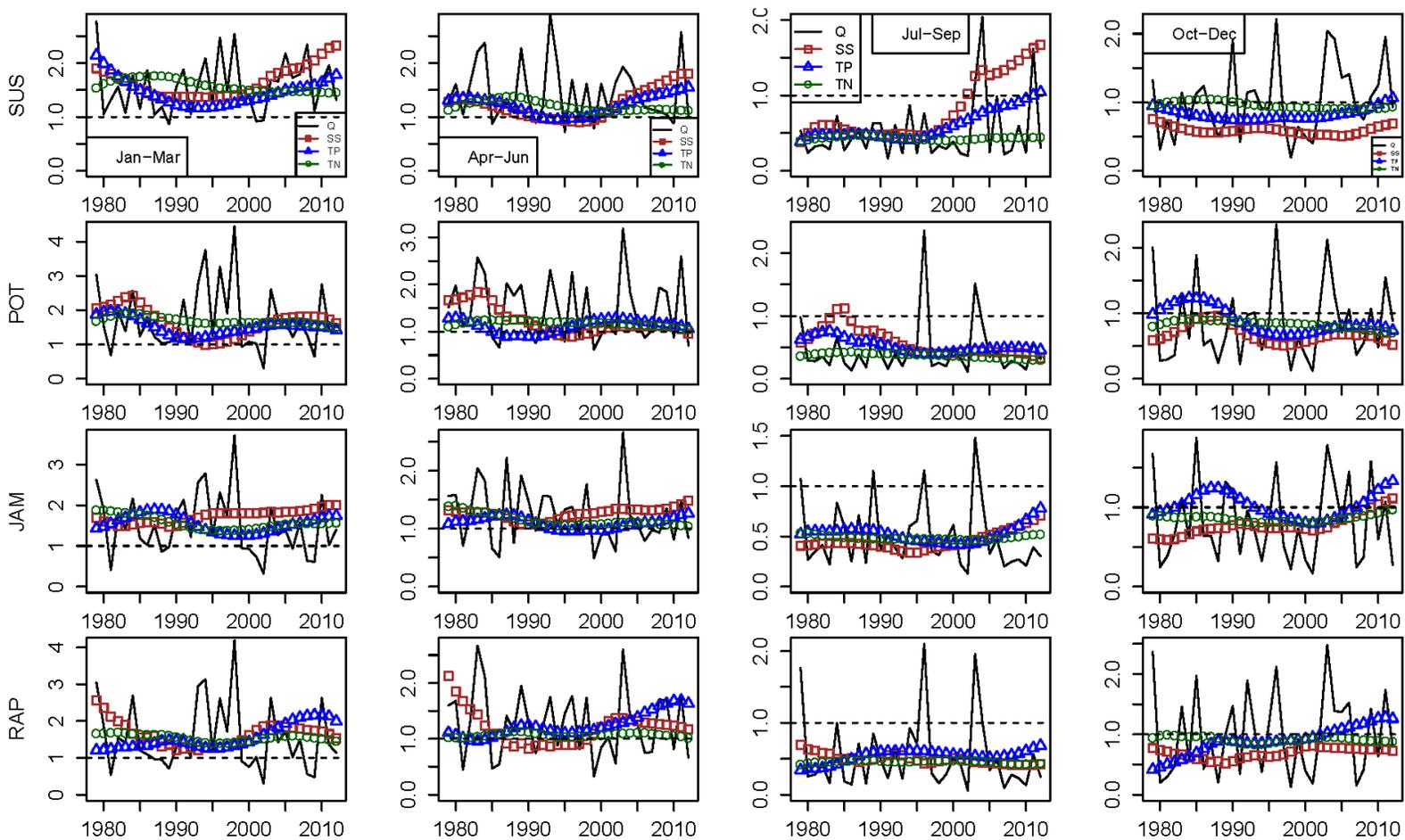


Figure 5.6. Seasonal averages of streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the four largest tributaries (SUS, Susquehanna; POT, Potomac; JAM, James; RAP, Rappahannock). Y-axis values are values scaled by the respective long-term medians (see Table 5.2).

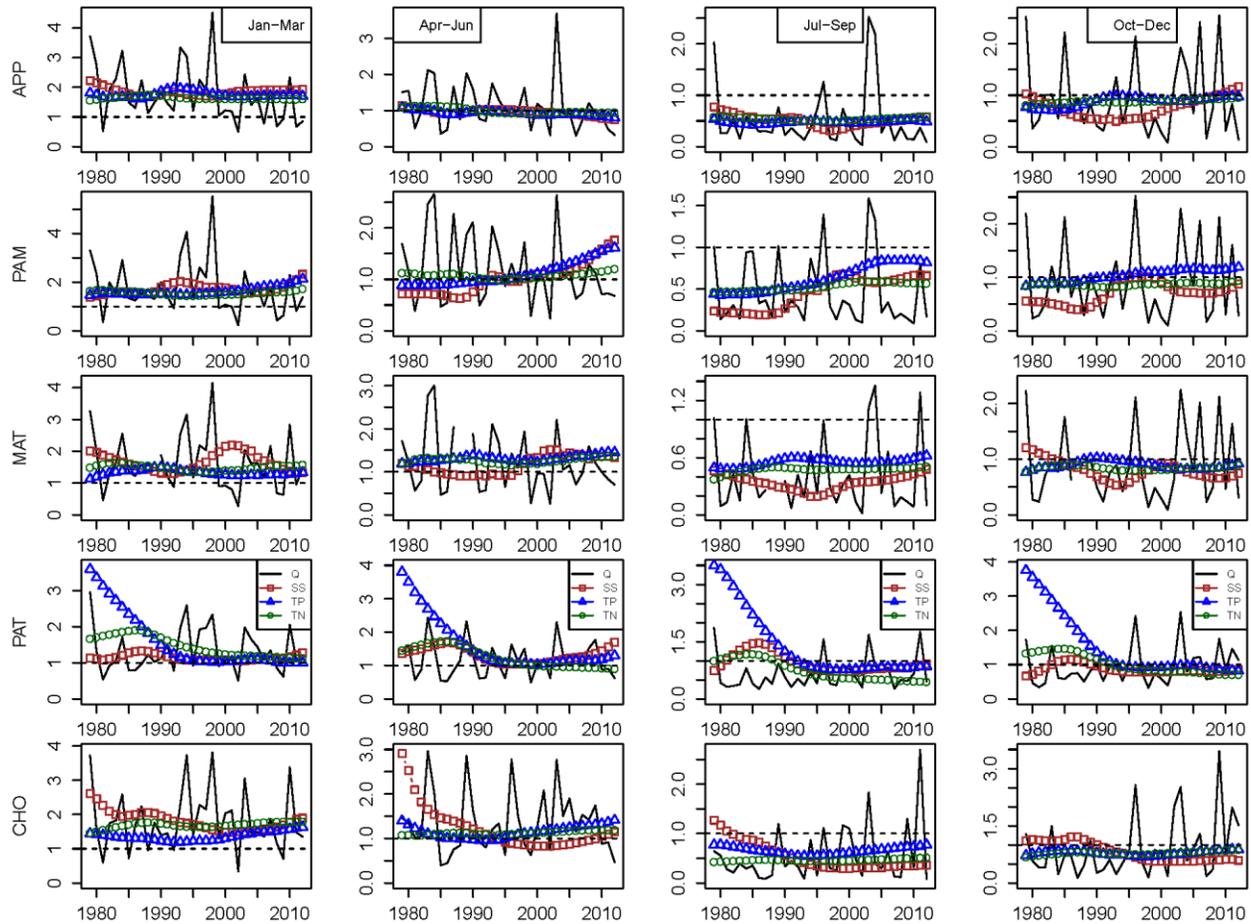


Figure 5.7. Seasonal averages of streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the five smallest tributaries (APP, Appomattox; PAM, Pamunkey; MAT, Mattaponi; PAT, Patuxent; CHO, Choptank). Y-axis values are values scaled by the respective long-term medians (see Table 5.2).

the 1970s. However, it remains unclear how much of this SS reduction can be related to land management practices.

### **5.5.2. Dissolved (DN, DP) and Dissolved-dominated (TN) Species**

Most tributaries have shown moderate long-term “rise-and-then-fall” trends in TN – *i.e.*, trends that are generally opposite to those of SS and TP. Notable decreases were observed in the three largest rivers (Susquehanna, Potomac, and James; Figure 5.6) and one small river (Patuxent; Figure 5.7). The effect is clearly related to DN, which showed strong decreases in Susquehanna (Figure D2g), Potomac (Figure D3g), James (Figure D4g), and Patuxent (Figure D9g). Similarly, DP has shown decreases in these rivers since the 1980s (Figures D2c, D3c, D4c, and D9c). The nontidal Patuxent is particularly remarkable in terms of the degree of reduction, largely owing to historical management controls on point source inputs (specifically the P-detergent ban and WWTP upgrade) (Boynton *et al.*, 2008) and, as recently reported elsewhere in the watershed, decreases in N loading from atmospheric deposition (Eshleman *et al.*, 2013; Linker *et al.*, 2013b). In fact, these efforts would be expected to achieve more apparent loading reduction in Patuxent than in the other tributaries, because Patuxent has been more dominated by point, urban, and atmospheric sources than the other tributaries (collectively representing 70% of N and 80% of P inputs).

In addition, we noted a typical progression in N decline – DN load (and thus the DN-dominated TN) in most tributaries generally showed strong declines in the late 1980s for about one decade, and then stabilized or even began to rise. This may be related to the response times of various management actions previously noted by Phillips and Lindsey (2003) – point-source controls since the 1980s provided immediate (days to months)

improvement in river water quality, whereas nonpoint source reductions or increases would need a much longer time (years to decades) for responses to be observed in the receiving rivers. Moreover, we noted that atmospheric deposition (AD) of TN has declined steadily in these watersheds (Shenk and Linker, 2013). Such trend contrasts between atmospheric input and riverine output imply that the benefits of AD reduction (due to Clean Air Act) may have been counteracted by increases in some other source inputs. In other words, these nine human-disturbed watersheds may have responded to reductions of atmospheric input in more complex ways than the upland forest watersheds analyzed by Eshleman *et al.* (2013).

Despite the general reductions in dissolved species in most tributaries and seasons, some exceptions from this general pattern were observed. Specifically, Choptank and Mattaponi have not shown a reduction and may even be rising in the DN-dominated TN load (Figure 5.7). In addition, DN and DP loads have been rising for decades in Choptank (Figures D10g and D10c) and Mattaponi (Figures D8g and D8c), which are located on opposite sides of Chesapeake Bay but both Coastal-Plain-affected. These cases may reflect lagged subsurface transport of land-applied nutrients. Nitrate concentrations in groundwater have been found typically to be higher than those in surface water (Lindsey *et al.*, 1997; Valiela *et al.*, 1999). Apparent ages of groundwater in the Chesapeake watershed can reach 20 years or more (Focazio *et al.*, 1997). Such long groundwater transit times inevitably delay the response of stream signals to watershed management practices (Lindsey *et al.*, 2003), and may be longer for the Delmarva Peninsula than the Piedmont region of the Bay watershed (Sanford and Pope, 2013). This “lag-time” effect has been noted for Choptank on the Eastern Shore, but less attention has been given to

Mattaponi River on the Western Shore. Given these trends, it is clear that the Coastal Plain areas will need to be carefully examined to better understand total loadings to the Bay, and also in regard to localized loadings in sensitive regions near these rivers.

### **5.5.3. Fractional Contributions of the Tributaries**

Here we shift focus from absolute loadings to fractional contributions (FCs) of each tributary, using the NTCBW load as a common denominator. In comparison to absolute loads, the FCs highlight “relative” progress in load reductions among the nontidal portions of the nine rivers. For brevity, we limit our discussion to several key results in regard to SS, TP, and TN. Results are shown in Figures 5.8-5.9 for the four largest and five smallest tributaries, respectively. Results of DP, PP, DN, and PN are shown in Appendix D4.

Remarkably, all seven upland rivers (*i.e.*, all except Choptank and Mattaponi) showed little temporal variability in their fractional contributions of TN (Figures 5.8-5.9). This stable behavior implies roughly similar long-term spatial patterns among these tributaries in terms of both N input and watershed modulation. In the latter regard, watersheds have been reported to convert variable climatic- and anthropogenic- driven input signals into comparatively persistent output signals at the edge of streams (Gall *et al.*, 2013). In comparison, the Coastal-Plain-affected rivers (Mattaponi and Choptank) showed general increases in their FC of TN in all seasons (Figure 5.9). Similarly, FCs of both DP and DN have also increased in these two rivers (Figure D18). Consistent with the discussion in the previous section, such rises may be related to historically rising accumulations of land-applied nutrients and their lagged subsurface transport. Although each represent 1% or less of river inputs from the NTCBW, the two rivers may be representing a typical

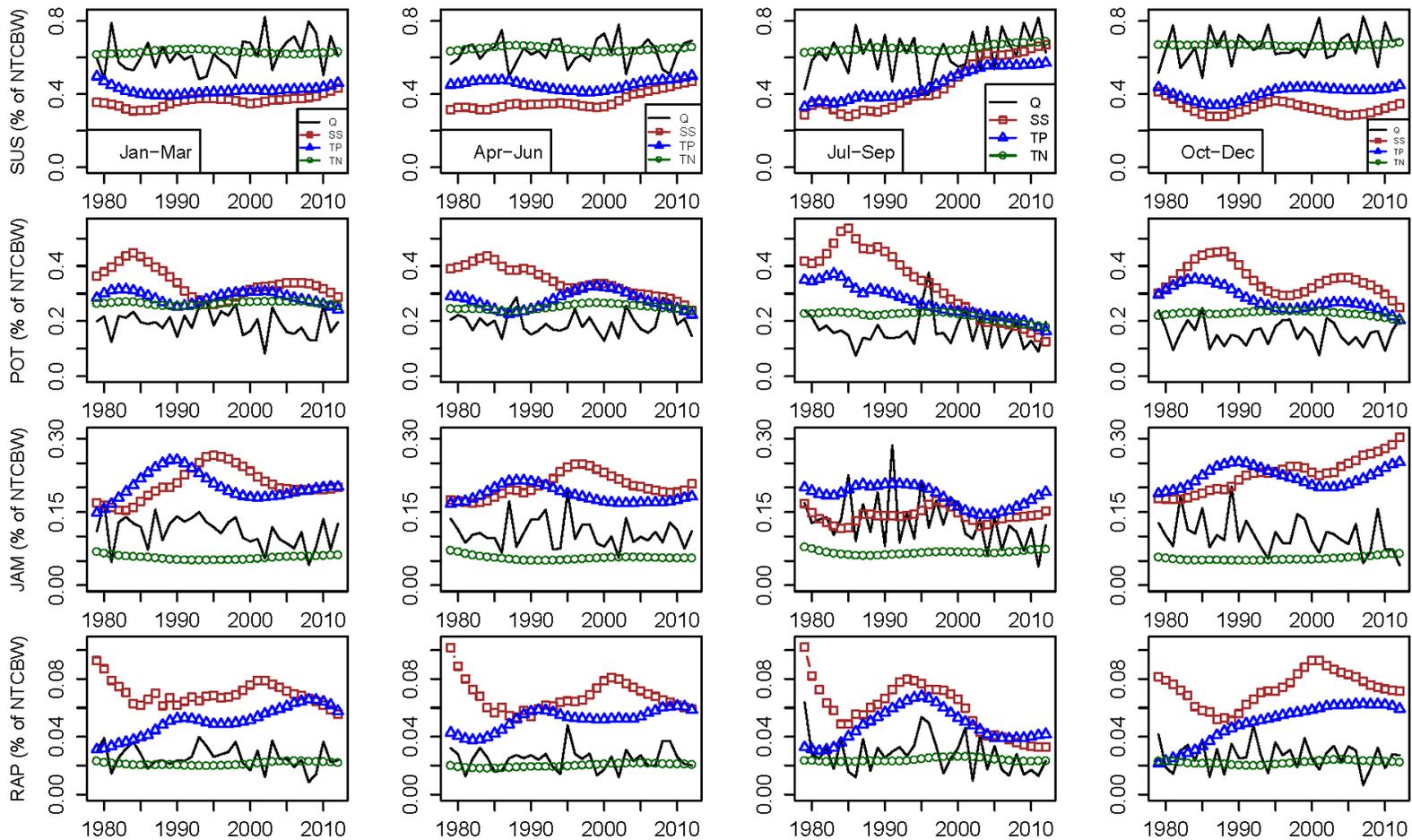


Figure 5.8. Seasonal fractional contributions of the four largest tributaries to the nontidal Chesapeake Bay watershed (NTCBW) for streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) (SUS, Susquehanna; POT, Potomac; JAM, James; RAP, Rappahannock).

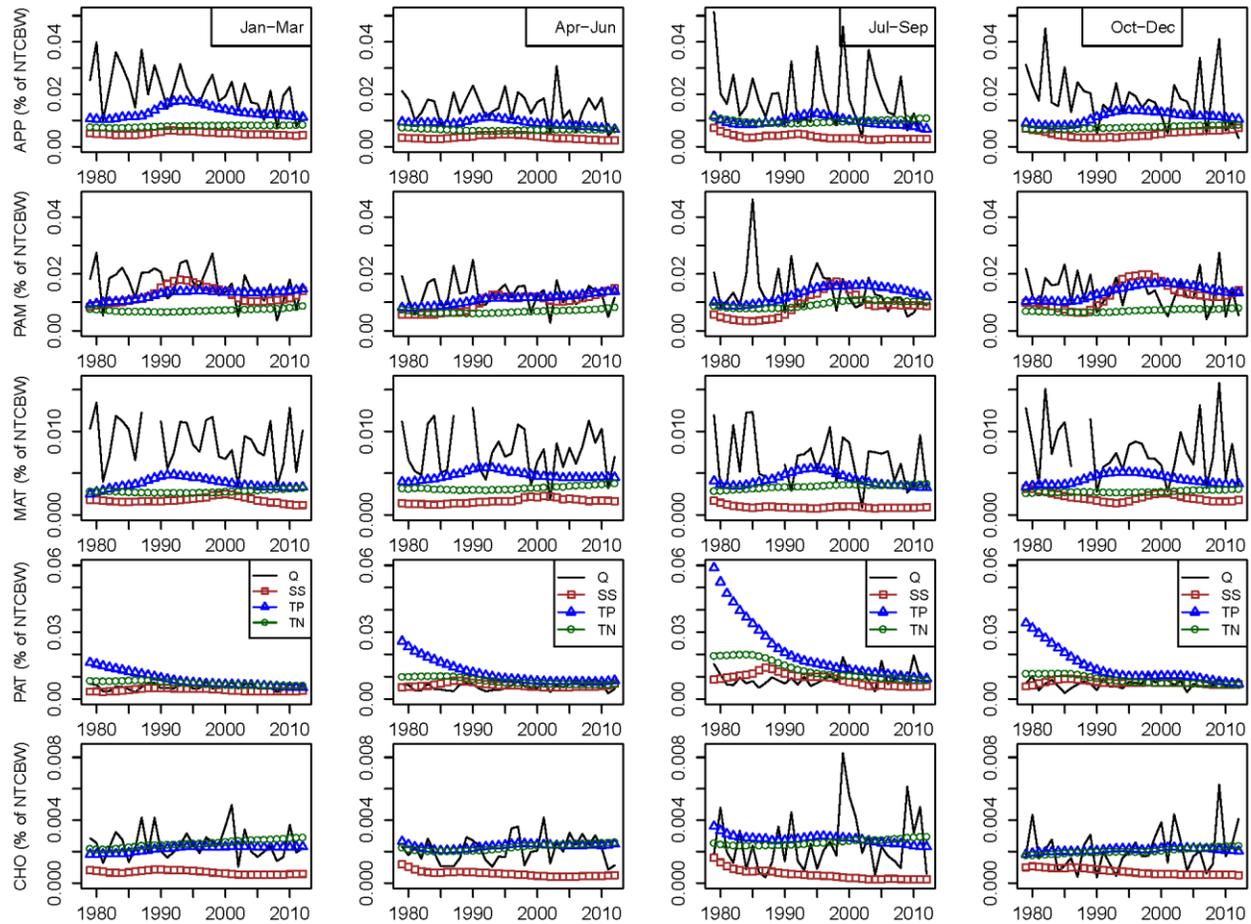


Figure 5.9. Seasonal fractional contributions of the five smallest tributaries to the nontidal Chesapeake Bay watershed (NTCBW) for streamflow and flow-normalized loadings of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) (APP, Appomattox; PAM, Pamunkey; MAT, Mattaponi; PAT, Patuxent; CHO, Choptank).

signature of the larger Coastal Plain areas that feed tidal regions, which have not been specifically reported in this study. In comparison to TN, SS and TP (both particulate-dominated) showed much stronger variability in regard to FC. The relative changes among tributaries in this regard have been strongly affected by changes in SS and TP retention within the Susquehanna reservoirs. Because of such retention, Susquehanna contributions of SS and TP have traditionally been about 30-50%, which are considerably smaller than its FC of streamflow (60-75%), whereas TN contribution has tended to scale more closely with streamflow (Figure 5.8). Recently, as the reservoir has been approaching storage capacity, Susquehanna's contributions of SS and TP have risen, especially in Jul-Sep. Although the past Susquehanna behavior has caused the FCs of SS and TP by the other three large rivers (Potomac, James, and Rappahannock) to be higher than that for flow in the past, the FCs of SS by James and of TP by James and Rappahannock have increased in recent years (Figure 5.8). These rises in FC, despite the increased absolute loadings from Susquehanna, reflect even stronger rises in absolute loadings from these rivers (Figure 5.6). The causes of such rises are an important matter for further investigation.

#### **5.5.4. Changes in N:P Ratios in the Tributaries**

The different observed trends in N and P loads mean that the N:P load ratio is also changing over time. In this regard, it is well known that the ratio of algal-available N and P can have important implications in regard to which nutrient may be the more important limiting factor for algal growth (Testa *et al.*, 2014). Depending on the relative importance of these load ratios for algal growth relative to internal nutrient cycling and the importance of either nutrient in the absolute sense or relative to other limiting factors

(such as silica and light), these ratios may help define the controlling nutrient conditions in the tributaries or their receiving waters (Fisher *et al.*, 1999; Kemp *et al.*, 2005; Conley *et al.*, 2009; Paerl, 2009). For brevity, we limit our discussion here to several key results in regard to TN:TP load ratios. WRTDS-based “flow-normalized” histories of TN:TP ratios are shown in Figures 5.10 (seasonal ratios) and 5.11 (annual ratios)<sup>16</sup> and WRTDS-based “true-condition” results are shown in Appendix D5. (Results of DN:DP and PN:PP ratios are also shown in Appendix D5.) Although the flow-normalized loadings are *model-estimated* (and therefore should be interpreted with caution), the results nonetheless provide possible insights into the systems’ historical and recent tendencies with respect to this important indicator under historically average streamflow conditions.

Overall, seasonal and annual TN:TP ratios have been declining at all nine RIM sites except Appomattox and Mattaponi Rivers, owing to the general trends of reducing TN and increasing TP loads. As a result, seasonal TN:TP ratios, on a molar basis, have been approaching 16:1 – the so-called “Redfield Ratio” (Redfield, 1958). These rivers have been generally transitioning from ratios normally associated with strong P-limitation to those normally associated with co-limitation (by both N and P) or even approaching N-limitation. In these rivers, the Jul-Sep season generally has the lowest TN:TP ratios among the four seasons and appear to generally have reached ratios associated with co-limitation. Following the convention in Qian *et al.* (2000), we show an approximate region of potential co-limitation in Figure 5.10 as the region between the dashed lines representing N:P ratios of 10 and 20. These results may have important implications for

---

<sup>16</sup> Relative to our published paper (Zhang *et al.*, 2015), Figure 5.11 has been added based on post-publication comments as an arguably more relevant measure of long-term trends.

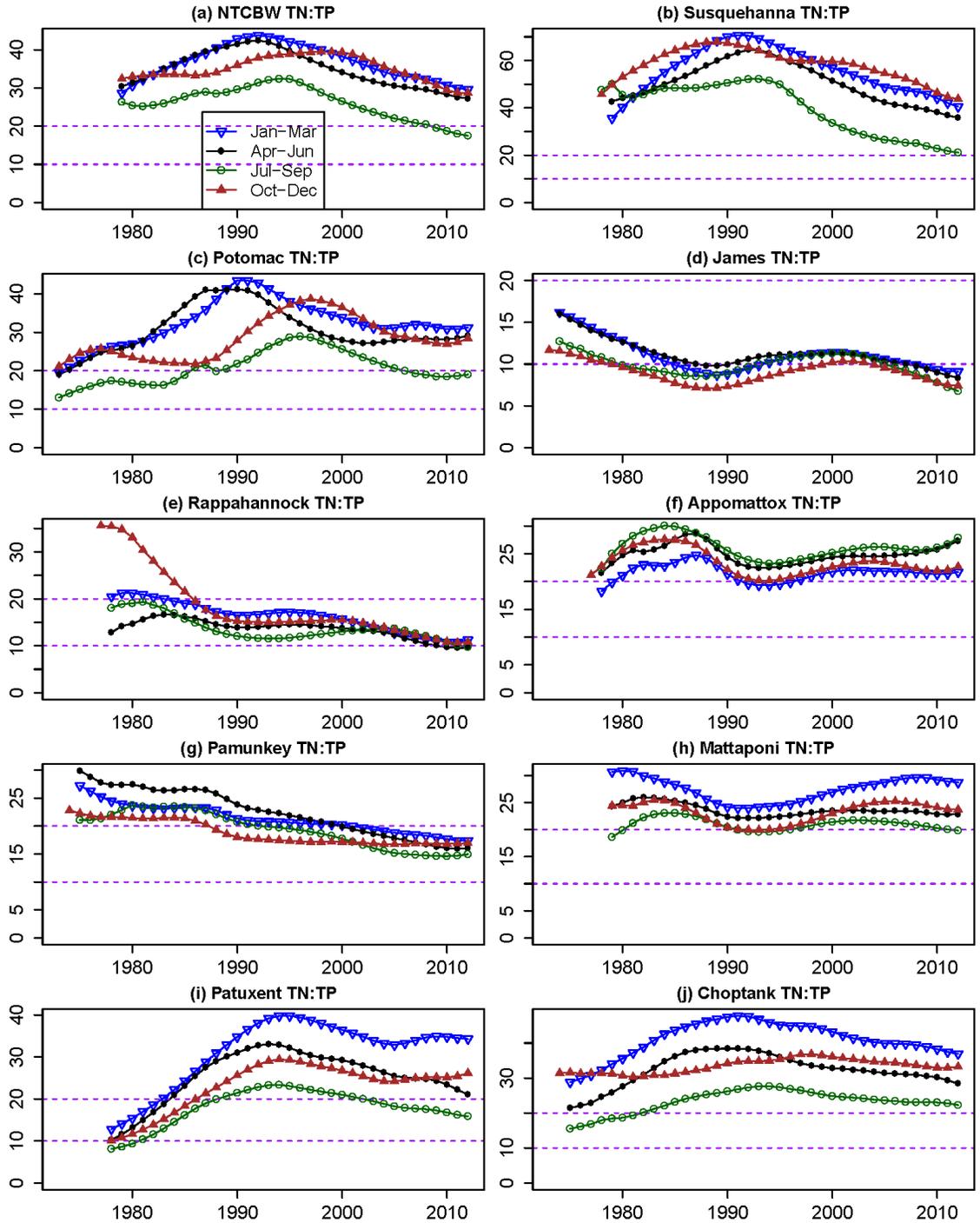


Figure 5.10. Seasonal TN:TP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW) and the nine individual tributaries, as calculated based on WRTDS flow-normalized estimates. On each panel, y-axis indicates the N:P ratios, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

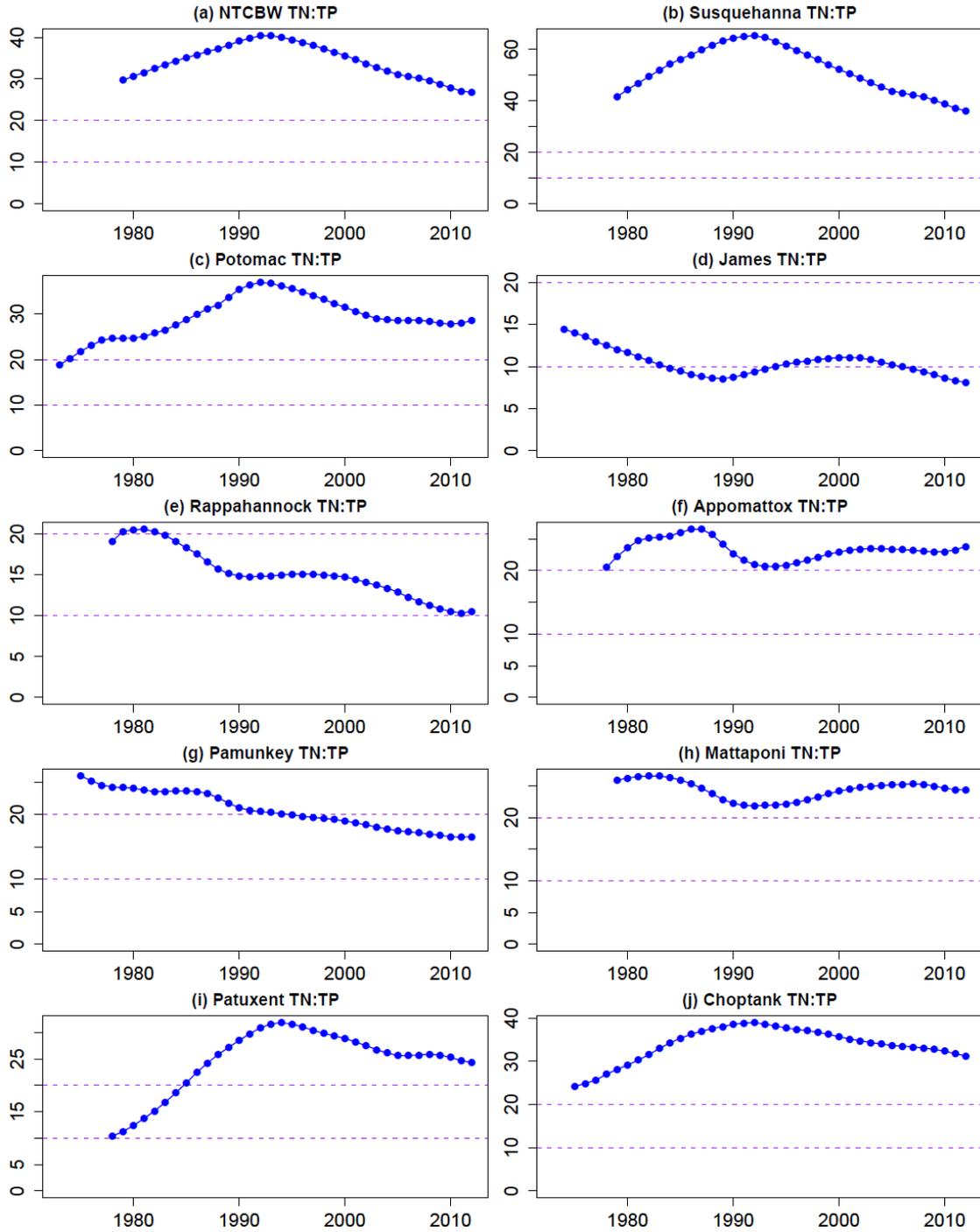


Figure 5.11. Annual TN:TP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW) and the nine individual tributaries, as calculated based on WRTDS flow-normalized estimates. On each panel, y-axis indicates the N:P ratios, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

the surrounding coastal regions because this is a critical period when the Bay undergoes dramatic changes in algal growth, hypoxia extent, and biogeochemical processes (Fisher *et al.*, 1999; Kemp *et al.*, 2005; Conley *et al.*, 2009; Paerl, 2009).

## **5.6. Next Steps: Explaining Changes and Trends**

Our analysis of historical loading changes in the NTCBW complements the prior work of others (Hirsch *et al.*, 2010; Hirsch, 2012; Langland *et al.*, 2012; Moyer *et al.*, 2012; Zhang *et al.*, 2013) and uses a data record that is now more than twice as long as that previously used by Sprague *et al.* (2000), who analyzed loading trends and source inputs for the nine rivers for 1985-1998. Such analyses of loadings are a critical and important basis for future work aimed at better understanding the impacts of numerous potential driving factors, including spatial and temporal changes in land use, fertilizer and manure application rates, atmospheric deposition, and point source loadings. In this regard, a collaboration of federal, state, private, and academic partners from throughout the Chesapeake watershed has begun work within the Chesapeake Bay Partnership to seek explanations for the various loading changes and trends (Keisman *et al.*, 2015).

## **5.7. Summary**

Reductions of nutrient and sediment loadings have been the foci of Chesapeake watershed management. To assess progress, we have analyzed long-term seasonal trends of flow-normalized loads at the fall-line of nine major tributaries from 1970s to 2012. Evaluations of loads by season revealed that nutrient and sediment load magnitudes have been generally highest in Jan-Mar and lowest in Jul-Sep, but the temporal trends have followed similar decadal-scale patterns in all seasons, with notable exceptions. Particulate

nutrients and SS loads have risen since the mid-1990s. The majority of these rises were from Susquehanna River and relate to diminished net trapping at Conowingo Reservoir. Rising SS trends were also observed, however, in other rivers. Moreover, the summed rise in particulate P load from other rivers is of similar magnitude as from Susquehanna. Dissolved nutrient loads have dropped in the upland (Piedmont-and-above) rivers, but risen in two small rivers within the Coastal Plain that are affected by lagged groundwater input. The two rivers may be representative of tidally influenced Coastal Plain regions that are not otherwise accounted by this work. In addition, analysis of fractional contributions revealed that the seven upland rivers yielded surprisingly steady fractions of N, suggesting consistent N input and modulation across the watershed. Finally, TN:TP ratios have declined in most rivers, suggesting the potential for changes in nutrient limitation in the downstream estuaries. Overall, this study demonstrates how integrated comparison of historical data can provide new insights on the impacts of management practice and highlights the value in and importance of maintaining long-term monitoring at multiple watershed locations. Such insights are critically needed to support further investigations about driving factors.

## **5.8. Supporting Information**

Supporting information to this chapter is provided in Appendix D. All derived data from this work, along with the raw river monitoring data, are stored at the publicly accessible Johns Hopkins University Data Archive via <http://dx.doi.org/10.7281/T1KW5CX5> (Zhang and Ball, 2014).

## 5.9. Acknowledgements

This work was funded by the U.S. Water Environment Research Federation (Grant No. U4R09), the National Science Foundation (Grants CBET-0854329 and CBET-1360415), and the Maryland Sea Grant under awards (NA10OAR4170072 and NA14OAR1470090) from the National Oceanic and Atmospheric Administration. We are grateful to Bob Hirsch, Doug Moyer, and Joel Blomquist (USGS) for sharing the latest WRTDS code and data. We thank Peter Wilcock, Dano Wilusz, Gary Shenk, Jeni Keisman, the Associated Editor (Song Qian), and anonymous reviewers for numerous useful comments. This work would not have been possible without the data made available through the USGS River Input Monitoring Program and the National Water Quality Assessment Program.

## 5.10. Literature Cited

Ator, S. W., J. W. Brakebill and J. D. Blomquist, 2011. Sources, fate, and transport of nitrogen and phosphorus in the Chesapeake Bay watershed: An empirical model. U.S. Geological Survey Scientific Investigations Report 2011-5167, Reston, VA, p. 27. <http://pubs.usgs.gov/sir/2011/5167/>.

Böhlke, J. K. and J. M. Denver, 1995. Combined Use of Groundwater Dating, Chemical, and Isotopic Analyses to Resolve the History and Fate of Nitrate Contamination in Two Agricultural Watersheds, Atlantic Coastal Plain, Maryland. *Water Resour. Res.* 31:2319-2339, DOI: 10.1029/95WR01584.

Boicourt, W. C., 1992. Influences of circulation processes on dissolved oxygen in the Chesapeake Bay. *In: Dissolved Oxygen in Chesapeake Bay*, D. Smith, M. Leffler

and G. Mackiernan (Editors). Maryland Sea Grant College, College Park, MD, pp. 7-59.

Boynton, W. R., J. D. Hagy, J. C. Cornwell, W. M. Kemp, S. M. Greene, M. S. Owens, J. E. Baker and R. K. Larsen, 2008. Nutrient Budgets and Management Actions in the Patuxent River Estuary, Maryland. *Estuaries Coasts* 31:623-651, DOI: 10.1007/s12237-008-9052-9.

Boynton, W. R. and W. M. Kemp, 2000. Influence of river flow and nutrient loads on selected ecosystem processes: A synthesis of Chesapeake Bay data. *In: Estuarine Science: A Synthesis Approach to Research and Practice*, J. E. Hobbie (Editors). Island Press, Washington, D.C., pp. 269-298.

Boynton, W. R., W. M. Kemp, J. M. Barnes, L. L. Matteson, F. M. Rohland, L. L. Megdeburger and B. J. Weaver, 1995. Ecosystem Processes Component, Level 1, Interpretive Report No. 12. Chesapeake Bay Laboratory, University of Maryland CBL Ref. No. 95-039, Solomons, MD, p. 101.  
[http://www.gonzo.cbl.umces.edu/documents/water\\_quality/Level1Report12.pdf](http://www.gonzo.cbl.umces.edu/documents/water_quality/Level1Report12.pdf).

Brady, D. C., J. M. Testa, D. M. Di Toro, W. R. Boynton and W. M. Kemp, 2013. Sediment flux modeling: Calibration and application for coastal systems. *Estuar. Coast. Shelf Sci.* 117:107-124, DOI: 10.1016/j.ecss.2012.11.003.

Brakebill, J. W., S. W. Ator and G. E. Schwarz, 2010. Sources of suspended-sediment flux in streams of the Chesapeake Bay Watershed: A regional application of the SPARROW model. *J. Am. Water Resour. Assoc.* 46:757-776, DOI: 10.1111/j.1752-1688.2010.00450.x.

Cohn, T. A., 2005. Estimating contaminant loads in rivers: An application of adjusted

- maximum likelihood to type 1 censored data. *Water Resour. Res.* 41:1-13, DOI: 10.1029/2004wr003833.
- Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.
- Conley, D. J., H. W. Paerl, R. W. Howarth, D. F. Boesch, S. P. Seitzinger, K. E. Havens, C. Lancelot and G. E. Likens, 2009. Controlling Eutrophication: Nitrogen and Phosphorus. *Science* 323:1014-1015, DOI: 10.1126/science.1167755.
- Eshleman, K. N., R. D. Sabo and K. M. Kline, 2013. Surface Water Quality Is Improving due to Declining Atmospheric N Deposition. *Environ. Sci. Technol.* 47:12193-12200, DOI: 10.1021/es4028748.
- Fenneman, N. M. and D. W. Johnson, 1946. Physiographic divisions of the conterminous U. S. U.S. Geological Survey, Reston, VA.  
<http://water.usgs.gov/GIS/metadata/usgswrd/XML/physio.xml>.
- Fisher, T. R., A. B. Gustafson, K. Sellner, R. Lacouture, L. W. Haas, R. L. Wetzel, R. Magnien, D. Everitt, B. Michaels and R. Karrh, 1999. Spatial and temporal variation of resource limitation in Chesapeake Bay. *Mar. Biol.* 133:763-778, DOI: 10.1007/s002270050518.
- Focazio, M. J., L. N. Plummer, J. K. Bohlke, E. Busenberg, L. J. Bachman and D. S. Powars, 1997. Preliminary Estimates of Residence Times and Apparent Ages of Ground Water in the Chesapeake Bay Watershed, and Water-Quality Data From a Survey of Springs. U.S. Geological Survey Water-Resources Investigations Report 97-4225, Richmond, VA, p. 75. <http://pubs.usgs.gov/wri/wri97-4225/>.

- Gall, H. E., J. Park, C. J. Harman, J. W. Jawitz and P. S. C. Rao, 2013. Landscape filtering of hydrologic and biogeochemical responses in managed catchments. *Landscape Ecol.* 28:651-664, DOI: 10.1007/s10980-012-9829-x.
- Gellis, A. C., C. R. Hupp, M. J. Pavich, J. M. Landwehr, W. S. L. Banks, B. E. Hubbard, M. J. Langland, J. C. Ritchie and J. M. Reuter, 2008. Sources, Transport, and Storage of Sediment at Selected Sites in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2008-5186, Reston, VA, p. 95. <http://pubs.usgs.gov/sir/2008/5186/>.
- Hagy, J. D., W. R. Boynton, C. W. Keefe and K. V. Wood, 2004. Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658, DOI: 10.1007/bf02907650.
- Hainly, R. A., L. A. Reed, H. N. J. Flippo and G. J. Barton, 1995. Deposition and simulation of sediment transport in the Lower Susquehanna River reservoir system. U.S. Geological Survey U.S. Geological Survey Water-Resources Investigations Report 95-4122, Lemoyne, PA, p. 39. <http://pubs.er.usgs.gov/publication/wri954122>.
- Hayhoe, K., C. P. Wake, T. G. Huntington, L. F. Luo, M. D. Schwartz, J. Sheffield, E. Wood, B. Anderson, J. Bradbury, A. DeGaetano, T. J. Troy and D. Wolfe, 2007. Past and future changes in climate and hydrological indicators in the US Northeast. *Clim. Dyn.* 28:381-407, DOI: 10.1007/s00382-006-0187-8.
- Helsel, D. R. and R. M. Hirsch, 2002. Statistical Methods in Water Resources. U.S. Geological Survey *Techniques of Water-Resources Investigations Book 4, Chapter A3*. U.S. Geological Survey, Reston, VA, p. 522.

<http://pubs.usgs.gov/twri/twri4a3/>.

Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.

Hirsch, R. M., 2014. Large Biases in Regression-Based Constituent Flux Estimates: Causes and Diagnostic Tools. *J. Am. Water Resour. Assoc.* 50:1401-1424, DOI: 10.1111/jawr.12195.

Hirsch, R. M., R. B. Alexander and R. A. Smith, 1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resour. Res.* 27:803-813, DOI: 10.1029/91WR00259.

Hirsch, R. M. and L. De Cicco, 2014. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data. U.S. Geological Survey Techniques and Methods Book 4, Chapter A10, Reston, VA, p. 94. <http://pubs.usgs.gov/tm/04/a10/>.

Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.

Horowitz, A. J., V. C. Stephens, K. A. Elrick and J. J. Smith, 2012. Concentrations and annual fluxes of sediment-associated chemical constituents from conterminous US coastal rivers using bed sediment data. *Hydrol. Process.* 26:1090-1114, DOI:

10.1002/hyp.8437.

- Johnes, P. J., 2007. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology* 332:241-258, DOI: 10.1016/j.jhydrol.2006.07.006.
- Keisman, J., J. Blomquist, S. Phillips, G. Shenk and E. Yagow, 2015. Estimating Land Management Effects on Water Quality Status and Trends. *Scientific and Technical Advisory Committee Publication Number 14-009*. Edgewater, Maryland, p. 28.
- Kemp, M. W. and W. R. Boynton, 1984. Spatial and temporal coupling of nutrient inputs to estuarine primary production: the role of particulate transport and decomposition. *Bull. Mar. Sci.* 35:522-535.
- Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C. Stevenson, 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* 303:1-29, DOI: 10.3354/meps303001.
- Langland, M. J., 2009. Bathymetry and sediment-storage capacity change in three reservoirs on the Lower Susquehanna River, 1996-2008. U.S. Geological Survey Scientific Investigations Report 2009-5110, Reston, VA, p. 21.  
<http://pubs.usgs.gov/sir/2009/5110/>.
- Langland, M. J., J. D. Blomquist, D. L. Moyer and K. E. Hyer, 2012. Nutrient and suspended-sediment trends, loads, and yields and development of an indicator of

- streamwater quality at nontidal sites in the Chesapeake Bay watershed, 1985-2010. U.S. Geological Survey Scientific Investigations Report 2012-5093, Reston, VA, p. 26. <http://pubs.usgs.gov/sir/2012/5093/pdf/sir2012-5093.pdf>.
- Langland, M. J. and R. A. Hainly, 1997. Changes in bottom-surface elevations in three reservoirs on the lower Susquehanna River, Pennsylvania and Maryland, following the January 1996 flood - implications for nutrient and sediment loads to Chesapeake Bay. U.S. Geological Survey Water-Resources Investigations Report 97-4138, Lemoyne, PA, p. 34. <http://pa.water.usgs.gov/reports/wrir97-4138.pdf>.
- Langland, M. J., D. L. Moyer and J. Blomquist, 2007. Changes in streamflow, concentrations, and loads in selected nontidal basins in the Chesapeake Bay Watershed, 1985-2006. U.S. Geological Survey Open File Report 2007-1372, Reston, VA, p. 68. <http://pubs.usgs.gov/of/2007/1372/>.
- Lindsey, B. D., C. A. Loper and R. A. Hainly, 1997. Nitrate in ground water and stream base flow in the Lower Susquehanna River Basin, Pennsylvania and Maryland. U.S. Geological Survey Water-Resources Investigations Report 97-4146, Lemoyne, PA, p. 66. [http://pa.water.usgs.gov/reports/wrir\\_97-4146.pdf](http://pa.water.usgs.gov/reports/wrir_97-4146.pdf).
- Lindsey, B. D., S. W. Phillips, C. A. Donnelly, G. K. Speiran, L. N. Plummer, J.-k. Böhlke, M. J. Focazio and W. C. Burton, 2003. Residence Times and Nitrate Transport in Ground Water Discharging to Streams in the Chesapeake Bay Watershed. New Cumberland, PA, pp. 201-201. U.S. Geological Survey Water-Resources Investigations Report 03-4035.
- Linker, L. C., R. A. Batiuk, G. W. Shenk and C. F. Cerco, 2013a. Development of the Chesapeake Bay Watershed Total Maximum Daily Load Allocation. *J. Am. Water*

*Resour. Assoc.* 49:986-1006, DOI: 10.1111/jawr.12105.

Linker, L. C., R. Dennis, G. W. Shenk, R. A. Batiuk, J. Grimm and P. Wang, 2013b.

Computing Atmospheric Nutrient Loads to the Chesapeake Bay Watershed and Tidal Waters. *J. Am. Water Resour. Assoc.* 49:1025-1041, DOI: 10.1111/jawr.12112.

Massoudieh, A., A. Gellis, W. S. Banks and M. E. Wieczorek, 2013. Suspended sediment source apportionment in Chesapeake Bay watershed using Bayesian chemical mass balance receptor modeling. *Hydrol. Process.* 27:3363-3374, DOI: 10.1002/hyp.9429.

Medalie, L., R. M. Hirsch and S. A. Archfield, 2012. Use of flow-normalization to evaluate nutrient concentration and flux changes in Lake Champlain tributaries, 1990–2009. *J. Great Lakes Res.* 38:58-67, DOI: 10.1016/j.jglr.2011.10.002.

Merritts, D., R. Walter, M. Rahnis, J. Hartranft, S. Cox, A. Gellis, N. Potter, W. Hilgartner, M. Langland, L. Manion, C. Lippincott, S. Siddiqui, Z. Rehman, C. Scheid, L. Kratz, A. Shilling, M. Jenschke, K. Datin, E. Cranmer, A. Reed, D. Matuszewski, M. Voli, E. Ohlson, A. Neugebauer, A. Ahamed, C. Neal, A. Winter and S. Becker, 2011. Anthropocene streams and base-level controls from historic dams in the unglaciated mid-Atlantic region, USA. *Philosophical Transactions of the Royal Society A* 369:976-1009, DOI: 10.1098/rsta.2010.0335.

Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.

- Mulholland, P. J. and W. R. Hill, 1997. Seasonal patterns in streamwater nutrient and dissolved organic carbon concentrations: Separating catchment flow path and in-stream effects. *Water Resour. Res.* 33:1297-1306, DOI: 10.1029/97wr00490.
- Murphy, R. R., 2014. Development and use of spatial interpolation methods to analyze trends in Chesapeake Bay seasonal hypoxia and stratification, Ph.D. Dissertation, The Johns Hopkins University, Baltimore, MD.
- Murphy, R. R., W. M. Kemp and W. P. Ball, 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* 34:1293-1309, DOI: 10.1007/s12237-011-9413-7.
- Murphy, R. R., E. Perlman, W. P. Ball and F. C. Curriero, 2014. Water-Distance-Based Kriging in Chesapeake Bay. *J. Hydrol. Eng.* 10.1061/(ASCE)HE.1943-5584.0001135, DOI: 10.1061/(ASCE)HE.1943-5584.0001135.
- Najjar, R. G., C. R. Pyke, M. B. Adams, D. Breitburg, C. Hershner, M. Kemp, R. Howarth, M. R. Mulholland, M. Paolisso, D. Secor, K. Sellner, D. Wardrop and R. Wood, 2010. Potential climate-change impacts on the Chesapeake Bay. *Estuar. Coast. Shelf Sci.* 86:1-20, DOI: 10.1016/j.ecss.2009.09.026.
- Pärn, J., G. Pinay and Ü. Mander, 2012. Indicators of nutrients transport from agricultural catchments under temperate climate: A review. *Ecol. Indicators* 22:4-15, DOI: 10.1016/j.ecolind.2011.10.002.
- Paerl, H. W., 2009. Controlling Eutrophication along the Freshwater–Marine Continuum: Dual Nutrient (N and P) Reductions are Essential. *Estuaries Coasts* 32:593-601, DOI: 10.1007/s12237-009-9158-8.
- Phillips, S. W. and B. D. Lindsey, 2003. The Influence of Ground Water on Nitrogen

- Delivery to the Chesapeake Bay. pp. 6-6. <http://pubs.usgs.gov/fs/2003/fs091-03/>.
- Pionke, H. B., W. J. Gburek and A. N. Sharpley, 2000. Critical source area controls on water quality in an agricultural watershed located in the Chesapeake Basin. *Ecol. Eng.* 14:325-335, DOI: 10.1016/S0925-8574(99)00059-2.
- Pritchard, D. W. and J. R. Schubel, 2001. Human influences on physical characteristics of the Chesapeake Bay. *In: Discovering the Chesapeake: the history of an ecosystem*, P. D. Curtin, G. S. Brush and G. W. Fisher (Editors). The Johns Hopkins University Press, Baltimore, MD, pp. 60-82.
- Qian, S. S., M. E. Borsuk and C. A. Stow, 2000. Seasonal and Long-Term Nutrient Trend Decomposition along a Spatial Gradient in the Neuse River Watershed. *Environ. Sci. Technol.* 34:4474-4482, DOI: 10.1021/es000989p.
- R Development Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3900051070. <http://www.r-project.org>.
- Redfield, A. C., 1958. The biological control of chemical factors in the environment. *Am. Sci.* 46:205-221, <http://www.jstor.org/stable/27827150>.
- Sanford, W. E. and J. P. Pope, 2013. Quantifying Groundwater's Role in Delaying Improvements to Chesapeake Bay Water Quality. *Environ. Sci. Technol.* 47:13330-13338, DOI: 10.1021/es401334k.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- Sprague, L. A., R. M. Hirsch and B. T. Aulenbach, 2011. Nitrate in the Mississippi River

and its tributaries, 1980 to 2008: are we making progress? *Environ. Sci. Technol.* 45:7209-7216, DOI: 10.1021/es201221s.

Sprague, L. A., M. J. Langland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. U.S. Geological Survey Water-Resources Investigations Report 00-4218, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).

Testa, J. M., D. C. Brady, D. M. Di Toro, W. R. Boynton, J. C. Cornwell and W. M. Kemp, 2013. Sediment flux modeling: Simulating nitrogen, phosphorus, and silica cycles. *Estuar. Coast. Shelf Sci.* 131:245-263, DOI: 10.1016/j.ecss.2013.06.014.

Testa, J. M., Y. Li, Y. J. Lee, M. Li, D. C. Brady, D. M. Di Toro, W. M. Kemp and J. J. Fitzpatrick, 2014. Quantifying the effects of nutrient loading on dissolved O<sub>2</sub> cycling and hypoxia in Chesapeake Bay using a coupled hydrodynamic-biogeochemical model. *J. Mar. Syst.* 139:139-158, DOI: 10.1016/j.jmarsys.2014.05.018.

U.S. Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment. Annapolis, MD.  
<http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html>.

U.S. Geological Survey, 2013. Surface-water data for the nation. Accessed 15 April 2013, <http://waterdata.usgs.gov/nwis/sw>.

USEPA, 2013. National Rivers and Streams Assessment 2008-2009: A Collaborative Survey (draft). EPA/841/D-13/001, Washington, D.C., p. 109.

[http://water.epa.gov/type/rs1/monitoring/riverssurvey/upload/NRSA0809\\_Report\\_Final\\_508Compliant\\_130228.pdf](http://water.epa.gov/type/rs1/monitoring/riverssurvey/upload/NRSA0809_Report_Final_508Compliant_130228.pdf).

USGS, 2012. Chesapeake Bay river input monitoring program. Accessed 1 March 2012, <http://va.water.usgs.gov/chesbay/RIMP/>.

Valiela, I., J. Costa, K. Foreman, J. M. Teal, B. Howes and D. Aubrey, 1999. Transport of groundwater-borne nutrients from watersheds and their effects on coastal waters. *Biogeochemistry* 10:177-197, DOI: 10.1007/bf00003143.

Walter, R. C. and D. J. Merritts, 2008. Natural streams and the legacy of water-powered mills. *Science* 319:299-304, DOI: 10.1126/science.1151716.

Wang, Y.-G., S. S. J. Wang and J. Dunlop, 2015. Statistical modelling and power analysis for detecting trends in total suspended sediment loads. *Journal of Hydrology* 520:439-447, DOI: 10.1016/j.jhydrol.2014.10.062.

Zhang, Q. and W. P. Ball, 2014. Data associated with Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. Johns Hopkins University Data Archive, Baltimore, MD. <http://dx.doi.org/10.7281/T1KW5CX5>, DOI: 10.7281/T1KW5CX5.

Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI: 10.1016/j.scitotenv.2013.02.012.

*Page intentionally left blank*

## **Chapter 6. What Can We Learn from Limited Data? Statistical Inferences and Uncertainties of Riverine Fluxes and Trends with Limited Sampling of Extreme-Flow Events<sup>17</sup>**

### **Abstract**

Sediment and particulate nutrient delivery from watersheds is often dominated by high-discharge events and some extreme discharges can deliver disproportionately large fractions of total flux. Unfortunately, river water-quality data are typically sparse in the high-discharge range, making flux prediction difficult and uncertain. In Susquehanna River, the largest tributary to Chesapeake Bay, sediment and particulate nutrient fluxes have increased dramatically in recent decades due to decreased trapping efficiency behind Conowingo Dam, as concluded based on previous statistical modeling of riverine monitoring data using the WRTDS (Weighted Regressions on Time, Discharge, and Season) method. Here we seek to systematically examine if these conclusions have been significantly affected by the limited number of concentration observations at extremely high discharges that have occurred during different times within hydrographs at different locations. Our analyses have provided further concrete evidence to support the prior conclusions and alleviate the above concern. The results confirm that WRTDS is a reasonably robust tool for inferring general water-quality conditions and trends from long-term monitoring data, but that uncertainties can sometimes be high during specific

---

<sup>17</sup> This chapter will be submitted for publication in a peer-reviewed journal. Bill Ball was involved in hypothesis development, study design, results interpretation, and editing. Ken Staver was involved in results interpretation. All figures, tables, and data were created by Qian Zhang.

events and need to be recognized. More generally, this work serves as an important case study that demonstrates the value of rigorous statistical approaches for analysis of sparse water-quality data while also emphasizing the uncertainties associated with modeled results, particularly during extreme-flow events.

## 6.1. Introduction

At many river monitoring sites around the world, water-quality data are usually collected at most 2-3 times per month and even less frequently for high discharges due to technical and safety issues. Therefore, statistical modeling is often needed to estimate constituent concentration and flux for each day in the record. Such daily estimates and their aggregates (*e.g.*, seasonal or annual averages) play an important role toward assessment against regulatory standards (*e.g.*, maximum concentration levels), establishment of restoration targets (*e.g.*, total maximum daily loads or TMDL), and calibration of watershed models (Hirsch *et al.*, 2010; Linker *et al.*, 2013; Shenk and Linker, 2013; Zhang *et al.*, 2015).

It has long been understood that uncertainties in the above type of statistical analyses are highest for times at the edges of the data record and for discharges and seasons where monitoring data are most limited (Cohn *et al.*, 1989; Harmel and King, 2005; Johnes, 2007; Hirsch *et al.*, 2010; Moyer *et al.*, 2012; Chanut *et al.*, 2016). In particular, monitoring data are typically sparse in the high-discharge range, making flux prediction difficult in this range, which often dominates sediment and particulate nutrient delivery from watersheds. In this work, we have revisited this general issue of statistical inferences and uncertainties of riverine fluxes and trends with limited sampling of extreme-flow events. Our focus is on the Chesapeake Bay watershed and our analyses are

directly useful toward Chesapeake Bay management. More broadly, our approaches and findings can also be relevant to analyses of riverine water-quality data elsewhere.

For the Chesapeake Bay watershed, TMDL limits have been established to regulate nutrient and sediment loadings from the watershed (U.S. Environmental Protection Agency, 2010). Susquehanna River is of particular importance as the largest tributary (Hagy *et al.*, 2004; Zhang *et al.*, 2015). For this river, several recent studies (Hirsch, 2012; Zhang *et al.*, 2013; Zhang *et al.*, 2016) have documented substantial increases in suspended sediment (SS) and total phosphorus (TP) loadings at Conowingo Dam in recent decades through analyses of decadal-scale monitoring data using the WRTDS (Weighted Regressions on Time, Discharge, and Season) method (Hirsch *et al.*, 2010). Two of these studies (Zhang *et al.*, 2013; Zhang *et al.*, 2016) have also analyzed data at locations upstream of Conowingo Reservoir and attributed the trends at Conowingo Dam to declined reservoir trapping performance. Specifically, the authors documented major increases in reservoir output loadings of SS and TP since around the mid-1990s, despite general declines in input loadings since the 1980s. These decadal-scale changes present critical challenges to the achievement of the TMDL and have become an important concern to Bay managers.

The above statistical analyses were conducted based on water-quality samples collected on ~37 days/year at the reservoir inlet and ~27 days/year at the reservoir outlet. All authors of the previous publications are aware of the above-listed issues and agree that statistical predictions for the most extreme flows are highly uncertain. For these flows (*e.g.*, 2004 Hurricane Ivan and 2011 Tropical Storm [TS] Lee), the timing and quality of sampling upstream and downstream of the reservoir system was simply too

sparse to allow definitive conclusions about trapping efficiency at and beyond those discharges (Robert Hirsch, personal communication, 2016). Because these extreme events typically occur around September, it was speculated that September results may be more heavily affected than other seasons and that a focus on September projections could cause some readers to draw potentially erroneous conclusions regarding the magnitude of the issue. For example, Hirsch (2012) documented 55% and 97% increase in flow-normalized loads of TP and SS, respectively, at Conowingo between 1996 and 2011. While the uncertainties were not well quantified, these estimates have been used as “truth” by recent studies toward examination of Conowingo infill effects on Bay hypoxia (The Lower Susquehanna River Watershed Assessment Team, 2015; Linker *et al.*, 2016).

Another important concern to Bay managers is nutrient and sediment transport during major storm events. In this regard, a discharge level of 11,300 m<sup>3</sup>/s (400,000 cfs) has been commonly used to represent the sediment scour threshold at Conowingo Dam (Gross *et al.*, 1978; Lang, 1982). For the period of 1978-2014, discharges above 11,300 m<sup>3</sup>/s only represent 0.13% of the entire record (18 out of 13,514 days), and hence their effects on estimated fluxes or trends are expected to be largely constrained. Nonetheless, these above-scour flows are of great interest to managers, because they can deliver disproportionately large amounts of river discharges and loadings. Specifically, during the 30-year period of 1985-2014, daily discharges above 11,300 m<sup>3</sup>/s have occurred in six months, namely, April 1993, January 1996, September 2004, June 2006, March 2011, and September 2011 (see Tables E1-E2). Although these six months represent only ~2% of days in this 30-year period, WRTDS estimates suggest that they account for 7% of

flow, 10% of total nitrogen (TN), 26% of TP, and 55% of SS from Conowingo Dam to the Bay in the same period (Zhang *et al.*, 2016).

Quantitative evaluation of the loading and trend uncertainties was made difficult by the complexity of estimation methods used, but new tools have recently been developed (Hirsch *et al.*, 2015). This work has been conducted to address an overall concern to Bay modelers and managers as to whether the limited number of concentration observations at extremely high discharges (which have occurred during different seasons and times within hydrographs) may have significantly affected WRTDS estimates at all discharge conditions (*i.e.*, high, middle, and low discharges), especially when samples are very limited in terms of total number and are not evenly distributed temporally and spatially over the record. Specifically,

- (1) The flows during Hurricane Ivan (2004) and TS Lee (2011) are so large and the concentrations so extreme that these samples might have a major effect on estimated loads in September. If so, the previously documented rise in annual load might be an artifact of the uneven occurrence and sampling of September high-flow events in the record, rather than temporal changes in reservoir performance.
- (2) The higher (but still limited) number of major storm samples in the recent record might have also biased the estimates at moderately high flows ( $< 11,300 \text{ m}^3/\text{s}$ ).
- (3) There are major differences in SS concentration that can occur for the same level of discharge between the rising and falling limbs of storm events (*i.e.*, hysteresis). A related issue is whether the upstream and downstream sites were similarly sampled during major storms in regard to sampling timing and sample count.

- (4) The above issues could affect loading estimates at Conowingo Dam in ways that are not mirrored at the upstream site (*i.e.*, Marietta), or vice versa, thereby leading to biased trend estimates on a long-term basis and/or biased mass-balance estimates (*i.e.*, net deposition or scour) during individual storms.
- (5) If the above issues are not conveyed properly, the uncertain estimates may be misunderstood as “truth” by some readers. This is of particular concern in the case of Conowingo Reservoir, where major management issues are under discussion.

## **6.2. Analyses & Results**

To more thoroughly investigate the uncertainty of the WRTDS estimates, new modeling experiments were conducted. Overall, our analyses (presented below) further confirm the prior general conclusions – *i.e.*, those regarding the temporal trends of increasing output and decreasing net deposition through the reservoir system, and also corroborate that uneven occurrence and sampling of September high-flow events have not had a major impact on these findings. At the same time, we are in a better position to elucidate the high uncertainty associated with projections during specific events or years.

### **6.2.1. WRTDS Mass-Balance Analysis for the Major Storm Events**

First, the differences in storm sampling between upstream (*i.e.*, Marietta) and downstream (*i.e.*, Conowingo) sites and their effects on mass-balance results were examined. Our focus was on the six major storm events in 1987-2014, during which daily discharge has exceeded the scour threshold (*i.e.*, 11,300 m<sup>3</sup>/s). Our results show that Marietta has a worse sampling coverage of these events than Conowingo (Table E2) – Marietta was not sampled on four dates (*i.e.*, 1996/1/20, 1996/1/21, 2006/06/29, and

2011/9/8), whereas Conowingo was not sampled on only one date (*i.e.*, 2006/6/29).

Based on this comparison, loads at Marietta and Conowingo as well as reservoir net deposition (or scour) were further quantified for three sampling scenarios:

- (1) estimation using the original concentration record;
- (2) estimation using “flow-censored”<sup>18</sup> (abridged) samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events. (Note: in our context, “censored samples” should not be confused with “samples with values less than detection limit”.)
- (3) estimation using an “artificial” record that combines the original record and artificial samples for the unsampled dates listed above – see details in Table E1.

Full comparison of daily loadings under these scenarios is presented in the Supporting Information (SS: Figure E1, Table E3; TP: Figure E2, Table E4; TN: Figure E3, Table E5). Full comparison of mass balances is presented in Table E2.

For brevity, two representative events, *i.e.*, Hurricane Ivan (2004) and TS Lee (2011), are selected and shown in Figure 6.1. For the peak-discharge day of both events, scenario #3 predicts a much higher SS concentration at Marietta than scenario #1, but the two scenarios have similar concentrations at Conowingo. Thus, scenario #3 is expected to predict a larger net deposition (or smaller net scour) during these storm events than scenario #1. In this regard, we have updated our previously published curve of sediment storage capacity in Conowingo Reservoir (Zhang *et al.*, 2016), with two added curves

---

<sup>18</sup> In our context, “flow-censored” samples are manually excluded from computational considerations in somewhat arbitrary way based on associated discharge values for comparative purposes. Such censoring should not be confused with the more conventional definition under which censored samples are those with values less than some nominal minimum reporting level.

that represent scenarios #2 and #3 above. Both scenarios predicted a much smaller net scour (*i.e.*, net loss of storage capacity) than scenario #1 during the two unusual years of 2004 and 2011 (Figure E4).

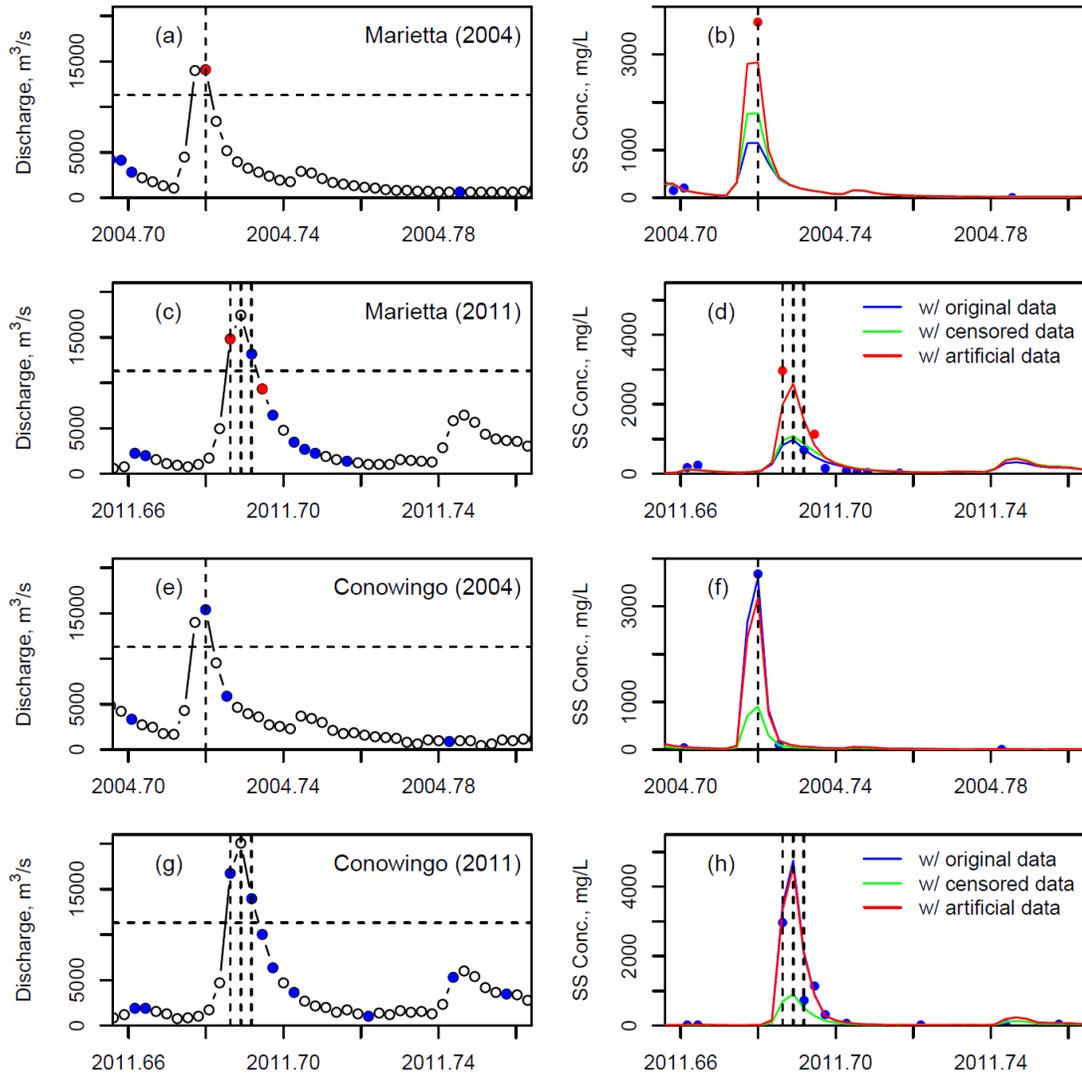


Figure 6.1. Comparison of hydrograph and suspended sediment (SS) chemograph for the Marietta and Conowingo sites during two major storm events, *i.e.*, 2004 Hurricane Ivan and 2011 Tropical Storm Lee, and WRTDS concentration estimates obtained based on three different sampling scenarios. Horizontal dashed lines indicate the scour threshold of  $11,300 \text{ m}^3/\text{s}$ . Vertical dashed lines indicate the key dates during each event, for which the differences of estimates among the three sampling scenarios are summarized in Table E2.

These storm-based analyses demonstrate that, during some (but not all) major storm events, WRTDS estimates at individual sites and hence mass-balance estimates can be heavily affected by the differences in upstream and downstream sampling. Such effects can be particularly relevant if these individual events or years (*e.g.*, 2004 Hurricane Ivan and 2011 TS Lee) are the main interests to managers. Below, the effects of these storm-flow samples on WRTDS regressions and flow-normalized trends are further examined.

### 6.2.2. WRTDS Weight Analysis

Here the weights assigned to individual samples in the WRTDS regression development were quantified and compared. Our results are summarized in Figure E5. Clearly, the relative weights become very small even for those extreme-events. Taking the Ivan sample (2004/09/20; 15,433 m<sup>3</sup>/s) as an example, its relative weight is:

- (i) 0.01 in the regression developed for an estimation day of 1990/09/01 with Q of 5,663 m<sup>3</sup>/s (200,000 cfs).
- (ii) 0.03 in the regression developed for an estimation day of 1990/09/01 with Q of 14,158 m<sup>3</sup>/s (500,000 cfs).
- (iii) 0.03 in the regression developed for an estimation day of 2005/09/01 with Q of 5,663 m<sup>3</sup>/s.
- (iv) 0.09 in the regression developed for an estimation day of 2005/09/01 with Q of 14,158 m<sup>3</sup>/s.

Interestingly, scenarios (ii) and (iii) have almost equal weights due to a trade-off between the relatively high discharge weight in (ii) and the high time and season weights in (iii).

Overall, this analysis shows that the storm samples are simply “diluted” by the large number of samples used in the regression model. The Ivan sample has a slightly higher

weight in scenario (iv), where it is more similar to the estimation day with respect to time, discharge, and season, but its effect on the developed regression is expected to be minor. In many cases, including even those “similar” in time or season (like scenario iii), the weight is very small.

### **6.2.3. WRTDS Regression Surface Analysis**

Here we evaluated if the exceptionally high concentrations during storm events have affected the entire WRTDS regression surface, *i.e.*, the estimated concentration as a function of time and discharge. Four different scenarios were considered:

- (i) estimation using all original samples.
- (ii) estimation using all samples but the Ivan sample (2004/09/20).
- (iii) estimation using all samples but the Lee and Ivan samples.
- (iv) estimation using all samples plus a “hypothetical” 1996/09/20 sample, assuming that the Ivan discharge and concentration also occurred on 1996/09/20.

The results (Figure E6) show that the presence (or absence) of the Ivan sample doesn’t affect the majority of the concentration regression surface. Instead, its effects are limited to the time-discharge space around 2004/09/20 and around the Ivan discharge. The estimation for dates during low to moderate flows has not been largely affected by this or other storm samples. This is consistent with the results of “Weight Analysis” above. In addition, when adding the “hypothetical” 1996/09/20 sample (scenario iv), the regression surface seems almost identical to the one without any manipulation (scenario i), which provides direct evidence to negate the speculation that WRTDS estimation has been biased by a lack of comparable storm-flow samples in the 1990s.

In addition, one can select a particular date (*e.g.*, 2005/09/01) and take the vertical slice on the regression surface corresponding to that date. This slice contains estimated concentration across the full discharge range and is equivalent to the C-Q curve presented in Figure 16 of Hirsch (2012). Here we re-visited the C-Q curves documented by Hirsch (2012) and Zhang *et al.* (2013) and further assigned confidence intervals (CIs) to these results using the recently developed bootstrap approach (Hirsch *et al.*, 2015). Specifically, the original concentration record was resampled (with replacement) 100 times and run WRTDS with each replicate record. The resulting 100 sets of estimates were used to quantify the CIs. For brevity, four dates were selected for comparison, namely, September 1 of 1990 and 2005 (after Hirsch, 2012) and May 15 of 1990 and 2005 (after Zhang *et al.*, 2013). The results are shown in Figure 6.2, with dashed lines indicating 95% CIs.

This figure reminds readers that one should be careful in the interpretation of Figure 16 of Hirsch (2012), because the most recent projection on the September date (*i.e.*, 2005/09/01) was most uncertain, likely due to heavy influences of recent September storms (*i.e.*, Hurricane Ivan and TS Lee). Nonetheless, both the September and May dates show clear separation between the 1990 (black) and 2005 (red) curves for a wide range of discharges, including sub-scour levels. These results highlight the relatively high uncertainty in projected C-Q relationships for September 1, while reaffirming the previously documented temporal changes in reservoir system function (Hirsch, 2012; Zhang *et al.*, 2013; Zhang *et al.*, 2016).

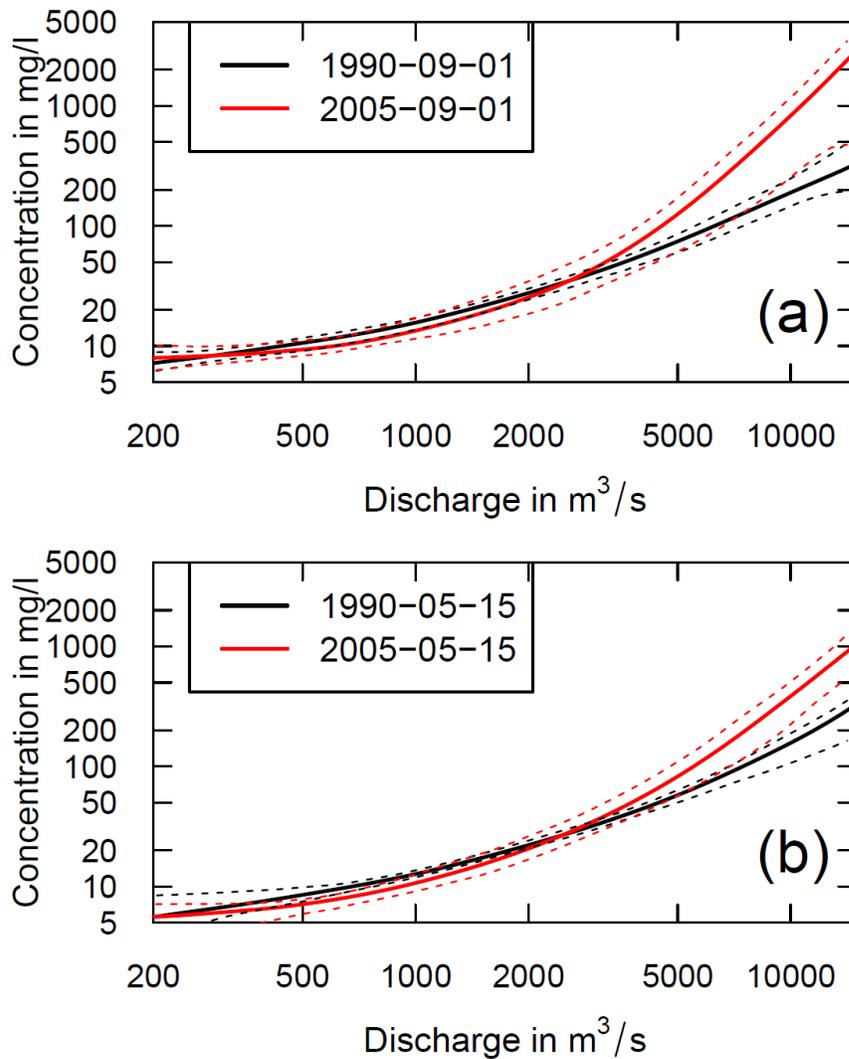


Figure 6.2. Estimated suspended sediment concentration as a function of discharge for the Conowingo site for four selected dates, *i.e.*, September 1 of 1990 and 2005 (after Hirsch, 2012) and May 15 of 1990 and 2005 (after Zhang *et al.*, 2013). Note that both axes are on logarithmic scale. Dashed lines indicate 95% confidence intervals, which were obtained from uncertainty analyses with 100 bootstrap model runs.

#### 6.2.4. WRTDS Flow-Normalized Loading Analysis

Finally, the effects of storm-flow samples on WRTDS trends were investigated. Time series of estimated flow-normalized loadings was reconstructed for (a) September only, (b) all non-September months, and (c) all months (Figure E7). As expected, the presence

of Ivan sample does affect the estimated trend in September – 440% vs. 340% increase, respectively, for scenarios with and without that sample. However, this sample has little effect on the estimated increase for the other 11 months of the year (67% vs. 64%). Therefore, the Ivan sample has a limited influence on the estimated annual increase (110% vs. 93%). Overall, these results are consistent with the results of “Weight Analysis” and “Regression Surface Analysis” above.

In addition, uncertainty analyses were conducted on the flow-normalized trends using the above-mentioned bootstrap approach (Hirsch *et al.*, 2015). As shown in Figure 6.3, the CI on estimated trend is generally wide at the edge years, which follows expectation. Importantly, the CI becomes very wide around 2004, likely due to effects of the unusually high discharge and concentration during Hurricane Ivan. Nonetheless, the overall rising trend since the mid-1990s is very robust.

Moreover, flow-normalized trends were calculated for two important scenarios. In one scenario, all scour-level (*i.e.*, > 11,300 m<sup>3</sup>/s) concentration samples were removed from the original record; this represents a hypothetical case of missed sediment concentration sampling for all scour-level discharges. In the other scenario, all scour-level concentration samples were removed, and additionally, all scour-level discharges were replaced with a nominal but very small value (*i.e.*, 20 m<sup>3</sup>/s); this represents a hypothetical case where no scour-level discharge occurred in the record (Robert Hirsch, personal communication, 2016). The trends are respectively shown as green and purple lines in Figure 6.3. Clearly, the prior conclusion of rising trend at Conowingo is also valid under these two scenarios, although the magnitude is smaller for the latter. Thus, we confirm that the rising trend since the mid-1990s has not been heavily affected by the

presence (or absence) of storm-event samples. Furthermore, such type of analysis was extended to SS and TP at both Marietta and Conowingo. The results (Figure E8) show overall declines at Marietta and rises at Conowingo, thereby providing further evidence to support the prior general conclusions (Hirsch, 2012; Zhang *et al.*, 2013; Zhang *et al.*, 2016).

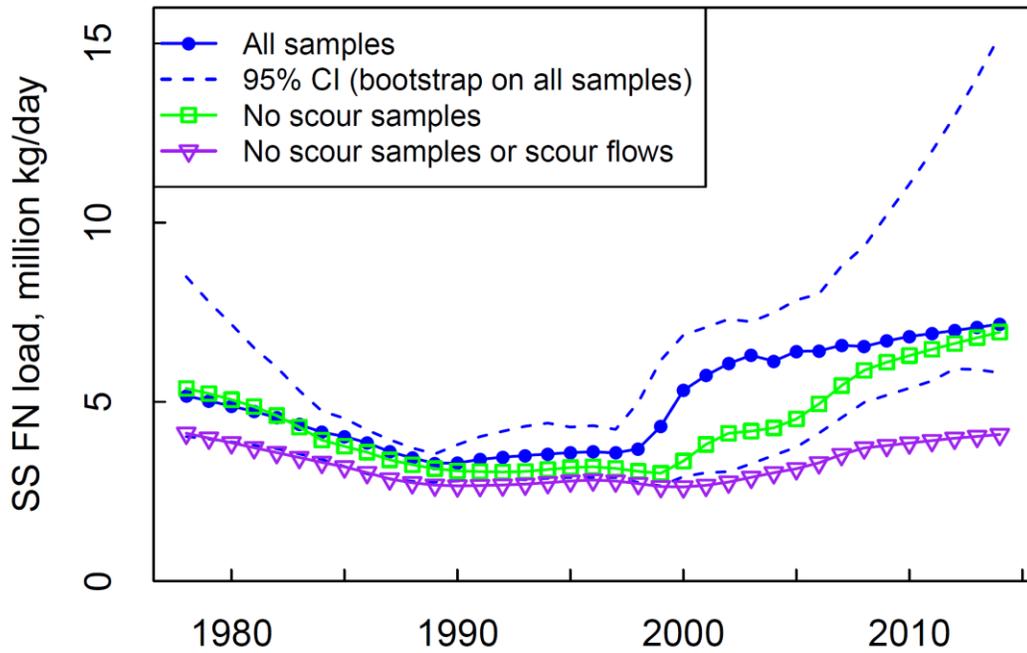


Figure 6.3. Flow-normalized annual load trends of suspended sediment at the Conowingo site, as obtained with different scenarios of sampling data. Solid blue line is based on the original concentration record. Dashed blue lines indicate 95% confidence intervals, as obtained from uncertainty analyses of 100 bootstrap runs. Solid green line is based on a flow-censored concentration record, which eliminates all concentration samples with discharge ( $Q$ )  $>$  scour threshold (*i.e.*,  $11,300 \text{ m}^3/\text{s}$ ). Solid purple line is based on censored records for both concentration and discharge, which eliminates all concentration samples with  $Q > 11,300 \text{ m}^3/\text{s}$  and replaces  $Q$  above  $11,300 \text{ m}^3/\text{s}$  with a nominal value (*i.e.*,  $20 \text{ m}^3/\text{s}$ ).

### 6.3. Summary, Implications, and Prospects

This work serves as an important case study that demonstrates the value of rigorous statistical approaches for analysis of sparse water-quality data while also emphasizing the uncertainties associated with modeled results. Our results confirm that WRTDS is a reasonably robust tool for inferring general water-quality conditions and trends from long-term monitoring data, but high uncertainties can sometimes occur with respect to “true-condition” loadings during the most extreme discharges and “flow-normalized” loadings during specific times and times of year that are proximate to those discharges. Compared with other regression-based approaches, WRTDS seeks to more properly model concentration variations in multiple dimensions (*i.e.*, time, discharge, and season) and, thanks to recent developments, its uncertainties can be explicitly quantified. This is particularly important when data can be confounded by issues of infrequent sampling, storm hysteresis, and temporal skewness of high-flow samples.

For Chesapeake Bay management and protection, the Chesapeake Bay Program needs a realistic representation of the current and future trapping efficiency of the reservoir system, preferably as a function of discharge and storage capacity. The Bay Program will need to evaluate how to represent these changes and the associated uncertainties in the ongoing upgrade of the Chesapeake Bay Watershed Model (Linker *et al.*, 2013; Shenk and Linker, 2013). We believe that Figures S20-S22 of Zhang *et al.* (2016) may be the best existing information to describe how the reservoir reach has historically transformed inputs into outputs and how this modulation has evolved over the last few decades. We suggest that models of this system should use data of this kind as a reasonable basis for calibration and verification, in order to ensure that the models can at least roughly “hind

cast” basic trends that have been inferred from past monitoring data. Better methods and estimates may be possible, but we have not seen any more compelling results from other methods for the Chesapeake system.

Last but not least, we echo Hirsch (2012) in emphasizing that further data collection is crucial toward improving estimation of inputs to and outputs from the reservoirs as well as the associated uncertainties. Particularly important and useful in this regard will be in situ high-frequency water-quality monitoring. Moreover, we encourage frequent re-analysis of both historical discrete samples and newly-collected in situ data to continue to advance our understanding of this complex and ever evolving system.

#### **6.4. Supporting Information**

Supporting information to this chapter is provided in Appendix E.

#### **6.5. Acknowledgements**

This work was supported by the Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090) and Maryland Water Resources Research Center (2015MD329B).

#### **6.6. Literature Cited**

Chanat, J. G., D. L. Moyer, J. D. Blomquist, K. E. Hyer and M. J. Langland, 2016.

Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year 2012. U.S. Geological Survey Scientific Investigations Report 2015-5133, Reston, VA, p. 76.  
<http://dx.doi.org/10.3133/sir20155133>.

- Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.
- Gross, M. G., M. Karweit, W. B. Cronin and J. R. Schubel, 1978. Suspended sediment discharge of the Susquehanna River to northern Chesapeake Bay, 1966 to 1976. *Estuaries* 1:106-110, <http://www.springerlink.com/index/r4796375t48061x3.pdf>.
- Hagy, J. D., W. R. Boynton, C. W. Keefe and K. V. Wood, 2004. Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658, DOI: 10.1007/bf02907650.
- Harmel, R. D. and K. W. King, 2005. Uncertainty in measured sediment and nutrient flux in runoff from small agricultural watersheds. *Trans. ASAE* 48:1713-1721.
- Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.
- Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.

- Johnes, P. J., 2007. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology* 332:241-258, DOI: 10.1016/j.jhydrol.2006.07.006.
- Lang, D. J., 1982. Water quality of the three major tributaries to the Chesapeake Bay, the Susquehanna, Potomac, and James Rivers, January 1979-April 1981. US Geological Survey, Water Resources Division Water-Resources Investigations Report 82-32, p. 64. <https://pubs.er.usgs.gov/publication/wri8232>.
- Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18. <http://dx.doi.org/10.3133/ofr20141235>.
- Linker, L. C., R. A. Batiuk, C. F. Cerco, G. W. Shenk, R. Tian, P. Wang and G. Yactayo, 2016. Influence of Reservoir Infill on Coastal Deep Water Hypoxia. *J. Environ. Qual.* 45, DOI: 10.2134/jeq2014.11.0461.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- The Lower Susquehanna River Watershed Assessment Team, 2015. Lower Susquehanna

River Watershed Assessment, Maryland and Pennsylvania. p. 192.

<http://dnr.maryland.gov/bay/lswa/report.htm>.

U.S. Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment. Annapolis, MD.

<https://www.epa.gov/chesapeake-bay-tmdl/chesapeake-bay-tmdl-document>.

Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI: 10.1016/j.scitotenv.2013.02.012.

Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.

Zhang, Q., R. M. Hirsch and W. P. Ball, 2016. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* 50:1877-1886, DOI: 10.1021/acs.est.5b04073.

*Page intentionally left blank*

## Chapter 7. An Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-Quality Data<sup>19</sup>

### Abstract

Derived from river monitoring data, concentration-discharge ( $C$ - $Q$ ) relationships are powerful indicators of export dynamics. Proper interpretation of such relationships can be made complex, however, if the relationships of  $\ln(C)$ - $\ln(Q)$  are nonlinear or if the relationships change over time, season, or discharge. Methods of addressing these issues by “binning” data or smoothing trends can introduce artifacts and ambiguities that obscure underlying interactions among time, discharge, and season. Here we illustrate these issues with examples and propose an alternative method that uses the regression coefficients of the recently-developed WRTDS (“Weighted Regressions on Time, Discharge, and Season”) model for examining riverine  $C$ - $Q$  relationships, including their uncertainty. The method is applied to sediment concentration data from Susquehanna River at Conowingo Dam (Maryland, USA) to illustrate how the WRTDS coefficients can be accessed and presented in ways that provide additional insights toward the interpretation of river water-quality data. The results clearly reveal that sediment concentration in the reservoir effluent has become more sensitive to discharge at moderate and high flows (but not very low flows) as it approaches sediment storage

---

<sup>19</sup> This chapter will be submitted for publication in a peer-reviewed journal. Bill Ball and Ciaran Harman were involved in hypothesis development, results interpretation, and editing. All figures and data were created by Qian Zhang.

capacity, reaffirming the recently-documented decadal-scale decline in reservoir trapping performance.

## 7.1. Introduction and Background

Derived from river monitoring data, concentration-discharge ( $C$ - $Q$ ) relationships are a powerful tool for understanding complex interactions between hydrologic and biogeochemical processes, including nutrient and sediment export dynamics (Evans and Davies, 1998; Chanut *et al.*, 2002; Godsey *et al.*, 2009; Meybeck and Moatar, 2012). In particular, the manner in which  $C$  varies with  $Q$  reflects the relative role of dilution in comparison to mechanisms of mobilization, such as dissolution and erosion (House and Warwick, 1998; Richardson, 2012; Bieroza and Heathwaite, 2015; Herndon *et al.*, 2015). In this context,  $C$ - $Q$  relationships have historically been classified into three broad categories – namely, (1) “dilution” (*i.e.*, negative relationship), as commonly observed for point-source dominated rivers; (2) “mobilization” (*i.e.*, positive relationship), as commonly observed for many nonpoint-source dominated rivers; and (3) “chemostasis” (*i.e.*,  $C$  invariant with  $Q$ ). In previous studies, the classified category has been shown to vary with constituent and site (Godsey *et al.*, 2009; Hirsch *et al.*, 2010; Meybeck and Moatar, 2012; Stallard and Murphy, 2014; Herndon *et al.*, 2015). In addition,  $C$ - $Q$  relationships may change over time in ways that can serve as useful indicators of changes in river system function (Richardson, 2012; Bieroza and Heathwaite, 2015; Burt *et al.*, 2015; Gray *et al.*, 2015; Zhang *et al.*, 2016b). These changes may in turn reflect anthropogenic activities such as land disturbance and watershed management, particularly if the stochastic properties of river discharge have remained roughly unchanged over time.

### 7.1.1. Complexities with *C-Q* Interpretation

In practice, log-log linear relationships have been used to characterize paired *C-Q* data – that is, by assuming  $C = a Q^b$  or  $\ln(C) = \ln(a) + b \ln(Q)$  (Walling, 1977; Crowder *et al.*, 2007; Meybeck and Moatar, 2012; Stallard and Murphy, 2014; Herndon *et al.*, 2015). As a closely related approach, other authors have characterized paired loading (*L*) vs. discharge relationships in a similar way, *i.e.*,  $\ln(L) = \ln(a') + b' \ln(Q)$  (Basu *et al.*, 2010; Musolff *et al.*, 2015; Zhang *et al.*, 2016a). The fitted slope of either approach (*b* or *b'*) can be useful for classification of export patterns in a given watershed, and for comparisons of export patterns among sites and/or constituents (Godsey *et al.*, 2009; Basu *et al.*, 2010; Meybeck and Moatar, 2012; Zhang *et al.*, 2016a). The validity of such relationships is confounded, however, by other factors that affect concentrations in natural streams and rivers. Such factors include:

- (1) Variations of the mathematical form of the relation for *C-Q* relations with discharge -- that is, non-linearity of the relationship of  $\ln(C)$  vs.  $\ln(Q)$ .
- (2) Variations of *C-Q* relations over time (“time effect” or “non-stationarity”).
- (3) Variations of *C-Q* relations with season (“seasonal effect”).

All of these complexities commonly exist in real water-quality data (Cohn, 1995; Horowitz, 2003; Crowder *et al.*, 2007; Hirsch *et al.*, 2010; Meybeck and Moatar, 2012; Hirsch, 2014; Warrick, 2014; Zhang *et al.*, 2016b), as illustrated by Figure 7.1 for the case of suspended sediment (SS) data from Susquehanna River at Conowingo, MD. This site lies downstream of Conowingo Reservoir, where prior studies have suggested long-term decline in the reservoir’s ability to trap sediment before it reaches Chesapeake Bay (Hirsch, 2012; Zhang *et al.*, 2013; The Lower Susquehanna River Watershed Assessment

Team, 2015; Zhang *et al.*, 2015; Zhang *et al.*, 2016). Ideally,  $C$ - $Q$  analysis would help us to reaffirm such questions quantitatively, but the above-noted complexities must be carefully accounted for in the analysis.

To address the issues identified above, one can imagine some simple solutions. In regard to issue #1, for example, one partial solution could be to fit non-linear and

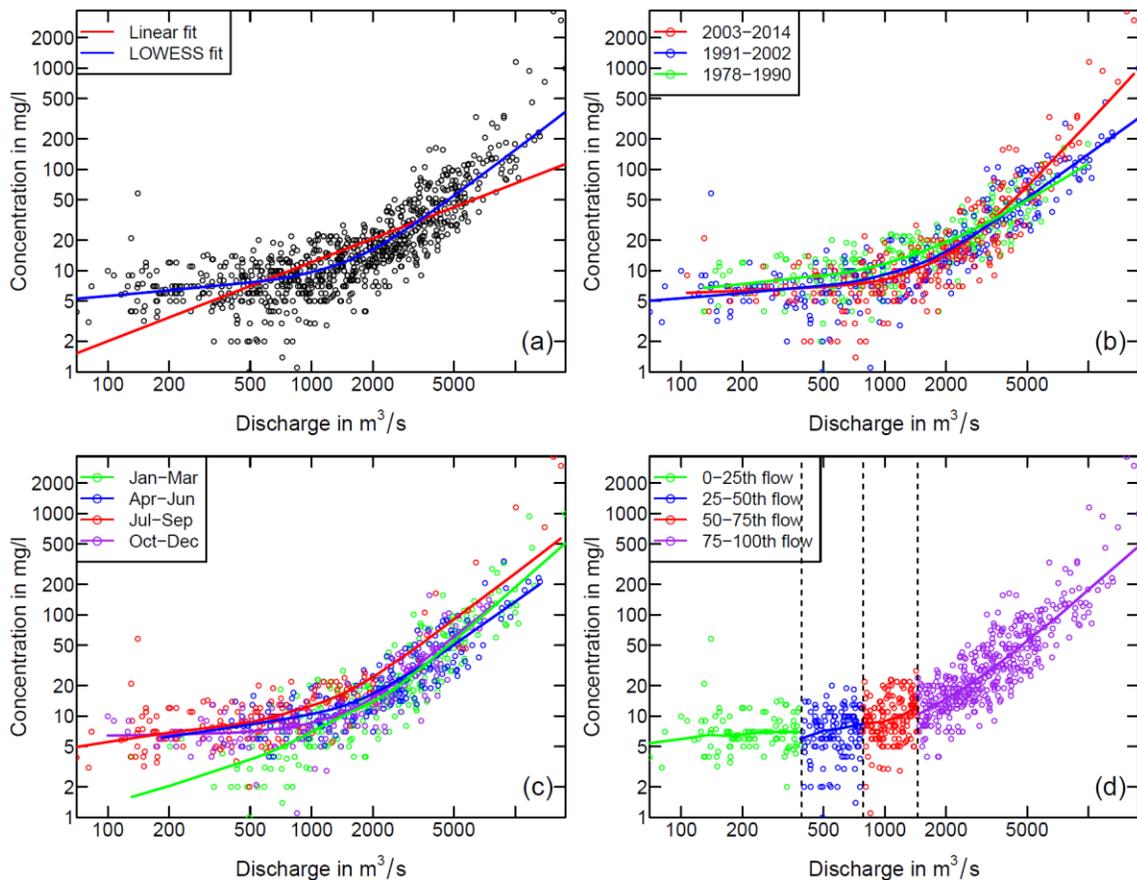


Figure 7.1. Illustration of common issues on interpretation of concentration-discharge ( $C$ - $Q$ ) relationships using the example of suspended sediment concentration data in Susquehanna River at Conowingo, MD. (a) Direct linear fit (red line) to the entire data sets assumes linearity in the  $C$ - $Q$  relationship. The LOWESS fit (blue line) doesn't invoke such an assumption. (b) LOWESS fits for  $C$ - $Q$  data in three different temporal periods. (c) LOWESS fits for  $C$ - $Q$  data in four different seasons. (d) LOWESS fits for  $C$ - $Q$  data in four quartiles of discharge.

preferably, non-parametric curves in the  $\ln(C)\sim\ln(Q)$  space, thereby allowing the fitted slope to vary. One such approach is to generate a LOWESS (locally weighted scatterplot smoothing) curve, which is largely data-driven and requires minimal use of statistical modeling (Figure 7.1a; also see Hirsch *et al.* (2010); Hirsch (2014); Warrick (2014); Zhang *et al.* (2016b) for examples). In regard to issues #2-#3, an intuitive solution would be to fit multiple  $C$ - $Q$  curves (*e.g.*, LOWESS) for data in separated bins with respect to time or season, and this could also be done with discharge, as a further means of addressing issue #1. Specifically,  $C$ - $Q$  data pairs may be separated into several non-overlapping periods (Figure 7.1b; also see Warrick (2014) and Zhang *et al.* (2016b) for examples), four seasons (or twelve months) of the year (Figure 7.1c; also see Hirsch (2014) for examples), or various quantiles of the discharge record (Figure 7.1d; also see Meybeck and Moatar (2012) for examples). In many cases, these approaches can provide more legitimate inferences on the underlying  $C$ - $Q$  relationships. However, these approaches give rise to several additional concerns:

- (1) LOWESS curves and other similar approaches require specifications of the smoothing window (span) and the degree of polynomials for the fitting. In this regard, the original single regression corresponds to setting the smoothing span equal to the entire length of data and the polynomial degree to one. Selections of shorter spans of data or higher degrees of polynomial are somewhat arbitrary.
- (2) Direct fitting to observed  $C$ - $Q$  data is further complicated by the fact that traditional (discrete) monitoring data are typically sparse and so similar numbers of samples may not be available for the various bins of time, season, or discharge.
- (3) The approach that fits multiple LOWESS curves using data from different “bins”

of time, discharge, or season is more sophisticated. This approach, however, requires making subjective choices on selection of bin intervals. Moreover, this approach makes the low-frequency  $C$ - $Q$  data even more sparse within each individual bin.

- (4) Interactions among time, season, and discharge need to be decoupled to the extent possible, yet this concern is simply overlooked by both linear and LOWESS fits.

Clearly, any of these issues can complicate or even mislead the interpretation of  $C$ - $Q$  relationships. In this context, the main objective of this article is to describe and evaluate a more robust, informative, and accessible solution for the interpretation of  $C$ - $Q$  patterns in riverine monitoring data through the recently-developed WRTDS (“Weighted Regressions on Time, Discharge, and Season”) method (Hirsch *et al.*, 2010). In the following sections, I introduce the rationale and algorithms of WRTDS, explain why it offers improvements in regard to the above-listed concerns, and present several specific ways of presenting WRTDS coefficients and using these to interpret  $C$ - $Q$  relationships in riverine monitoring data.

### 7.1.2. Solution: The WRTDS Method

WRTDS has been developed by the U.S. Geological Survey (USGS) to improve the statistical estimation of daily concentrations and loadings based on historically common practices of discrete (and often low-frequency) water-quality sampling for concentration. Similar to many other regression-based approaches, WRTDS makes use of time, discharge, and season as explanatory variables:

$$\ln(C_i) = \beta_{0,i} + \beta_{1,i}t_i + \beta_{2,i} \ln(Q_i) + \beta_{3,i} \sin(2\pi t_i) + \beta_{4,i} \cos(2\pi t_i) + \varepsilon_i \quad (7.1)$$

where  $t_i$  is time in decimal years,  $C_i$  is daily concentration at time  $t_i$ ,  $Q_i$  is daily discharge at time  $t_i$ ,  $\beta_{0,i} \sim \beta_{4,i}$  are fitted coefficients, and  $\varepsilon_i$  is the error term. The 1<sup>st</sup> - 3<sup>rd</sup> terms on the right of Equation (7.1) represent the intercept, time effects, and discharge effects, respectively, whereas the 4<sup>th</sup> and 5<sup>th</sup> terms collectively represent cyclical seasonal effects.

For each day of estimation, WRTDS pre-screens all available concentration samples to select at least 100 samples that are sufficiently “close” to that estimation day. This proximity is evaluated with three distances, *i.e.*, time, discharge, and seasonal distances. Using the tri-cube weight function and selected half-window widths, these distances are converted to time, discharge, and seasonal weights, respectively, and their product is the total weight assigned to that sampled day. Reasonable default half-window widths suggested by Hirsch and De Cicco (2015) for many rivers (and specifically the Susquehanna) are 7 years, 0.5 year, and 2 natural log units for time, season, and discharge, respectively. The selected samples are then used to fit Equation (7.1) and the fitted coefficients are used to estimate  $\ln(C_i)$ , which is then transformed back to concentration. WRTDS has been documented to offer better estimation results than previous approaches, in large part because it does not rely on assumptions about the homoscedasticity of model errors, constancy of seasonal trends in concentration, or constancy of the  $C$ - $Q$  relationship (Hirsch *et al.*, 2010; Moyer *et al.*, 2012; Chanat *et al.*, 2016).

A key feature of WRTDS is that a unique model (*i.e.*, Equation 7.1) is developed for each day in the record, thereby allowing the dependencies of concentration on time, discharge, and season to be variable across the space. Consequently, WRTDS is largely immune to the issues listed above:

- (1) It does not assume a linear  $\ln(C) \sim \ln(Q)$  relation or any other functional forms at the scale of the whole dataset. Instead, it conducts “local” linear fit at many points in the time-discharge space in a consistent manner.
- (2) It allows the  $C$ - $Q$  relation to vary with time.
- (3) It allows the  $C$ - $Q$  relation to vary with season.
- (4) It allows the  $C$ - $Q$  relation to vary with discharge.
- (5) It avoids the specification of any smooth window or polynomial degree (though the choice of half-window width does introduce an element of subjective choice).
- (6) It avoids making subjective choices regarding how to separate intervals for time, discharge, or season. Instead, it always evaluates all available concentration samples and uses the most “relevant” data for each day of estimation, thereby making full use of the sparse samples.
- (7) Its model framework (*i.e.*, Equation 7.1) can, at least partially, decouple the interactions among time, discharge, and season.
- (8) It is less sensitive to the quality and quantity of typically sparse high-flow samples in the record, as compared with the common approaches discussed above.

These advantages suggest that the WRTDS model coefficients, specifically  $\beta_2$  coefficients (which reflect  $\partial \ln C / \partial \ln Q$ ; see Equation 1), may provide a clearer and more nuanced picture of riverine  $C$ - $Q$  relationships than previous methods. These coefficients are hidden from user view in current software implementations, but they can be conveniently extracted with additional coding. Such coding has been developed in this work and is made available to the readers – see below.

In this article, I describe the application of such an analysis and present several ways of reporting the  $\beta_2$  coefficients to illustrate how this approach can add insights to the analysis of a decadal-scale water-quality record. In particular, I use the example of SS concentrations in Susquehanna River at Conowingo, MD, for which daily streamflow data and more sparse SS concentration data were collected from the USGS National Water Information System (NWIS) Web Interface (U.S. Geological Survey, 2014) for the period between 1984 and 2015.

## 7.2. Methods: WRTDS $\beta_2$ Extraction and Use

The WRTDS method is available for public use through the *EGRET* (Exploration and Graphics for RivEr Trends) package, which currently in version 2.2.0 (Hirsch and De Cicco, 2015) and is run in the *R* statistical programming language (R Development Core Team, 2014). In its current implementation, WRTDS establishes a set of evenly-spaced grid points on a surface defined by time ( $t$ ) and  $\ln(Q)$ . The standard grid design is fully described in the user manual (Hirsch and De Cicco, 2015). Briefly, for  $x$ -axis, time grid values are spaced 0.0625 years ( $1/16^{\text{th}}$  of a year) apart from the beginning year to the end year of the record. For  $y$ -axis, 14 grid values are spaced with equal distance in log space for the discharge range from 5 percent below the minimum  $Q$  to 5 percent above the maximum  $Q$  in the record. At each grid point, WRTDS develops a separate weighted-regression model using Equation (7.1), which results in an estimated concentration “surface” as functions of  $t$  and  $\ln(Q)$ . Daily concentration is then estimated using a bi-linear interpolation of this concentration surface (Hirsch and De Cicco, 2015).

For the work described herein, I have installed the *EGRET* package (version 2.2.0) and made modifications that allow users to extract and save  $\beta_2$  coefficients from each

regression made. In brief, the *EGRET* codes were modified to run the same regression model (*i.e.*, Equation 7.1) on the same  $t$ - $Q$  grid but to extract and store  $\beta_2$  coefficients instead of concentration estimates. In a similar way, several existing *EGRET* plotting functions were modified to provide contours of the estimated  $\beta_2$  coefficients in lieu of concentrations, as presented subsequently. The original versions of these functions are described in the user manual (Hirsch and De Cicco, 2015), whereas the modified functions are documented at the publicly accessible Johns Hopkins University Data Archive via <http://dx.doi.org/10.7281/T18G8HM0> (Zhang and Ball, 2016). In addition, variance inflation factor (VIF; Kutner *et al.*, 2004) was calculated for each regression model developed in WRTDS to quantify the severity of collinearity among the independent variables. In general, multicollinearity is considered high if  $VIF > 5$  (Kutner *et al.*, 2004). In our case, the VIFs are always  $< 2$ , indicating that collinearity is not an issue.

The estimated  $\beta_2$  coefficients can be used to categorize export patterns – namely, (1) “dilution” (*i.e.*,  $\beta_2 < 0$ ); (2) “chemostasis” (*i.e.*,  $\beta_2 = 0$ ); and (3) “mobilization” (*i.e.*,  $\beta_2 > 0$ ). To obtain estimates of the uncertainty of the results, we have employed the “block-bootstrap” method of Hirsch *et al.* (2015), which involves resampling (with replacement) of the raw concentration data to obtain multiple realizations (*e.g.*, 50) of representative data sets and associated  $\beta_2$  estimates. See Zhang *et al.* (2016b) for more details of this approach.

In the remainder of this article, I present application of the above-described methods to our selected case study and provide our interpretation of these results. In addition to confirming some previous findings with respect to SS trends in Susquehanna River, this

exercise provides additional insights about the causative factors and illustrates the utility of the proposed approach.

### 7.3. Results and Discussion

#### 7.3.1. Visualization of WRTDS $\beta_2$ Coefficients

The estimated  $\beta_2$  coefficients for the case of suspended sediment in Susquehanna River at Conowingo, MD, are shown in Figure 7.2 as a contour plot against axes of time and discharge. This plot provides a rich set of information regarding  $\beta_2$  coefficients. First, Figure 7.2 reveals that the coefficients clearly vary with discharge, with higher discharges generally corresponding to larger coefficients – *i.e.*, the relationship is non-linear, with a stronger influence of discharge on concentration at high discharge. Second, Figure 7.2 also reveals that the coefficients vary with time and season and that these temporal variations also appear to be discharge-dependent. For low discharges (*e.g.*, 100 m<sup>3</sup>/s), the coefficients are positive (*i.e.*, “mobilization” effect) for most of the time but are negative (*i.e.*, “dilution” effect) during most months in the late 1990s and during some months in the other years. For high discharges (*e.g.*, 10,000 m<sup>3</sup>/s), the coefficients are always positive, but with very distinctly different magnitudes between the pre-2000 period ( $\beta_2$ : 1~2) and the post-2000 period ( $\beta_2$ : 2~2.5). This functional change reaffirms the recently-documented fact that the trapping efficiency of Conowingo Reservoir has diminished during the recent period, which has been attributed to the condition that sediment storage behind Conowingo Dam is approaching the reservoir’s capacity (Hirsch, 2012; Zhang *et al.*, 2013; Friedrichs *et al.*, 2014; Langland, 2015; The Lower Susquehanna River Watershed Assessment Team, 2015; Zhang *et al.*, 2016b). Moreover, Figure 7.2 also

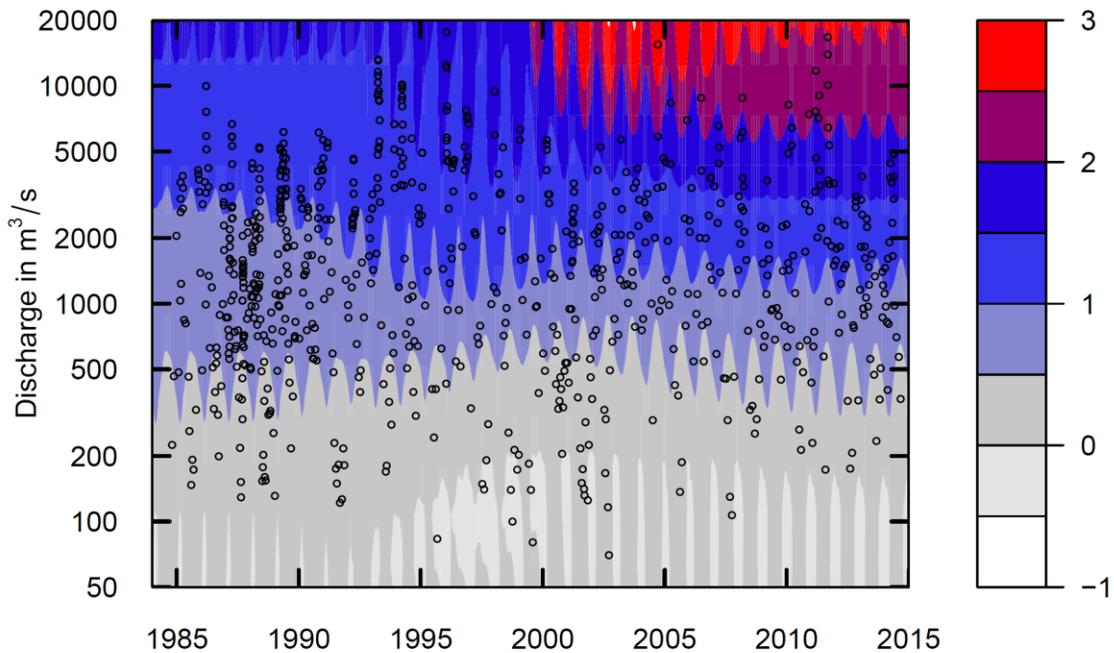


Figure 7.2. Contour plot showing estimated WRTDS  $\beta_2$  coefficients as a function of time and discharge for suspended sediment in Susquehanna River at Conowingo, MD. Black open circles indicate the  $t$ - $Q$  combinations where concentration samples were taken.

provides the additional insight that the  $\beta_2$  coefficient has become larger since 2000. In other words, the sensitivity of sediment concentration to discharge has been increasing since around the year 2000. This interpretation of the evolution of the reservoir's response to discharge would have been obscured if concentration trends were analyzed over time or discharge alone – *i.e.*, without considering the evolution of the  $\partial(\ln C)/\partial(\ln Q)$  relationship. Finally, an important advantage of WRTDS over other statistical approaches is that the results for low to moderate discharges are less sensitive to the quality and quantity of typically sparse high-flow samples in the record. In this regard, Figure 7.2 also shows (as open circles) the times and discharges at which measured concentration samples were taken. This display provides useful information to readers and sets a reminder with respect to where the samples are more sparse and the results less certain. In

this case, samples are clearly more sparse at discharges exceeding scour levels (*i.e.*, > 11,300 m<sup>3</sup>/s) for this reservoir (Gross *et al.*, 1978; Lang, 1982).

An additional use of Figure 7.2 is toward direct comparisons of coefficients among selected years for any given level of discharge. For example, one can use Figure 7.2 to compare the changes in coefficient between the years 1990 and 2010, which are plotted in Figure 7.3. The results clearly indicate that the sensitivity of concentration to discharge has increased in all seasons at most discharge conditions (> ~1,000 m<sup>3</sup>/s), although the sensitivity has actually decreased in most seasons at very low discharges (< ~500 m<sup>3</sup>/s).

### 7.3.2. Trends of WRTDS $\beta_2$ Coefficients with Respect to Season and Discharge

The estimated  $\beta_2$  coefficients can be grouped by discharge percentiles to reveal discharge dependency (Figure 7.4a). Overall, these coefficients follow a monotonic

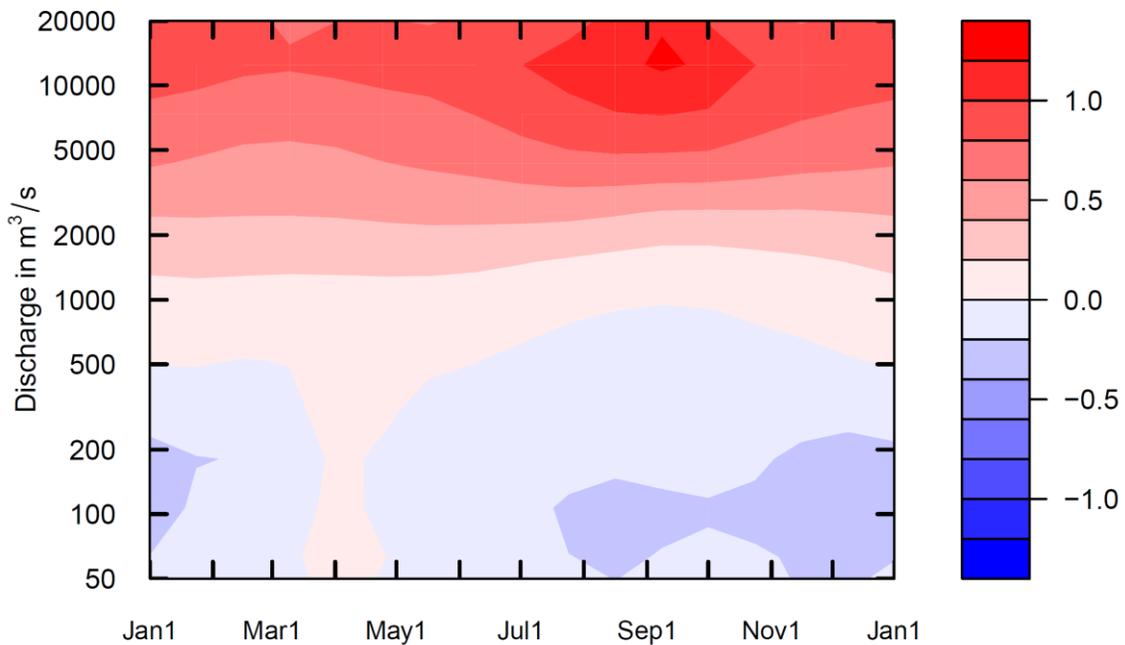


Figure 7.3. Contour plot showing estimated change in  $\beta_2$  coefficients between year 1990 and year 2010 for suspended sediment in Susquehanna River at Conowingo, MD.

pattern with respect to discharge, with highest values occurring at the highest discharge interval (90<sup>th</sup>~100<sup>th</sup> percentile) and lowest values in the lowest discharge interval (0<sup>th</sup>~10<sup>th</sup> percentile). For the latter, the coefficients have a median of around zero and exhibit chemostasis or even dilution effects. Such low-flow patterns indicate the likely existence of discharge thresholds for sediment mobilization, below which runoff generation in the watershed has presumably been insufficient to mobilize sediments, or river discharge is perhaps hydrologically disconnected from sediment source areas (Shanley *et al.*, 2011). In addition, there could be discharge thresholds above which the scour of sediments from riverbanks and riverbeds increases disproportionately with discharge. Here the sampling site is at the outlet of a large reservoir, whose effects must also be considered. Interestingly, the  $\beta_2$  coefficients at the upstream Marietta sampling location (reservoir inlet) also monotonically increase with discharge, but are always above 0.5, suggesting mobilization behavior at all discharges (data not shown). The low-flow chemostasis at Conowingo reflects processes of sediment removal and local generation that are insensitive to rate of discharge in the low-flow range. One possible scenario, for example, is that under such low-flow conditions, the low SS concentrations observed at Conowingo represent a combination of very fine influent SS and newly generated (autochthonous) materials. In this regard, we might expect the former concentrations to increase with increasing discharge rate (reduced sedimentation) and the latter to decrease (reduced residence time for growth), perhaps in an approximately compensatory manner.

Similarly, the  $\beta_2$  coefficients can also be grouped by calendar months to reveal overall intra-annual (seasonal) changes in  $C$ - $Q$  relationships (Figure 7.4b). Overall, the

coefficients follow a cyclic pattern, with highest values occurring in March-April (early spring) and lowest values in July-September (summer). The coefficients in summer months are close to zero (median:  $\sim 0.3$ ), exhibiting the tendency toward chemostasis. In other words, sediment concentration has very a weak sensitivity to changes in discharge during the summer months and a much stronger sensitivity in the spring months.

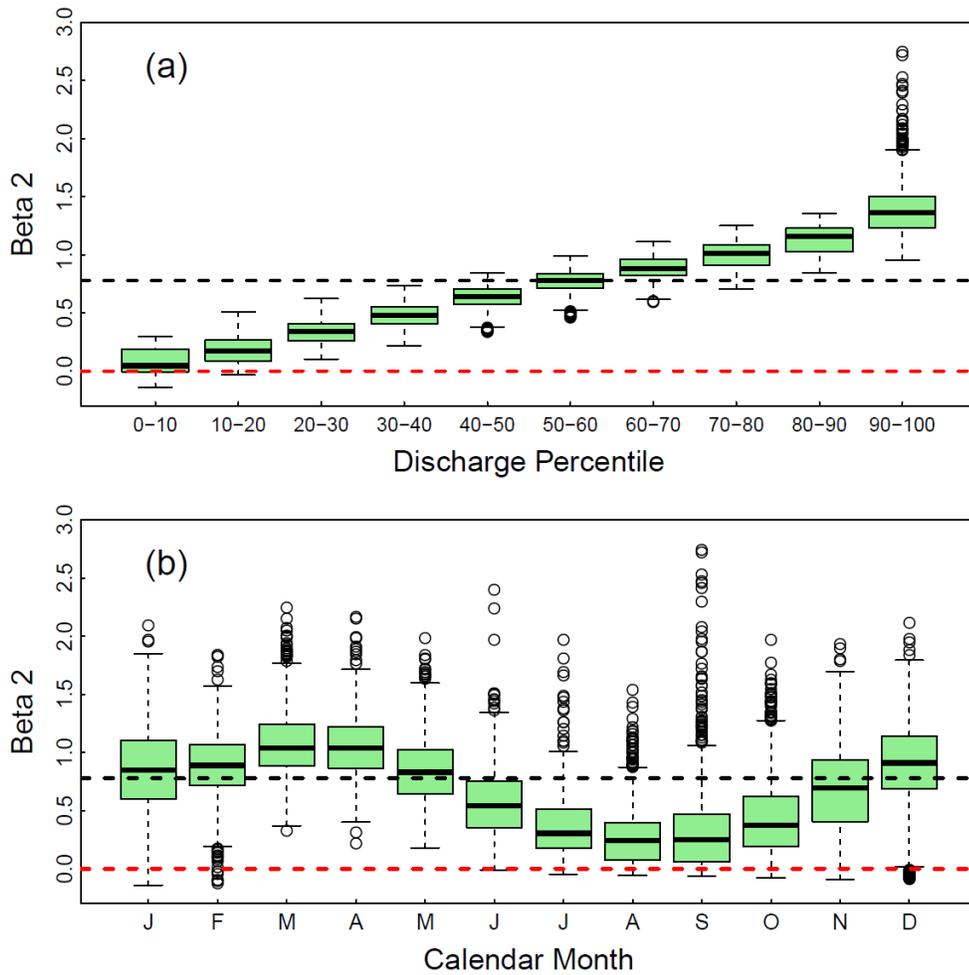


Figure 7.4. Boxplots of estimated daily  $\beta_2$  coefficients by (a) discharge percentile and (b) calendar month for suspended sediment in Susquehanna River at Conowingo, MD. The dashed black lines indicates the estimated slope (0.83) obtained from linear fit to all data – see Figure 7.1a. The red dashed line indicates no sensitivity to discharge (*i.e.*, chemostasis).

However, discharge tends to be higher in spring than in summer, so this pattern would be expected even without additional seasonal effects. To isolate seasonal effects and also to illustrate the usefulness of the contour plot (Figure 7.2), we selected a representative discharge (885 m<sup>3</sup>/s – approximately the median value of the daily discharge record), and extracted  $\beta_2$  coefficients from the  $\beta_2$  surface for this discharge throughout the period of record. These coefficients are again plotted by calendar months (Figure 7.5). Clearly, Figure 7.5 shows a much weaker seasonal variability in  $\beta_2$  coefficients than Figure 7.4b, reaffirming the importance of decoupling interactions between season and discharge when interpreting C-Q relationships. Nonetheless, Figure 7.5 still shows a cyclic pattern, with higher coefficients in spring than summer. Such pattern may relate to intra-annual changes in ground cover, which is a hypothesis worthy of further exploration.

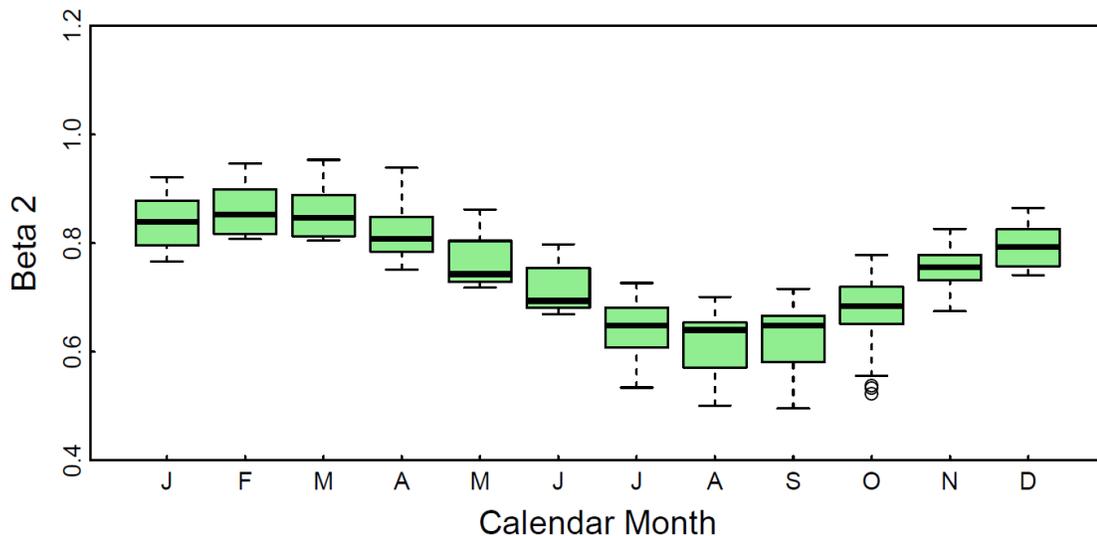


Figure 7.5. Boxplots of estimated  $\beta_2$  coefficients by calendar month for a selected discharge, *i.e.*, 885 m<sup>3</sup>/s (approximately the median value of the daily discharge record), for suspended sediment in Susquehanna River at Conowingo, MD. These coefficients were extracted from the  $\beta_2$  surface (*i.e.*, Figure 7.2) for this particular discharge throughout the period of record.

Finally, the dashed black lines in the two panels of Figure 7.4 represent the fitted slope ( $\sim -0.83$ ) obtained from a single linear regression on all  $\ln(C) \sim \ln(Q)$  data (see Figure 7.1). This value, however, is clearly not representative of the estimated coefficients in many discharge intervals (Figure 7.4a) and many calendar months (Figure 7.4b). Such deviations further highlight the shortcomings of assuming fixed slope without regard to season or discharge – not only do such approaches lose information, but they can also lead, in some cases, to inappropriate conclusions.

### 7.3.3. Temporal Trends in WRTDS $\beta_2$ Coefficients with Uncertainty Analysis

The WRTDS framework allows us to examine long-term trends in the estimated  $\beta_2$  coefficients at different discharges and at different times of the year. The extracted trends separate out the effects of inter-annual discharge variability, and can thus provide a more focused view of temporal changes in river water quality that can lead to new insights about causative factors.

As an example of such an analysis, Figure 7.6 shows the annual averages of estimated  $\beta_2$  coefficients for three selected discharges, as derived from the contour plot (Figure 7.2). The three selected discharges are 181; 2,550; and 7,330  $\text{m}^3/\text{s}$  and represent low-, mid-, and high-flow conditions, respectively. In addition, we use Figure 7.6 to illustrate how the method can be used to provide uncertainty analyses on the estimated  $\beta_2$  coefficients. Specifically, the figure shows the 90% confidence interval (dashed lines) obtained using 50 model runs on concentration data that were obtained with the aforementioned method of resampling with replacement (Hirsch *et al.*, 2015).

For the low-flow condition (black lines),  $\beta_2$  coefficients show a moderate decline in the period of record, with a major decline occurring in the late 1990s. By contrast,  $\beta_2$

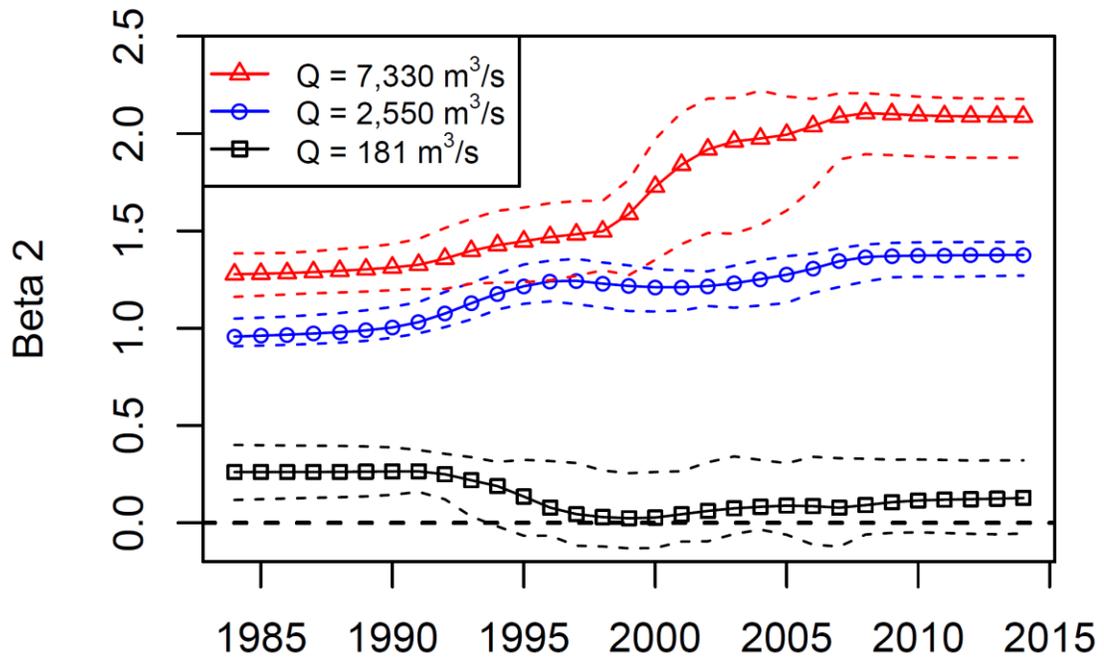


Figure 7.6. Annual averages of estimated  $\beta_2$  coefficients for three selected discharges for suspended sediment in Susquehanna River at Conowingo, MD. These coefficients were extracted from the  $\beta_2$  surface (*i.e.*, Figure 7.2) for these hypothetical cases of constant discharge throughout the period of record. The dashed lines represent the 90% confidence interval for coefficients derived from 50 replicates model runs that were based on resampled (with replacement) concentration data – see Section 7.2.

coefficients for the mid-flow condition (blue lines) show the opposite trend, with a moderate increase during the period of record and a major increase in the late 1990s. For the high-flow condition (red lines), the trend is similar to that for mid-flow, except with a much stronger increase and with the major increase occurring not only in the late 1990s but also through early 2000s. These observed trends reflect best estimates of the effect of reservoir system change on the nature of the  $\ln(C)$  vs.  $\ln(Q)$  relationship at each selected discharge. The physical implication of these results is that sediment concentration in the reservoir effluent has become more sensitive to discharge at moderate and high flows as

it approaches sediment storage capacity. These  $\beta_2$ -inferred system changes have occurred at discharges well below the literature-documented scour threshold (*i.e.*, 11,300 m<sup>3</sup>/s) (Gross *et al.*, 1978; Lang, 1982). This is consistent with our previous conclusion, based on discharge-binned analysis of loading estimates, that net deposition rates are decreasing even in the absence of scour (Zhang *et al.*, 2016b).

Notably, there is a much wider confidence interval for the 2000-2010 period for the high-flow results. This is due to the very limited quantity of sediment samples during Hurricane Ivan in September 2004 as well as the unusually high sediment concentration observed during this event (3,980 mg/L; highest concentration in the entire record). In this regard, the confidence interval serves to inform the reader about the higher uncertainty in the Ivan time interval, even while providing added support for the overall conclusion of long-term rising trend in  $\beta_2$ .

#### **7.3.4. Limitations of the Proposed Approach**

This proposed approach has several limitations in terms of applicability. First, the approach requires an adequate number of samples to properly conduct the analysis over the intended record, *e.g.*, ~800 sampled days in the 30-year record addressed here. Hirsch *et al.* (2010) recommended concentration data coverage of at least 20 years, coupled with a sampling frequency of at least 10 samples per year and a complete record of daily discharge. These data requirements can be met by many water-quality monitoring sites in the United States and elsewhere. Second, the approach assumes smooth water quality changes (Hirsch *et al.*, 2010). Rapid changes in water quality dynamics may be obscured by smoothing introduced by the method. Third, as with other methods such as those described in Section 7.2, this approach relies on an adequate quality and quantity of

samples at high discharges where major flux is involved, although the method is better than many at minimizing the effect of high-discharge samples on analysis of low- or mid-level discharge conditions. Finally, this approach has been developed for long-term flux and trend analysis from traditional (low-frequency) monitoring data and is no substitute for continuous (high-frequency) data where shorter-term variability (*e.g.*, storm-based  $C$ - $Q$  hysteresis) is important to consider (Outram *et al.*, 2014; Pellerin *et al.*, 2014; Bieroza and Heathwaite, 2015; Bowes *et al.*, 2015; Halliday *et al.*, 2015; Miller *et al.*, 2016). In this regard, adding the capability of evaluating sub-daily dynamics has been recognized as an important topic for further research (Hirsch *et al.*, 2010).

#### **7.4. Summary**

Riverine concentration-discharge ( $C$ - $Q$ ) relationships are powerful indicators that can provide important clues toward understanding nutrient and sediment export dynamics from river systems, and the analysis of such relations has been a long-standing topic of importance in hydrologic literature. There are several issues, however, that can complicate or even mislead the interpretation of  $C$ - $Q$  relationships. These include non-linear  $\ln(C)$ - $\ln(Q)$  relations; variations in  $C$ - $Q$  relations over time and season; and different variations for different levels of discharge. In this context, the recently developed WRTDS method offers a means of addressing these concerns. Specifically, the  $\beta_2$  coefficients present a reasonably robust, informative, and accessible product of the WRTDS software that can provide new insights to the interpretation of river water-quality data, as demonstrated here with a case study for sediment concentration in Susquehanna River at Conowingo Dam. For this case, the approach was able to provide more nuanced and insightful evidence about previously reported long-term decline in

reservoir trapping performance. The case study also highlights an additional benefit of the method, which is the ability to perform uncertainty analyses. All of these benefits can be achieved by running additional codes within the *EGRET* package – such codes are made available to users through a DOI-referenced archive site that will be maintained for at least five years after publication (Zhang and Ball, 2016).

### **7.5. Supporting Information**

The modified *EGRET* functions, along with example data sets and *R* scripts that demonstrate the proposed approach, are documented at the publicly accessible Johns Hopkins University Data Archive via <http://dx.doi.org/10.7281/T18G8HM0> (Zhang and Ball, 2016).

### **7.6. Acknowledgements**

This work was supported by the Maryland Water Resources Research Center (2015MD329B), Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090), and National Science Foundation (CBET-1360415). We acknowledge Bob Hirsch, Doug Moyer, and Joel Blomquist (USGS) as well as Hugh Ellis and Dano Wilusz (JHU) for insightful discussions.

### **7.7. Literature Cited**

Basu, N. B., G. Destouni, J. W. Jawitz, S. E. Thompson, N. V. Loukinova, A. Darracq, S. Zanardo, M. Yaeger, M. Sivapalan, A. Rinaldo and P. S. C. Rao, 2010. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophys. Res. Lett.* 37:L23404, DOI: 10.1029/2010gl045168.

- Bieroza, M. Z. and A. L. Heathwaite, 2015. Seasonal variation in phosphorus concentration–discharge hysteresis inferred from high-frequency in situ monitoring. *Journal of Hydrology* 524:333-347, DOI: 10.1016/j.jhydrol.2015.02.036.
- Bowes, M. J., H. P. Jarvie, S. J. Halliday, R. A. Skef, A. J. Wade, M. Loewenthal, E. Gozzard, J. R. Newman, E. J. Palmer-felgate, R. A. Skeffington, A. J. Wade, M. Loewenthal, E. Gozzard, J. R. Newman and E. J. Palmer-felgate, 2015. Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration-flow relationships. *Sci. Total Environ.* 511:608-620, DOI: 10.1016/j.scitotenv.2014.12.086.
- Burt, T. P., F. Worrall, N. J. K. Howden and M. G. Anderson, 2015. Shifts in discharge-concentration relationships as a small catchment recover from severe drought. *Hydrol. Process.* 29:498-507, DOI: 10.1002/hyp.10169.
- Chanat, J. G., D. L. Moyer, J. D. Blomquist, K. E. Hyer and M. J. Langland, 2016. Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year 2012. U.S. Geological Survey Scientific Investigations Report 2015-5133, Reston, VA, p. 76. <http://dx.doi.org/10.3133/sir20155133>.
- Chanat, J. G., K. C. Rice and G. M. Hornberger, 2002. Consistency of patterns in concentration-discharge plots. *Water Resour. Res.* 38:1-10, DOI: 10.1029/2001WR000971.
- Cohn, T. A., 1995. Recent advances in statistical methods for the estimation of sediment

- and nutrient transport in rivers. *Rev. Geophys.* 33:1117-1117, DOI: 10.1029/95RG00292.
- Crowder, D. W., M. Demissie and M. Markus, 2007. The accuracy of sediment loads when log-transformation produces nonlinear sediment load–discharge relationships. *Journal of Hydrology* 336:250-268, DOI: 10.1016/j.jhydrol.2006.12.024.
- Evans, C. and T. D. Davies, 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. *Water Resour. Res.* 34:129-137, DOI: 10.1029/97WR01881.
- Friedrichs, C., T. Dillaha, J. Gray, R. Hirsch, A. Miller, D. Newburn, J. Pizzuto, L. Sanford, J. Testa, G. V. Houtven and P. Wilcock, 2014. Review of the Lower Susquehanna River Watershed Assessment. Chesapeake Bay Program Scientific and Technical Advisory Committee Report No. 14-006, Edgewater, Maryland, p. 40.
- Godsey, S. E., J. W. Kirchner and D. W. Clow, 2009. Concentration-discharge relationships reflect chemostatic characteristics of US catchments. *Hydrol. Process.* 23:1844-1864, DOI: 10.1002/hyp.7315.
- Gray, A. B., G. B. Pasternack, E. B. Watson, J. A. Warrick and M. A. Goñi, 2015. Effects of antecedent hydrologic conditions, time dependence, and climate cycles on the suspended sediment load of the Salinas River, California. *Journal of Hydrology* 525:632-649, DOI: 10.1016/j.jhydrol.2015.04.025.
- Gross, M. G., M. Karweit, W. B. Cronin and J. R. Schubel, 1978. Suspended sediment discharge of the Susquehanna River to northern Chesapeake Bay, 1966 to 1976.

*Estuaries* 1:106-110, <http://www.springerlink.com/index/r4796375t48061x3.pdf>.

Halliday, S. J., R. A. Skeffington, A. J. Wade, M. J. Bowes, E. Gozzard, J. R. Newman, M. Loewenthal, E. J. Palmer-Felgate and H. P. Jarvie, 2015. High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. *Hydrol. Process.* 10.1002/hyp.10453, DOI: 10.1002/hyp.10453.

Herndon, E. M., A. L. Dere, P. L. Sullivan, D. Norris, B. Reynolds and S. L. Brantley, 2015. Landscape heterogeneity drives contrasting concentration–discharge relationships in shale headwater catchments. *Hydrol. Earth Syst. Sci.* 19:3333-3347, DOI: 10.5194/hess-19-3333-2015.

Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.

Hirsch, R. M., 2014. Large Biases in Regression-Based Constituent Flux Estimates: Causes and Diagnostic Tools. *J. Am. Water Resour. Assoc.* 50:1401-1424, DOI: 10.1111/jawr.12195.

Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.

Hirsch, R. M. and L. De Cicco, 2015. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0,

- February 2015). U.S. Geological Survey Techniques and Methods Book 4, Chapter A10, Reston, VA, p. 93. <http://dx.doi.org/10.3133/tm4A10>.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Horowitz, A. J., 2003. An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrol. Process.* 17:3387-3409, DOI: 10.1002/hyp.1299.
- House, W. A. and M. S. Warwick, 1998. Hysteresis of the solute concentration/discharge relationship in rivers during storms. *Water Res.* 32:2279-2290, DOI: 10.1016/S0043-1354(97)00473-9.
- Kutner, M., C. Nachtsheim and J. Neter, 2004. *Applied Linear Statistical Models*, McGraw-Hill Education, ISBN 0073014664.
- Lang, D. J., 1982. Water quality of the three major tributaries to the Chesapeake Bay, the Susquehanna, Potomac, and James Rivers, January 1979-April 1981. US Geological Survey, Water Resources Division Water-Resources Investigations Report 82-32, p. 64. <https://pubs.er.usgs.gov/publication/wri8232>.
- Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18. <http://dx.doi.org/10.3133/ofr20141235>.
- Meybeck, M. and F. Moatar, 2012. Daily variability of river concentrations and fluxes:

- Indicators based on the segmentation of the rating curve. *Hydrol. Process.* 26:1188-1207, DOI: 10.1002/hyp.8211.
- Miller, M. P., A. J. Tesoriero, P. D. Capel, B. A. Pellerin, K. E. Hyer and D. A. Burns, 2016. Quantifying watershed-scale groundwater loading and in-stream fate of nitrate using high-frequency water quality data. *Water Resour. Res.* 52:330-347, DOI: 10.1002/2015WR017753.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Musolff, A., C. Schmidt, B. Selle and J. H. Fleckenstein, 2015. Catchment controls on solute export. *Adv. Water Resour.* 86:133-146, DOI: 10.1016/j.advwatres.2015.09.026.
- Outram, F. N., C. E. M. Lloyd, J. Jonczyk, C. M. H. Benskin, F. Grant, M. T. Perks, C. Deasy, S. P. Burke, A. L. Collins, J. Freer, P. M. Haygarth, K. M. Hiscock, P. J. Johnes and A. L. Lovett, 2014. High-frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of the 2011–2012 drought in England. *Hydrol. Earth Syst. Sci.* 18:3429-3448, DOI: 10.5194/hess-18-3429-2014.
- Pellerin, B. A., B. A. Bergamaschi, R. J. Gilliom, C. G. Crawford, J. Saraceno, C. P. Frederick, B. D. Downing and J. C. Murphy, 2014. High frequency measurement of nitrate concentration in the Lower Mississippi River, USA. *Environ. Sci. Technol.* 48:12612-12619, DOI: 10.1021/es504029c.

- R Development Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3900051070. <http://www.r-project.org>.
- Richardson, M. C., 2012. Contributions of streamflow variability, concentration-discharge shifts and forested wetlands to terrestrial-aquatic solute export in Precambrian Shield headwater catchments. *Ecohydrology* 5:596-612, DOI: 10.1002/eco.244.
- Shanley, J. B., W. H. McDowell and R. F. Stallard, 2011. Long-term patterns and short-term dynamics of stream solutes and suspended sediment in a rapidly weathering tropical watershed. *Water Resour. Res.* 47:W07515, DOI: 10.1029/2010WR009788.
- Stallard, R. F. and S. F. Murphy, 2014. A Unified Assessment of Hydrologic and Biogeochemical Responses in Research Watersheds in Eastern Puerto Rico Using Runoff-Concentration Relations. *Aquat. Geochem.* 20:115-139, DOI: 10.1007/s10498-013-9216-5.
- The Lower Susquehanna River Watershed Assessment Team, 2015. Lower Susquehanna River Watershed Assessment, Maryland and Pennsylvania. p. 192. <http://dnr.maryland.gov/bay/lswa/report.htm>.
- U.S. Geological Survey, 2014. Surface-water data for the nation.
- Walling, D. E., 1977. Assessing the accuracy of suspended sediment rating curves for a small basin. *Water Resour. Res.* 13:531-538, DOI: 10.1029/WR013i003p00531.
- Warrick, J. A., 2014. Trend analyses with river sediment rating curves. *Hydrol. Process.* 949:936-949, DOI: 10.1002/hyp.10198.

- Zhang, Q. and W. P. Ball, 2016. Data associated with Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-Quality Data. Johns Hopkins University Data Archive, Baltimore, MD.  
<http://dx.doi.org/10.7281/T18G8HM0>, DOI: 10.7281/T18G8HM0.
- Zhang, Q., W. P. Ball and D. L. Moyer, 2016a. Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns. *Sci. Total Environ.* 563-564: 1016-1029, DOI: 10.1016/j.scitotenv.2016.03.104.
- Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI: 10.1016/j.scitotenv.2013.02.012.
- Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.
- Zhang, Q., R. M. Hirsch and W. P. Ball, 2016b. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* 50:1877-1886, DOI: 10.1021/acs.est.5b04073.

## Chapter 8. Non-stationary Concentration-Discharge Relationships: A Synthesis of Nutrient and Sediment Patterns in the Major Tributaries to Chesapeake Bay<sup>20</sup>

### Abstract

Derived from river monitoring data, concentration-discharge ( $C-Q$ ) relationships are indicators of export dynamics. Here we provide a data-driven, top-down synthesis of  $C-Q$  patterns in nine major tributaries of Chesapeake Bay for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN). The recently-developed WRTDS (Weighted Regressions on Time, Discharge, and Season) method was adopted to make robust interpretation of  $C-Q$  relationships. Unlike many previous  $C-Q$  studies that focused on stormflow conditions, this approach allows simultaneous examination of various discharge conditions within an uncertainty framework. Our synthesis has offered several new insights on the complexity of watershed function (*i.e.*, concentration sensitivity to discharge). First, constituent export has been dominated by mobilization patterns for SS and TP (particulate-dominated species) and chemostasis patterns for TN (dissolved-dominated species) under many discharge conditions. The general lack of dilution patterns may suggest that none of the three constituents has been supply-limited in these watersheds. In addition, SS and TP coefficients have followed a clear positive monotonic pattern with respect to discharge, exhibiting threshold behaviors (*i.e.*, chemostasis at low discharge *vs.* mobilization at high discharge). Moreover, for many site-constituent

---

<sup>20</sup> This chapter will be submitted for publication in a peer-reviewed journal. Bill Ball was involved in hypothesis development, results interpretation, and editing. All figures, tables, and data were created by Qian Zhang.

combinations, coefficients show clear temporal non-stationarity in  $C$ - $Q$  relationships under selected fixed discharges, possibly reflecting changes in dominant watershed sources due to management actions. These results highlight the potential pitfalls of assuming fixed  $C$ - $Q$  relationships over discharge or time. The WRTDS coefficients provide an alternative approach for interpretation of water-quality data and for generation of sensible hypotheses on dominant processes in different watersheds. More generally, the approach is also adaptable to other river systems and our synthesis effectively illustrates the value of long-term data collection at multiple locations in the watersheds.

### **8.1. Introduction**

Derived from river monitoring data, concentration-discharge ( $C$ - $Q$ ) relationships are a powerful tool for understanding complex interactions between hydrologic and biogeochemical processes, including nutrient and sediment export dynamics (Evans and Davies, 1998; Chanut *et al.*, 2002; Godsey *et al.*, 2009; Meybeck and Moatar, 2012). In particular,  $C$ - $Q$  relationships have been commonly classified into three broad categories – namely, (1) “dilution” (*i.e.*, negative relationship); (2) “mobilization” (*i.e.*, positive relationship); and (3) “chemostasis” (*i.e.*,  $C$  invariant with  $Q$ ). Such relationships have often been found to vary with constituent and with site (Godsey *et al.*, 2009; Hirsch *et al.*, 2010; Meybeck and Moatar, 2012; Stallard and Murphy, 2014; Herndon *et al.*, 2015). In general, these patterns are largely controlled by the spatial availability and distribution of constituent sources in the watershed as well as their hydrological connectivity to the stream. Particularly, dilution responses can occur when anthropogenic point sources (*e.g.*, wastewater treatment plants) or other spatially distinct and flow-independent sources (*e.g.*, mineral dissolution from base-flow pathways) are dominant and are more

concentrated than nonpoint sources in the watershed. Mobilization responses can occur when otherwise disconnected solute or sediment sources become connected to water flow paths during elevated discharges. These conceptualizations have formed the foundations to the development of component mixing models for interpretation of event-scale concentration data in terms of contributions from deep subsurface, shallow subsurface, and surface water sources (Evans and Davies, 1998; Chanat *et al.*, 2002; Bieroza and Heathwaite, 2015) and riverine loading apportionment models for analysis of decadal-scale records (Bowes *et al.*, 2008; Bowes *et al.*, 2009). As a relatively less familiar concept, chemostasis has been recently documented for nutrients and weathering products in a range of watersheds and has been attributed to constant fluxes of release from legacy stores that have been accumulated historically from sources such as agricultural input, atmospheric deposition, and mineral formation and deposition (Godsey *et al.*, 2009; Basu *et al.*, 2010; Basu *et al.*, 2011; Thompson *et al.*, 2011; Herndon *et al.*, 2015).

While many prior studies of  $C$ - $Q$  relationships focus on the interpretation of event-scale data, particularly storm hysteresis, and the development of component mixing models for inferring source water (Evans and Davies, 1998; House and Warwick, 1998; Outram *et al.*, 2014; Bieroza and Heathwaite, 2015; Burt *et al.*, 2015), we believe that decadal-scale low-frequency data also have merits. In particular, such long-term data can reveal temporal changes in  $C$ - $Q$  relationships, which in turn may reflect anthropogenic activities such as land disturbance and watershed management (Bieroza and Heathwaite, 2015; Burt *et al.*, 2015; Gray *et al.*, 2015; Zhang *et al.*, 2016c). In fact, several recent studies have focused on the synthesis of long-term data at multiple sites (Godsey *et al.*,

2009; Basu *et al.*, 2010; Thompson *et al.*, 2011). A common feature of these analyses is to study patterns using long-term monitoring data from multiple watersheds for parsimonious representation of dominant watershed processes. This type of approach has been referred to as “top-down” approach, as opposed to “bottom-up” approach that employs sophisticated and distributed models that capture all relevant processes and parameters (Basu *et al.*, 2011).

While the adoption of log-linear  $C$ - $Q$  relationships (or its modified form, loading-discharge relationships) has been a popular practice in the hydrological literature, there are several issues noted with this approach that can complicate or even mislead the interpretation. These include non-linear  $\ln(C)$ - $\ln(Q)$  relations; variations in  $C$ - $Q$  relations over time and season; and different variations for different levels of discharge. Zhang *et al.* (2016b) – and the references therein – has thoroughly discussed these issues with real water-quality data. In this regard, Zhang *et al.* (2016b) proposes use of the recently-developed WRTDS (“Weighted Regressions on Time, Discharge, and Season”) method (Hirsch *et al.*, 2010) as an alternative and improved approach. The authors have also provided examples to illustrate how the WRTDS model’s coefficients (which were previously hidden from user view) can be accessed, organized, and presented in various ways to provide improved interpretation of riverine  $C$ - $Q$  relationships.

In the Chesapeake Bay watershed, reduction of total nitrogen (TN), total phosphorus (TP), and suspended sediment (SS) loadings has long been a management focus toward controlling Bay eutrophication and hypoxia (Kemp *et al.*, 2005; Murphy *et al.*, 2011; Shenk and Linker, 2013). Better understanding of nutrient and sediment export from different regions of the watershed is critically needed for assessing historical

management progress and for shaping future strategies. In this context, the overall objective of this work was to apply the WRTDS approach to better interpret the nature of and change in  $C$ - $Q$  relationships for major constituents (*i.e.*, SS, TP, and TN) at multiple long-term sites across the Chesapeake watershed (Figure 8.1). In this synthesis, we seek answers to the following questions:

- (1) How does the nature of  $C$ - $Q$  relationships (*i.e.*, dilution, chemostasis, and mobilization) vary by discharge condition and how do the patterns compare among sites and species?
- (2) Are  $C$ - $Q$  relationships temporally varying (*i.e.*, non-stationary) under different discharge conditions and how do the patterns compare among sites and species?

To our knowledge, this is the first top-down analysis of  $C$ - $Q$  patterns in these major Chesapeake tributaries. This analysis has benefited from the high-quality long-term monitoring data at multiple locations across the Chesapeake Bay watershed, which represent diverse characteristics in terms of land use, physiography, and hydrological settings. More broadly, our approach can be adapted to river systems elsewhere for understanding patterns and controls of  $C$ - $Q$  relationships in traditional low-frequency river water-quality monitoring data.

## **8.2. Methods**

### **8.2.1. Study Sites**

This work focuses on nine major non-tidal tributaries to Chesapeake Bay, namely, Susquehanna, Potomac, James, Rappahannock, Appomattox, Pamunkey, Mattaponi, Patuxent, and Choptank (Figure 8.1; Table 8.1). Since the 1980s, these rivers have been

monitored at their fall-line (divide of tidal and non-tidal areas) locations by the U.S. Geological Survey (USGS) River Input Monitoring (RIM) Program. Collectively, these nine sites account for ~93% of non-tidal discharge and ~77% of total freshwater discharge to Chesapeake Bay between 1991 and 2000 (Shenk and Linker, 2013). The Choptank is located entirely in the coastal plain and may represent be representative of the much larger Eastern Shore area. On the Western Shore, only Mattaponi draws a substantial portion of its water from coastal plain areas. The other seven tributaries are dominated by upland physiographic provinces, namely, piedmont, Blue Ridge, valley and ridge, and Appalachian plateau (Shenk and Linker, 2013).

Among the nine RIM tributaries, the Susquehanna is the largest, and contributed ~62% of river discharge, ~65% of TN load, ~46% of TP load, and ~41% of SS load to the Bay between 1979 and 2012 (Zhang *et al.*, 2015). The relatively lower fractional contributions of TP and SS likely reflect historical retention within the Lower Susquehanna River Reservoir System (LSRRS). The largest and most-downstream member of the LSRRS, Conowingo Reservoir, is reportedly over 90% full in terms of sediment storage (Langland, 2015), accompanied by substantial recent decline in net trapping of sediment and particulate nutrients (Hirsch, 2012; Zhang *et al.*, 2013; Zhang *et al.*, 2016c). Above the LSRRS, six sites have been monitored by the Susquehanna River Basin Commission (SRBC) since the 1980s, with three of them (*i.e.*, Towanda, Danville, and Marietta) on the main-stem of Susquehanna and the other three (*i.e.*, Lewisburg, Newport, and Conestoga) on tributaries to Susquehanna (Figure 8.1; Table 8.1).

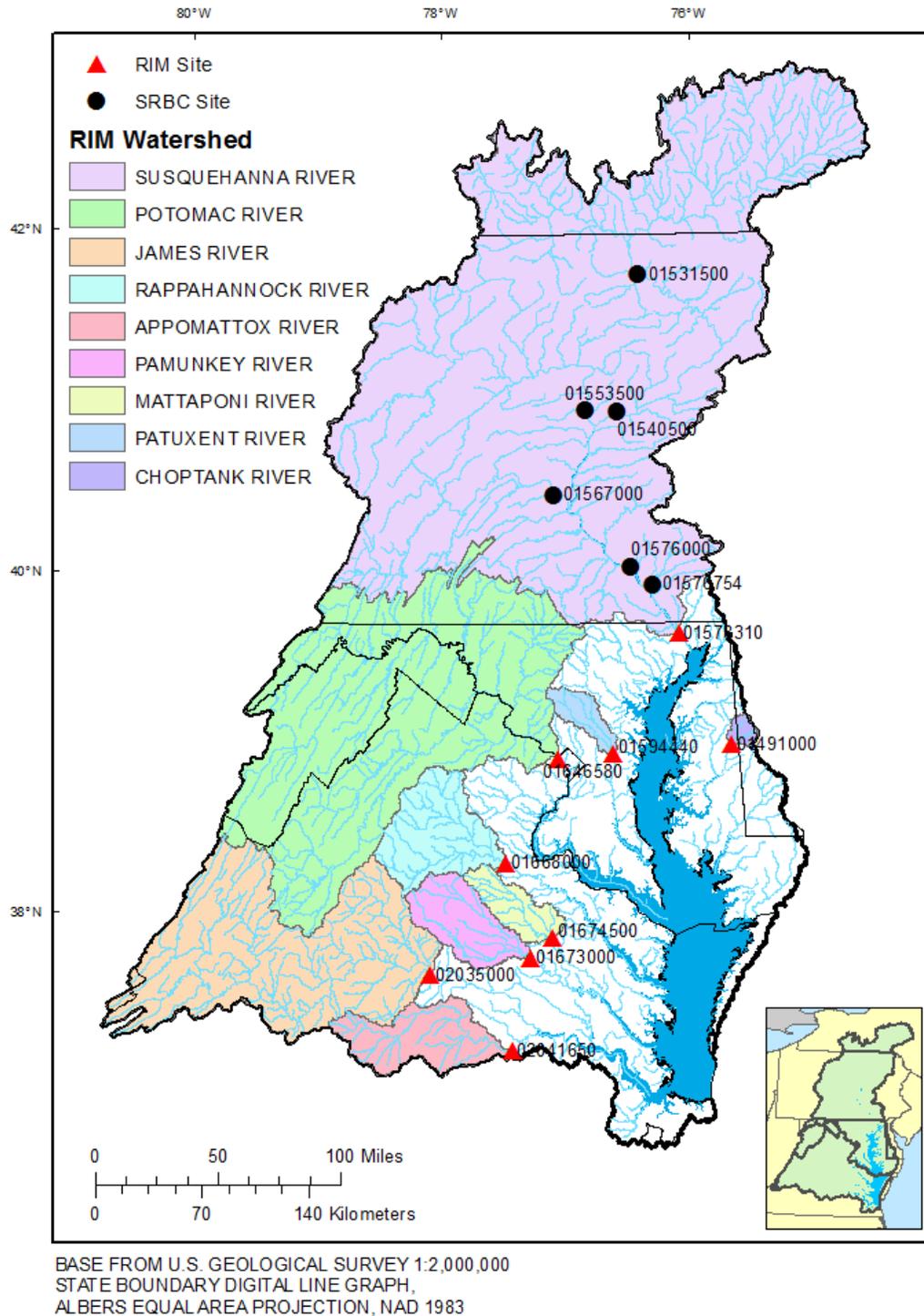


Figure 8.1. Chesapeake Bay watershed and the 15 monitoring sites that include nine River Input Monitoring (RIM) sites on the fall-line of nine major tributaries and six Susquehanna River Basin Commission (SRBC) sites at upstream locations within the Susquehanna River basin. This figure was modified after Figure 1 in Moyer *et al.* (2012).

Table 8.1. Details of the 15 long-term monitoring sites in the Chesapeake Bay watershed. <sup>a</sup>

Station Number	River sites	Drainage area, km <sup>2</sup>	Annual river flow in 1984-2014		Upstream land use (percent)				
			Average flow, m <sup>3</sup> /s	Average yield, m/yr	Urban	Agricultural	Forested	Other	
RIM Sites	01578310	Susquehanna River near Conowingo, MD	70,189	1147	0.52	2	29	67	2
	01646580	Potomac River at Chain Bridge, Washington D.C.	30,044	338	0.35	3	35	61	1
	02035000	James River at Cartersville, VA	16,213	199	0.39	1	16	80	3
	01668000	Rappahannock River near Fredericksburg, VA	4,144	49	0.38	1	36	61	2
	02041650	Appomattox River at Matoaca, VA	3,471	33	0.30	1	20	72	7
	01673000	Pamunkey River near Hanover, VA	2,800	28	0.31	1	24	68	7
	01674500	Mattaponi near Beulahville, VA	1,557	15	0.30	1	19	69	11
	01594440	Patuxent River at Bowie, MD	901	11	0.38	13	41	38	8
	01491000	Choptank River near Greensboro, MD	293	4.2	0.45	1	50	29	20
SRBC Sites	01576000	Susquehanna River at Marietta, PA	67,314	1114	0.52	4	30	64	2
	01540500	Susquehanna River at Danville, PA	29,008	475	0.52	5	33	60	2
	01531500	Susquehanna River at Towanda, PA	20,194	325	0.51	4	35	60	1
	01553500	West Branch Susquehanna River at Lewisburg, PA	17,765	310	0.55	2	15	81	2
	01567000	Juniata River at Newport, PA	8,687	126	0.46	2	28	69	1
	01576754	Conestoga River at Conestoga, PA	1,217	19	0.50	8	54	37	1

<sup>a</sup> modified from Table 3 and Table 8 in Sprague *et al.* (2000)

### 8.2.2. Monitoring Data

At each of the 15 sites, daily discharge data were compiled from the USGS National Water Information System (NWIS) (U.S. Geological Survey, 2014). In addition, SS, TP, and TN concentration data were compiled from NWIS for the nine RIM sites (U.S. Geological Survey, 2014) and from SRBC for the six SRBC sites (Susquehanna River Basin Commission, 2014). These 15 sites are among the most densely sampled long-term stations within the Chesapeake Bay Nontidal Water-Quality Monitoring Network (Chanat *et al.*, 2016). Temporal coverages of water-quality samples are summarized in Table 8.2. The average number of days sampled varies with sites, ranging between 12.6–39.4 days/year (median = 25.7) for SS, 20.8–40.4 days/year (median = 28.3) for TP, and 20.8–39.4 days/year (median = 27.6) for TN.

In general, water-quality concentration samples at each site were collected across the full range of hydrological conditions in each year and comprised at least one sample in each calendar month and eight targeted samples during times of stormflow (Chanat *et al.*, 2016; Zhang *et al.*, 2016a). Consequently, these sites have been sampled at least 20 days per year (Table 8.2). The only exceptions are SS sampling at Potomac (Maryland) and all five RIM rivers in Virginia, where only 12.6–15.8 days per year were sampled on average for two reasons: (a) stormflow sampling was not implemented in the early years of the record, and (b) SS sampling was discontinued at the Virginia sites in the late 1990s. To further examine the representativeness of water-quality sampling with respect to flow conditions, distributions of discharge on days with water-quality samples and discharge on all days in the record were compared at each individual site. Results show that SS (Figure F1), TP (Figure F2), and TN (Figure F3) have been sampled at each site

Table 8.2. Temporal coverage of observed water-quality data at the 15 Chesapeake sites.

(T<sub>Start</sub>: first sampled day; T<sub>End</sub>: last sampled day; #<sub>Sampled</sub>: total number of sampled days; f<sub>Sampling</sub>: average number of sampled days per year.)

River sites	Suspended Sediment (SS)			Total Phosphorus (TP)			Total Nitrogen (TN)			
	T <sub>Start</sub>	T <sub>End</sub>	# <sub>Sampled</sub> (f <sub>Sampling</sub> )	T <sub>Start</sub>	T <sub>Start</sub>	T <sub>End</sub>	T <sub>Start</sub>	T <sub>End</sub>	# <sub>Sampled</sub> (f <sub>Sampling</sub> )	
USGS RIM Sites	Susquehanna River near Conowingo, MD	1984/10/25	2014/9/3	799 (26.6/yr)	1984/10/25	2014/9/3	808 (26.9/yr)	1984/10/25	2014/9/3	804 (26.8/yr)
	Potomac River at Chain Bridge, Washington D.C.	1984/11/14	2014/9/9	463 (15.4/yr)	1984/10/9	2014/9/9	1191 (39.7/yr)	1984/10/2	2014/9/9	1183 (39.4/yr)
	James River at Cartersville, VA	1984/10/30	2014/9/2	475 (15.8/yr)	1984/10/30	2014/9/2	814 (27.1/yr)	1984/10/30	2014/9/2	811 (27.0/yr)
	Rappahannock River near Fredericksburg, VA	1984/10/9	2014/9/9	381 (12.7/yr)	1984/10/9	2014/9/9	763 (25.4/yr)	1984/10/9	2014/9/9	756 (25.2/yr)
	Appomattox River at Matoaca, VA	1984/11/20	2014/9/4	385 (12.8/yr)	1984/11/20	2014/9/4	779 (26.0/yr)	1984/11/20	2014/9/4	773 (25.8/yr)
	Pamunkey River near Hanover, VA	1984/10/17	2014/9/16	386 (12.9/yr)	1984/10/17	2014/9/16	831 (27.7/yr)	1984/10/17	2014/9/16	828 (27.6/yr)
	Mattaponi near Beulahville, VA	1984/10/17	2014/9/30	378 (12.6/yr)	1984/10/17	2014/9/30	819 (27.3/yr)	1984/10/17	2014/9/30	813 (27.1/yr)
	Patuxent River at Bowie, MD	1984/11/28	2014/9/26	771 (25.7/yr)	1984/10/24	2014/9/26	849 (28.3/yr)	1984/10/24	2014/9/26	771 (25.7/yr)
	Choptank River near Greensboro, MD	1984/10/19	2014/9/25	613 (20.4/yr)	1984/10/19	2014/9/25	625 (20.8/yr)	1984/10/19	2014/9/25	624 (20.8/yr)
SRBC Sites	Susquehanna River at Marietta, PA	1986/10/7	2014/9/29	1037 (37.0/yr)	1986/10/7	2014/9/29	1044 (37.3/yr)	1986/10/7	2014/9/29	1043 (37.3/yr)
	Susquehanna River at Danville, PA	1984/10/11	2014/9/30	1183 (39.4/yr)	1984/10/11	2014/9/30	1213 (40.4/yr)	1984/10/11	2014/9/30	1082 (36.1/yr)
	Susquehanna River at Towanda, PA	1988/10/5	2014/9/15	990 (38.0/yr)	1984/10/10	2014/9/15	1061 (35.4/yr)	1984/10/10	2014/9/15	931 (31.0/yr)
	West Branch Susquehanna River at Lewisburg, PA	1984/10/11	2014/9/30	1130 (37.7/yr)	1984/10/11	2014/9/30	1165 (38.8/yr)	1984/10/11	2014/9/30	1031 (34.4/yr)
	Juniata River at Newport, PA	1984/10/10	2014/9/17	1006 (33.5/yr)	1984/10/10	2014/9/17	1050 (35.0/yr)	1984/10/10	2014/9/17	914 (30.5/yr)
	Conestoga River at Conestoga, PA	1984/10/18	2014/9/29	1076 (35.9/yr)	1984/10/18	2014/9/29	1065 (35.5/yr)	1984/10/18	2014/9/29	1014 (33.8/yr)

with generally good coverage of high-discharge conditions. In fact, water-quality samples have been collected with a median discharge that tends to bias toward the higher end.

### 8.2.3. WRTDS Method

The WRTDS method (Hirsch *et al.*, 2010) was applied to analyze  $C$ - $Q$  relationships. This method was originally developed to improve the statistical estimation of daily concentrations and loadings based on low-frequency concentration samples. It uses time, season, and discharge as model explanatory variables:

$$\ln(C_i) = \beta_{0,i} + \beta_{1,i}t_i + \beta_{2,i} \ln(Q_i) + \beta_{3,i} \sin(2\pi t_i) + \beta_{4,i} \cos(2\pi t_i) + \varepsilon_i \quad (8.1)$$

where  $t_i$  is time in decimal years,  $C_i$  is daily concentration at time  $t_i$ ,  $Q_i$  is daily discharge at time  $t_i$ ,  $\beta_{0,i} \sim \beta_{4,i}$  are fitted coefficients, and  $\varepsilon_i$  is the error term. The 1<sup>st</sup> - 3<sup>rd</sup> terms on the right of Equation (8.1) represent the intercept, time effects, and discharge effects, respectively, whereas the 4<sup>th</sup> and 5<sup>th</sup> terms collectively represent cyclical seasonal effects. For each day of estimation, WRTDS pre-screens all available concentration samples and selects the most relevant samples to fit Equation (8.1), with the “relevancy” being quantified and compared on three dimensions, *i.e.*, time, discharge, and season. The fitted coefficients are used to estimate  $\ln(C_i)$  on the estimation day with known values of  $t_i$  and  $Q_i$ . The estimation process is fully described in Hirsch and De Cicco (2015).

To expedite the estimation process, WRTDS establishes a set of evenly-spaced grid points on a surface defined by  $t$  and  $\ln(Q)$ . The standard grid design is fully described in the user manual (Hirsch and De Cicco, 2015). Briefly, for the  $x$ -axis, time grid values are spaced 0.0625 years (1/16<sup>th</sup> of a year) apart from the beginning year to the end year of the record. For the  $y$ -axis, 14 grid values are spaced with equal distance in log space for the

discharge range from 5 percent below the minimum  $Q$  to 5 percent above the maximum  $Q$  in the record.

For interpretation of  $C$ - $Q$  relationships, Zhang *et al.* (2016b) recommended that WRTDS  $\beta_2$  coefficients can offer an effective solution to those issues listed above:

- (1) It does not assume a strictly linear  $\ln(C) \sim \ln(Q)$  relation. Instead, it conducts “local” fitting at many points in the time-discharge space in a consistent manner.
- (2) It allows the  $C$ - $Q$  relation to vary with time, discharge, and season.
- (3) It avoids the specification of any smoothing window or polynomial degree.
- (4) It avoids making subjective choices on separated intervals for time, discharge, or season.
- (5) It can, at least partially, decouple the interactions among time, discharge, and season.
- (6) It is less sensitive to the quality and quantity of typically sparse high-flow samples in the record, as compared with other approaches.

#### **8.2.4. Data Analysis**

For each constituent at each site, I implemented WRTDS using the *EGRET* (Exploration and Graphics for RivEr Trends) package version 2.2.0 (Hirsch and De Cicco, 2015) in R 3.1.0 (R Development Core Team, 2014). Some necessary modifications of the existing *EGRET* functions were made to extract and visualize the estimated  $\beta_2$  coefficients, which are made available to users through the Johns Hopkins University Data Archive (Zhang and Ball, 2016). In brief, the *EGRET* functions were modified to run the same regression model (*i.e.*, Equation 8.1) on the same  $t$ - $Q$  grid but to extract and store  $\beta_2$  coefficients instead of concentration estimates. The estimated  $\beta_2$

coefficients are then used to categorize export patterns – namely, (1) “dilution” (*i.e.*,  $\beta_2 < 0$ ); (2) “chemostasis” (*i.e.*,  $\beta_2 = 0$ ); and (3) “mobilization” (*i.e.*,  $\beta_2 > 0$ ). For illustration, the estimated  $\beta_2$  coefficients for TP in Susquehanna River at Conowingo, MD, are shown in Figure 8.2 as a contour plot against axes of  $t$  and  $\ln(Q)$ .

To obtain uncertainty estimates on  $\beta_2$  coefficients for each constituent at each site, we have adopted the block-bootstrap method of Hirsch *et al.* (2015) to re-sample (with replacement) the original concentration data to obtain 50 realizations of representative sets and re-run the model estimation with each replicate (Hirsch *et al.*, 2015). The resulting 50 sets of  $\beta_2$  coefficients were used to calculate 90% confidence intervals,

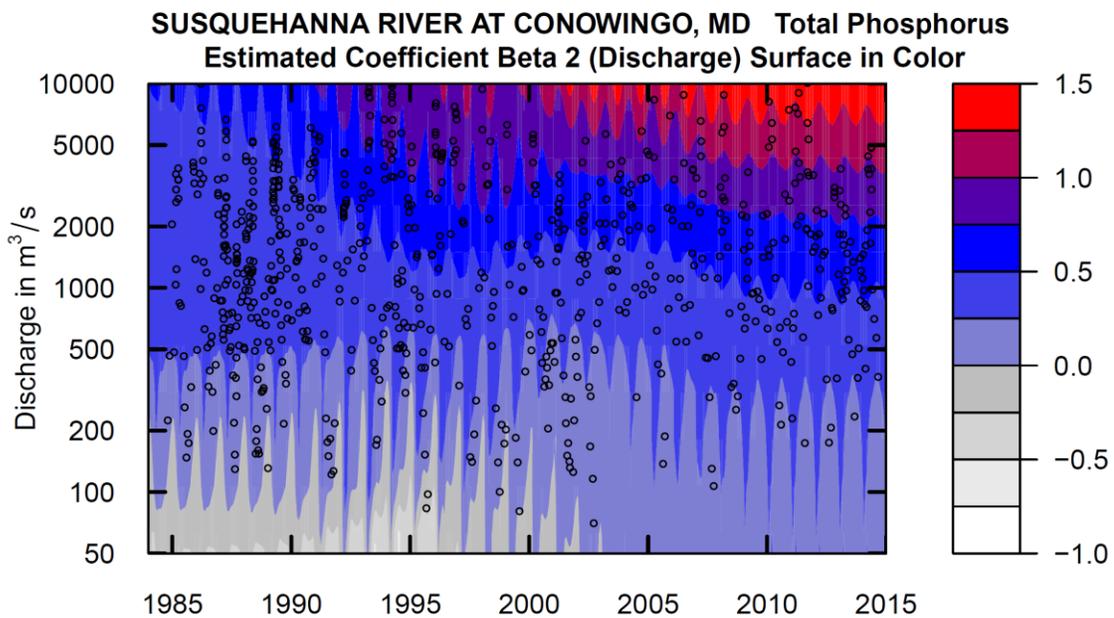


Figure 8.2. Contour plot showing estimated WRTDS  $\beta_2$  coefficients as a function of time and discharge for total phosphorus in Susquehanna River at Conowingo, MD. Black open circles indicate the time-discharge combinations where concentration samples have been taken. The  $\beta_2$  coefficients correspond to three broad categories, namely, (1) dilution (*i.e.*,  $\beta_2 < 0$ ); (2) chemostasis ( $\beta_2 \approx 0$ ); and (3) mobilization ( $\beta_2 > 0$ ).

which are shown as dashed lines on subsequent plots. This approach is more fully described in Zhang *et al.* (2016b).

Moreover, to better understand spatial and temporal patterns of  $\beta_2$  coefficients, I compiled and analyzed watershed source input data in the drainage basins above each site, which were made available to us by the Chesapeake Bay Program Office (Shenk and Linker, 2013). These data include atmospheric deposition, fertilizer, manure, and point-source contributions from each drainage basin between 1984 and 2011 – see Zhang *et al.* (2016a) for data details.

### **8.3. Results**

#### **8.3.1. Changes in Coefficients with Discharge**

Here we address the first question posited above: *How does the nature of C-Q relationships (i.e., dilution, chemostasis, and mobilization) vary by discharge condition and how do the patterns compare among sites and species?* In this analysis, daily estimates of  $\beta_2$  coefficients are grouped by discharge percentiles to reveal discharge-related patterns. Results are shown in Figures 8.3-8.5 for SS, TP, and TN, respectively, and are elaborated below.

##### **8.3.1.1. SS**

At the 15 sites, SS coefficients (Figure 8.3) show predominantly mobilization effects (*i.e.*,  $\beta_2 > 0$ ) across all discharge intervals. The few exceptions include several sites at the lowest discharge interval (*i.e.*, Towanda, Lewisburg, Conowingo, Rappahannock, Appomattox, Mattaponi, and Choptank) and one site at the highest discharge interval

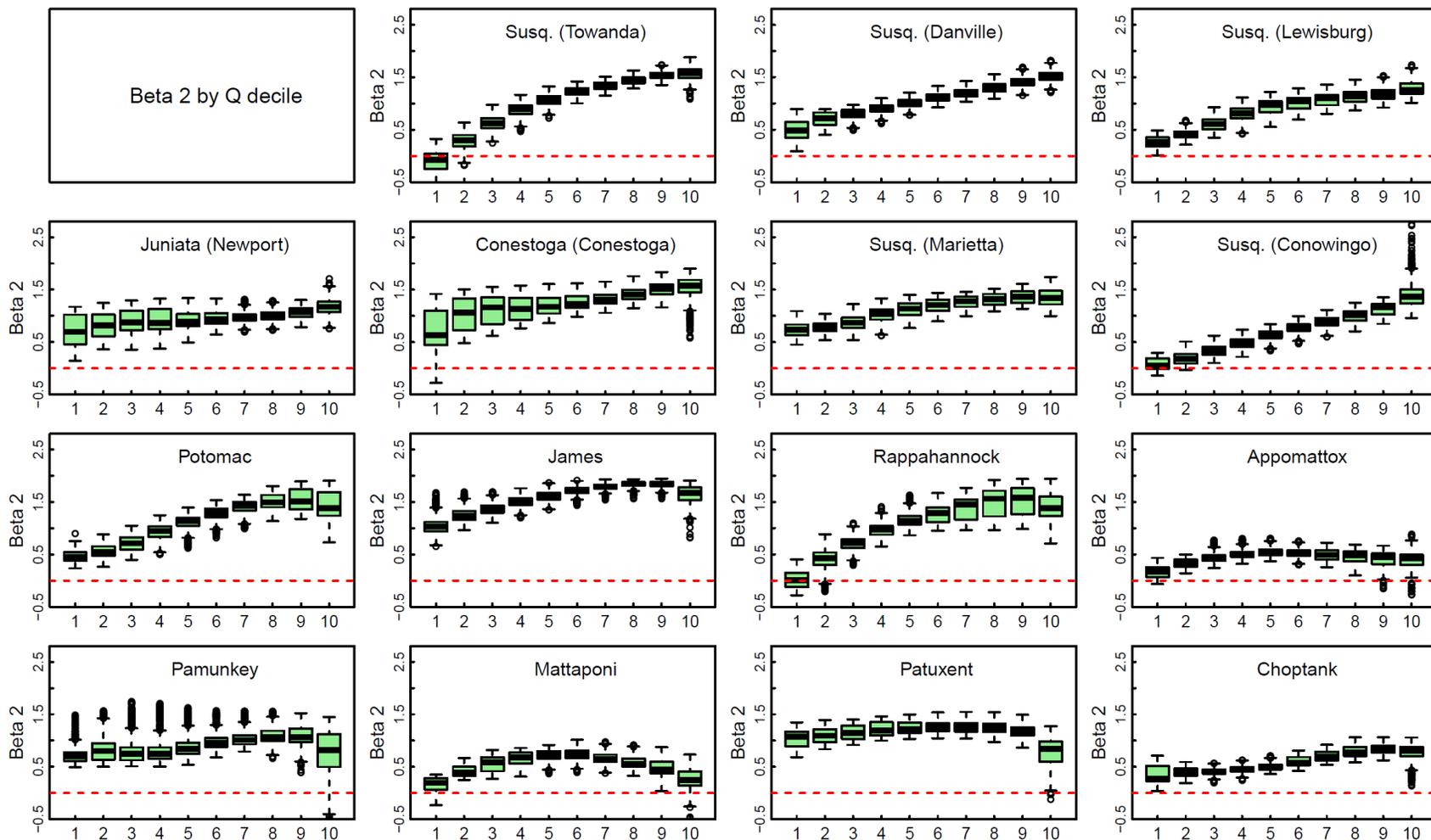


Figure 8.3. Boxplot summary of estimated WRTDS  $\beta_2$  coefficients by discharge decile for suspended sediment (SS) at the 15 Chesapeake sites. (X-axis: 1 = 0<sup>th</sup>~10<sup>th</sup>, 2 = 10<sup>th</sup>~20<sup>th</sup> ..., 9 = 80<sup>th</sup>~90<sup>th</sup>, 10 = 90<sup>th</sup>~100<sup>th</sup>.)

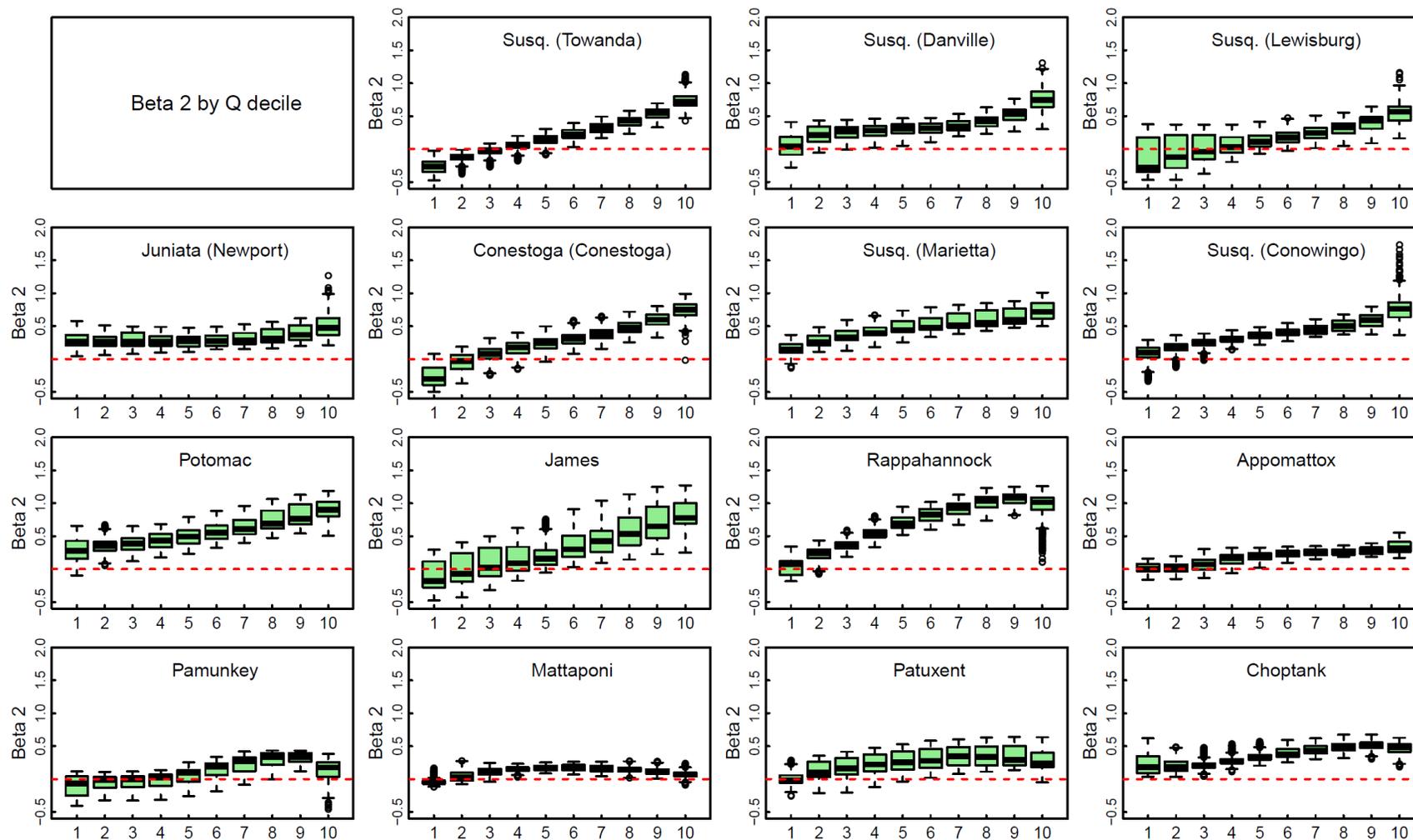


Figure 8.4. Boxplot summary of estimated WRTDS  $\beta_2$  coefficients by discharge decile for total phosphorus (TP) at the 15 Chesapeake sites. (X-axis: 1 = 0<sup>th</sup>~10<sup>th</sup>, 2 = 10<sup>th</sup>~20<sup>th</sup>..., 9 = 80<sup>th</sup>~90<sup>th</sup>, 10 = 90<sup>th</sup>~100<sup>th</sup>.)

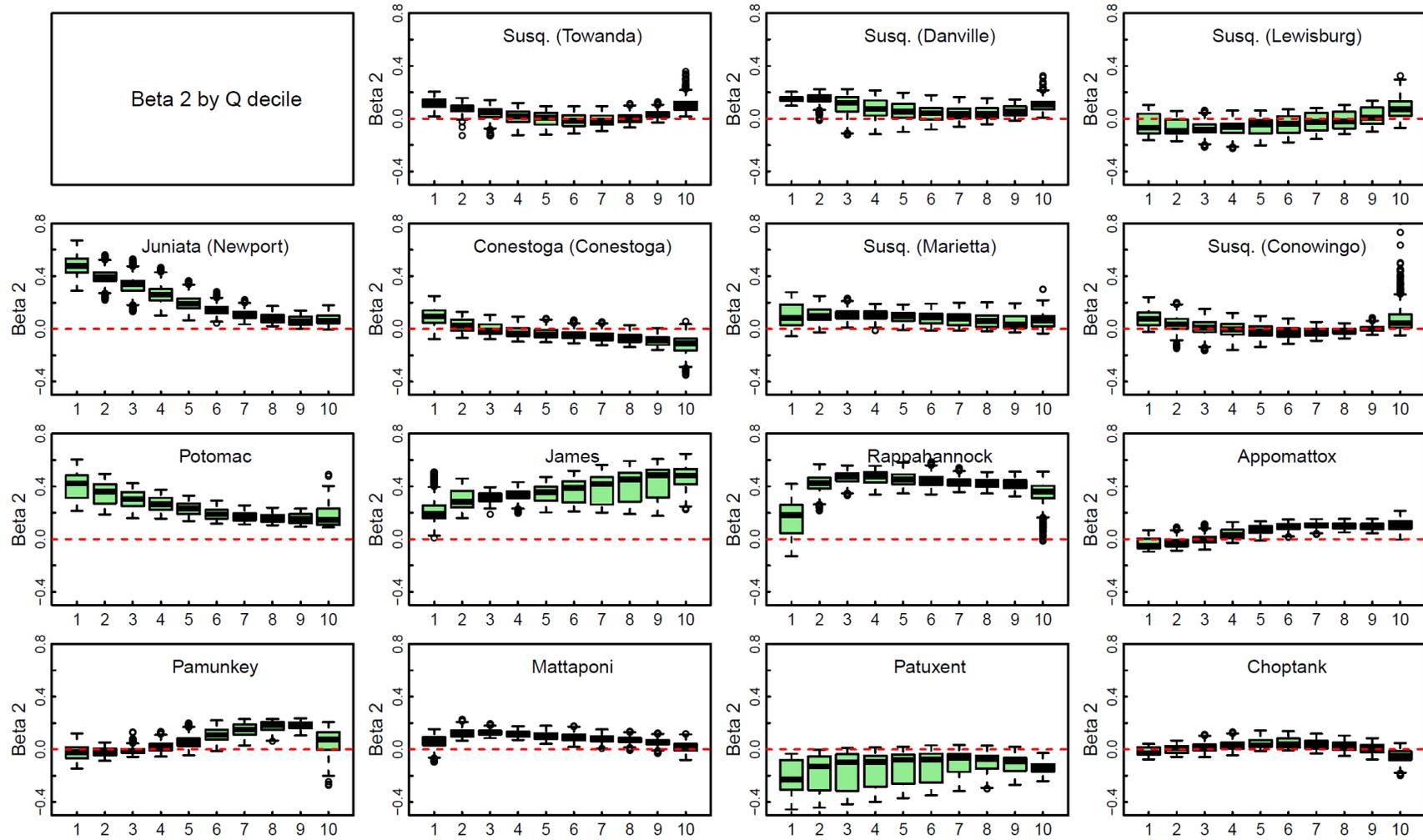


Figure 8.5. Boxplot summary of estimated WRTDS  $\beta_2$  coefficients by discharge decile for total nitrogen (TN) data at the 15 Chesapeake sites. (X-axis: 1 = 0<sup>th</sup>~10<sup>th</sup>, 2 = 10<sup>th</sup>~20<sup>th</sup>..., 9 = 80<sup>th</sup>~90<sup>th</sup>, 10 = 90<sup>th</sup>~100<sup>th</sup>.)

(*i.e.*, Mattaponi). In these cases, the coefficients are closer to zero, thereby exhibiting chemostasis ( $\beta_2 \approx 0$ ) or even dilution ( $\beta_2 < 0$ ).

At most sites, SS coefficients (Figure 8.3) follow a positive monotonic pattern with respect to discharge, with highest values occurring at the highest discharge interval (*i.e.*, 90<sup>th</sup>~100<sup>th</sup> percentile) and lowest values at the lowest discharge interval (*i.e.*, 0<sup>th</sup>~10<sup>th</sup> percentile). Deviations from this general pattern are observed in several cases. At the low discharge intervals, Pamunkey and Choptank do not show larger coefficients at relatively higher discharges. At the high discharge intervals, coefficients show three types of pattern: (a) continue increasing with discharge (Danville, Lewisburg, Newport, Conestoga, Conowingo), (b) levelling off (Towanda, Marietta, and Choptank), and (c) decline with discharge (Potomac, James, Rappahannock, Appomattox, Pamunkey, Mattaponi, and Patuxent). These latter two types also deviate from the positive monotonic pattern.

#### **8.3.1.2. TP**

Similar to SS, TP coefficients (Figure 8.4) also show predominantly mobilization effects across all discharge intervals, but the coefficients are always  $< 2.0$  and are much smaller than SS coefficients. (Note that these coefficients are intended for comparing the sensitivities of concentration to discharge but not for comparing the concentration magnitudes; see Equation 8.1.) The few exceptions include all sites at the lowest discharge interval and two sites at the highest discharge interval (Mattaponi and Pamunkey). In these cases, the coefficients are consistently around zero, exhibiting chemostasis or even dilution.

TP coefficients (Figure 8.4) also follow a positive monotonic pattern with respect to discharge. Deviations from this pattern are observed at both low-discharge and high-discharge ends. At the low discharge intervals, coefficients do not show increases against discharge at four sites (Newport, Appomattox, Pamunkey and Choptank). At the high discharge intervals, coefficients show decline against discharge at several sites, namely, Rappahannock, Pamunkey, Mattaponi, Patuxent, and Choptank.

#### **8.3.1.3. TN**

Unlike SS or TP, TN coefficients (Figure 8.5) do not show predominantly mobilization effects. Instead, near-zero coefficients (chemostasis) are much more commonly observed for many discharge intervals. Compared with SS and TP, TN coefficients are generally much smaller (all  $< 0.8$ ). Non-chemostatic patterns are nonetheless noted: mobilization effects are observed with Newport, Potomac, James, and Rappahannock for most discharge intervals, whereas dilution effects are observed with Patuxent for all discharge intervals and Conestoga for high discharge intervals.

TN coefficients (Figure 8.5) generally do not follow a positive monotonic pattern with respect to discharge; the discharge-dependency is rather diverse. Among the 15 sites, positive monotonic patterns are only observed with Lewisburg, James, Appomattox, and Pamunkey. Among the other sites, negative monotonic patterns are observed with Newport, Conestoga, and Potomac, whereas near-flat patterns are observed with Marietta, Conowingo, Mattaponi, and Choptank.

### **8.3.2. Temporal Changes in Coefficients for Selected Discharge Conditions**

Here we address the second question posited above: *Are C-Q relationships temporally varying (i.e., non-stationary) under different discharge conditions and how do*

*the patterns compare among sites and species?* In this analysis, the estimated  $\beta_2$  coefficients are grouped by year under three selected fixed discharge conditions. Such examination separately accommodates the effects of inter-annual discharge variability on estimated coefficients and can thus provide a more focused view of temporal changes in riverine export pattern. Recall that WRTDS model coefficients were estimated and stored on 14 discharge levels, so we selected the discharge levels closest to the 10<sup>th</sup>, 60<sup>th</sup>, and 99.5<sup>th</sup> percentiles of the daily discharge record to represent the low-, mid-, and high-discharges conditions, respectively. Estimated  $\beta_2$  coefficients for these selected discharges (*i.e.*, horizontal slices of the contour plot) were extracted from the contour plot (*e.g.*, Figure 8.2) and their annual-average values were calculated, which are plotted in Figures 8.6-8.8 for SS, TP, and TN, respectively. In addition, 90% confidence intervals obtained using the aforementioned bootstrap method are shown as dashed lines in these figures. For ease of discussion, the period-of-record changes in estimated annual  $\beta_2$  coefficient ( $\Delta$ ) were also quantified, *i.e.*,  $\Delta = \beta_{2,2014} - \beta_{2,1984}$ . Moreover, the probability of  $\Delta > 0$  (“ $P_{\Delta > 0}$ ”) or  $\Delta < 0$  (“ $P_{\Delta < 0}$ ”) was calculated based on the 50 runs. These results are summarized in Table 8.3 and elaborated below.

### **8.3.2.1. SS**

At high discharges (red lines in Figure 8.6; Table 8.3), SS coefficients have declined at 10 of the 15 sites. The largest decline occurred at Pamunkey ( $\Delta = -0.91$ ), whereas the largest rise occurred at Conowingo (+0.81). Both changes are robust based on the 50 replicate runs. At middle discharges (blue lines), SS coefficients have declined at only 6 of the 15 sites, with the largest decline again occurred at Pamunkey (-0.35) and the largest rise at Rappahannock (+0.26). At low discharges (black lines), SS coefficients

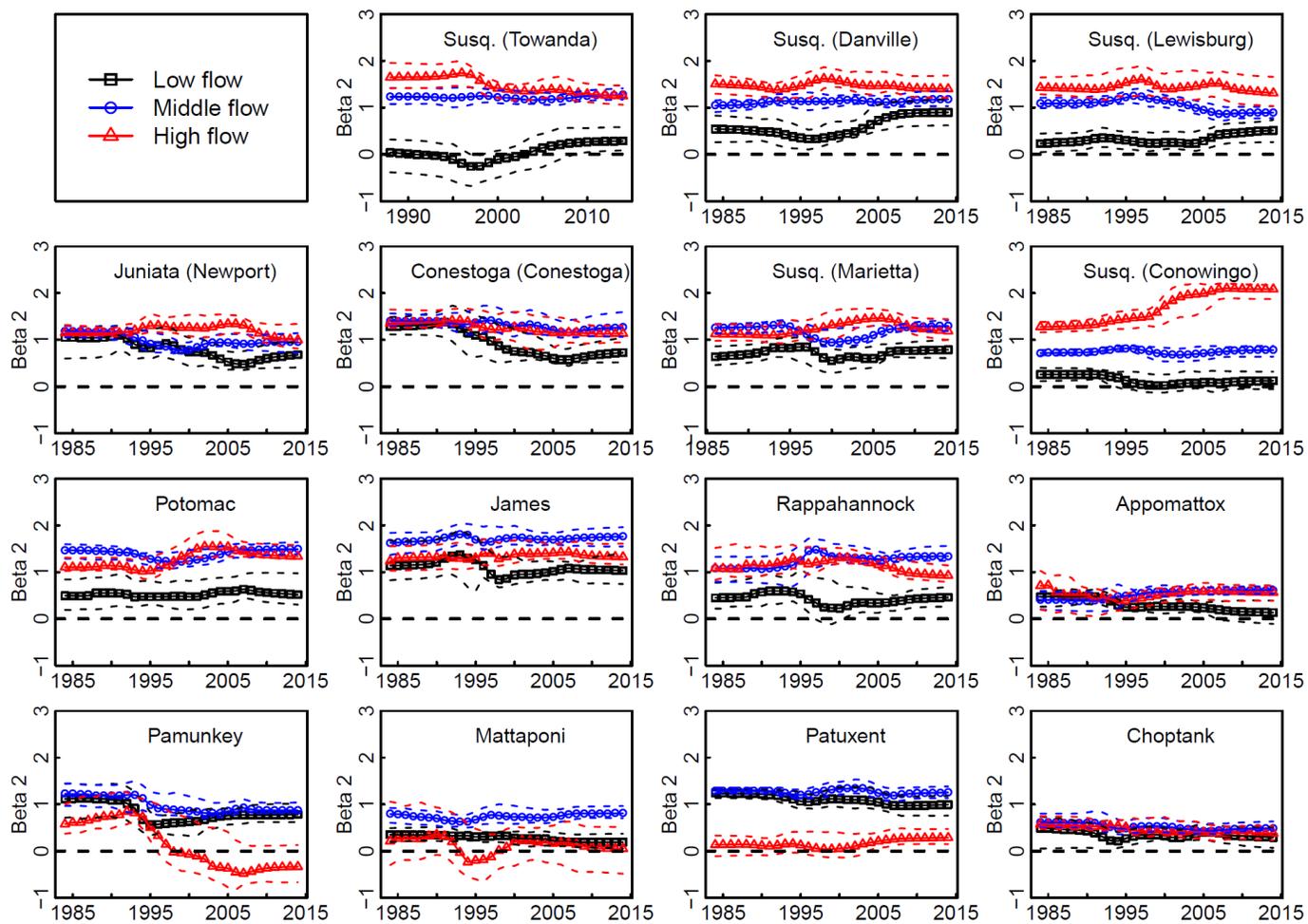


Figure 8.6. Annual averages of estimated WRTDS  $\beta_2$  coefficients for three selected discharges for suspended sediment (SS) at the 15 Chesapeake sites. Dashed lines represent the 90% confidence interval as derived from 50 bootstrap runs based on resampled data.

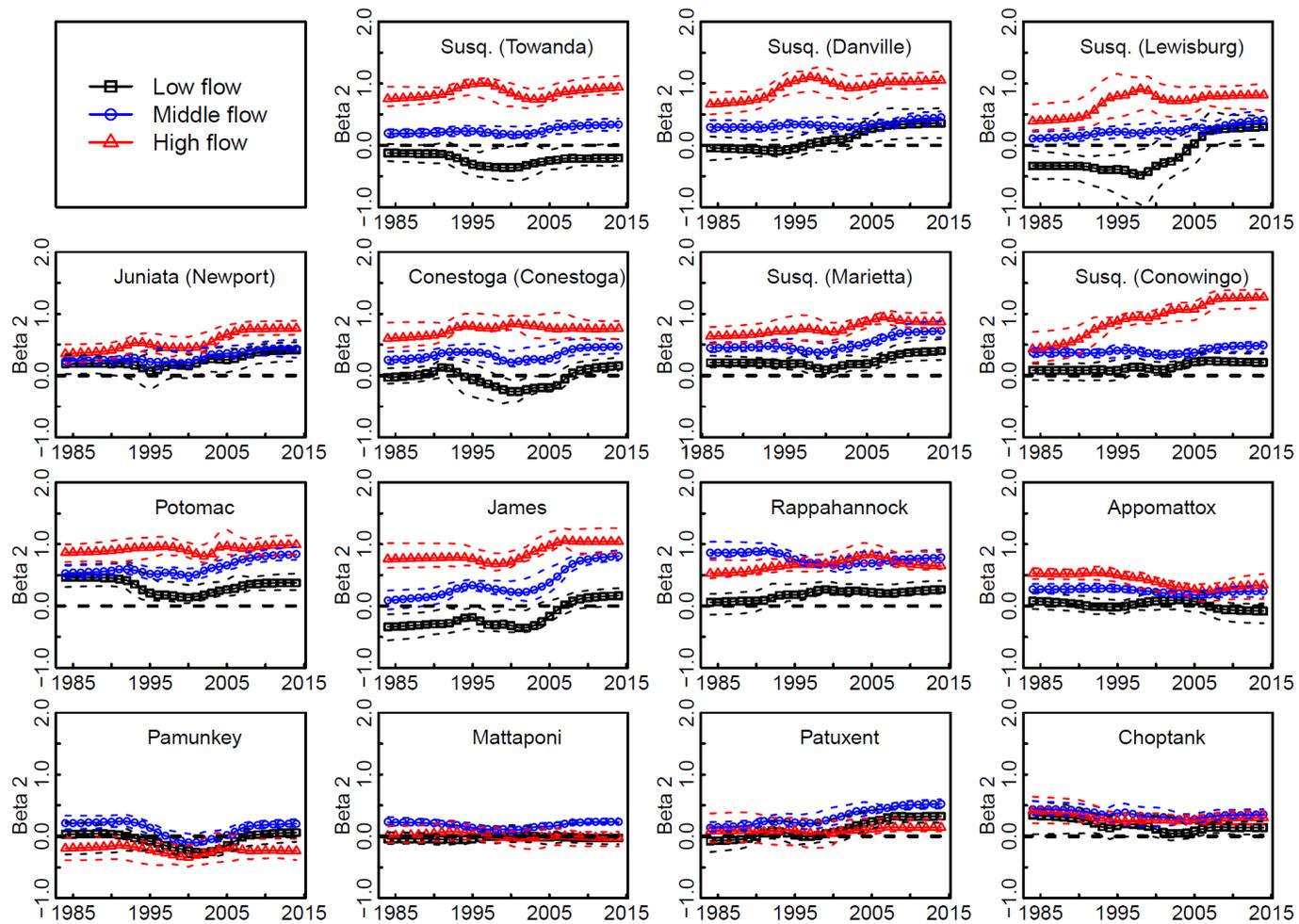


Figure 8.7. Annual averages of estimated WRTDS  $\beta_2$  coefficients for three selected discharges for total phosphorus (TP) at the 15 Chesapeake sites. Dashed lines represent the 90% confidence interval as derived from 50 bootstrap runs based on resampled data.

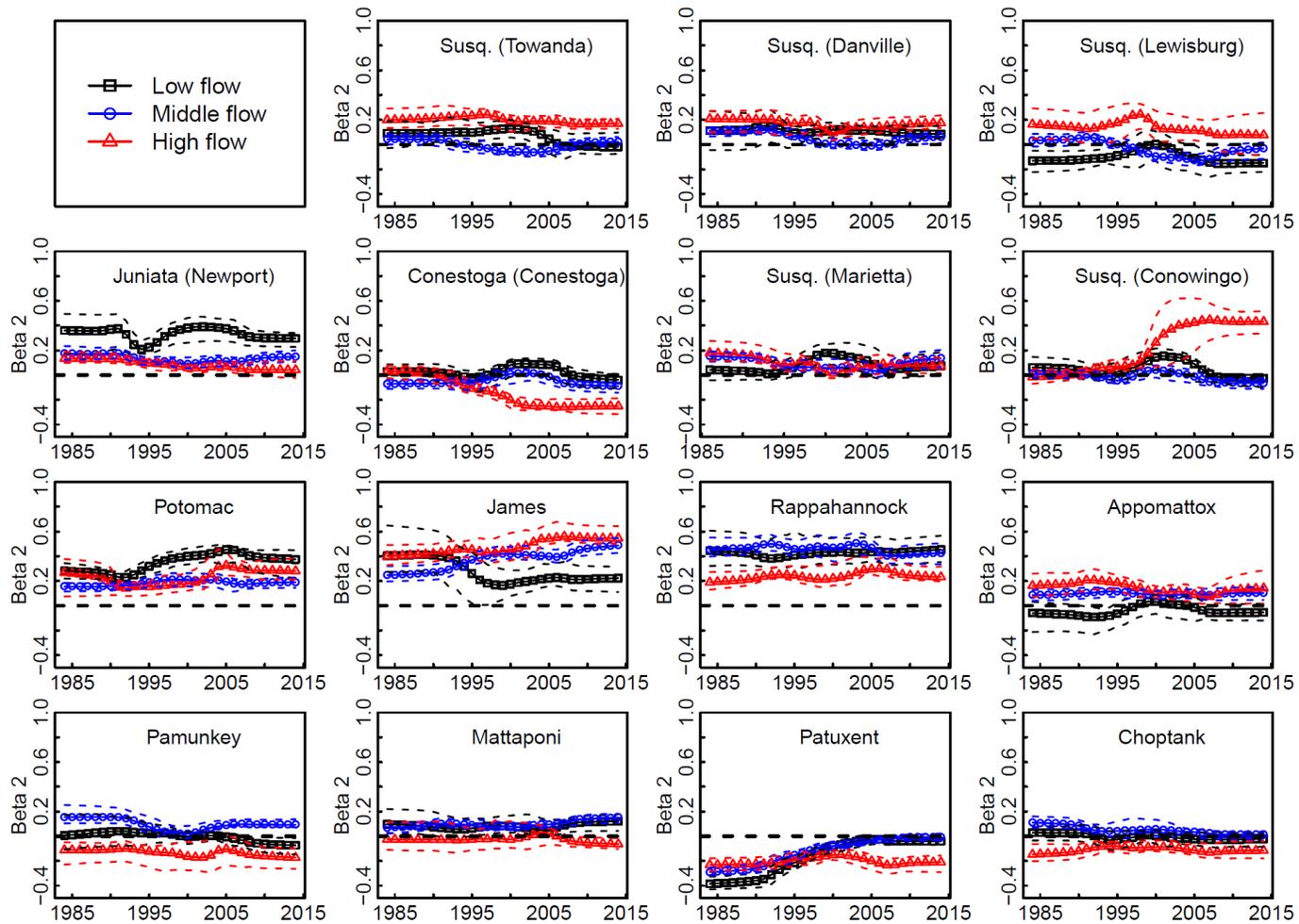


Figure 8.8. Annual averages of estimated WRTDS  $\beta_2$  coefficients for three selected discharges for total nitrogen (TN) at the 15 Chesapeake sites. Dashed lines represent the 90% confidence interval as derived from 50 bootstrap runs based on resampled data.

Table 8.3. Period-of-record changes ( $\Delta$ ) in estimated WRTDS  $\beta_2$  coefficients at the 15 Chesapeake sites under three different discharge conditions. ( $\Delta_{\text{Low}}$ :  $\Delta$  under low-discharge condition;  $\Delta_{\text{Mid}}$ :  $\Delta$  under mid-discharge condition;  $\Delta_{\text{High}}$ :  $\Delta$  under high-discharge condition;  $P_{\Delta>0}$ : probability of positive change [ $\Delta > 0$ ] observed in the 50 replicate runs; pink cells:  $\Delta > 0$ ; green cells:  $\Delta < 0$ ; yellow cells:  $P_{\Delta>0} > 0.9$  or  $P_{\Delta<0} > 0.9$ .)

River Sites		Suspended Sediment (SS)						Total Phosphorus (TP)						Total Nitrogen (TN)					
		$\Delta_{\text{Low}}$	$P_{\Delta>0}$	$\Delta_{\text{Mid}}$	$P_{\Delta>0}$	$\Delta_{\text{High}}$	$P_{\Delta>0}$	$\Delta_{\text{Low}}$	$P_{\Delta>0}$	$\Delta_{\text{Mid}}$	$P_{\Delta>0}$	$\Delta_{\text{High}}$	$P_{\Delta>0}$	$\Delta_{\text{Low}}$	$P_{\Delta>0}$	$\Delta_{\text{Mid}}$	$P_{\Delta>0}$	$\Delta_{\text{High}}$	$P_{\Delta>0}$
RIM Sites	Conowingo	-0.13	0.18	0.06	0.63	0.81	1.00	0.13	0.94	0.11	0.96	0.84	1.00	-0.09	0.06	-0.07	0.02	0.45	1.00
	Potomac	0.02	0.59	0.02	0.65	0.23	0.96	-0.09	0.22	0.30	1.00	0.13	0.84	0.09	0.82	0.03	0.88	0.01	0.69
	James	-0.10	0.37	0.14	0.75	0.09	0.65	0.51	1.00	0.72	1.00	0.28	0.94	-0.18	0.06	0.24	1.00	0.15	1.00
	Rappahannock	0.01	0.53	0.26	0.90	-0.16	0.33	0.21	0.96	-0.08	0.20	0.12	0.86	0.01	0.55	-0.03	0.29	0.04	0.86
	Appomattox	-0.33	0.02	0.21	0.96	-0.15	0.59	-0.16	0.10	-0.02	0.16	-0.19	0.10	0.01	0.63	0.02	0.67	-0.02	0.33
	Pamunkey	-0.33	0.16	-0.35	0.06	-0.91	0.00	0.03	0.61	0.00	0.43	-0.04	0.45	-0.08	0.12	-0.06	0.12	-0.06	0.25
	Mattaponi	-0.16	0.18	0.01	0.71	-0.16	0.31	0.01	0.51	0.01	0.61	-0.03	0.31	0.02	0.49	0.08	0.98	-0.04	0.27
	Patuxent	-0.25	0.02	-0.03	0.31	0.15	0.86	0.40	1.00	0.38	1.00	0.06	0.75	0.34	1.00	0.28	1.00	0.02	0.59
	Choptank	-0.20	0.20	-0.12	0.20	-0.20	0.06	-0.20	0.16	-0.09	0.20	-0.11	0.14	-0.05	0.12	-0.10	0.00	0.03	0.63
SRBC Sites	Marietta	0.15	0.78	0.04	0.65	0.06	0.63	0.19	0.94	0.27	1.00	0.23	0.96	0.03	0.69	-0.02	0.31	-0.11	0.02
	Danville	0.36	0.92	0.14	0.90	-0.10	0.31	0.40	0.98	0.15	0.92	0.39	1.00	-0.03	0.49	-0.04	0.06	-0.03	0.43
	Towanda	0.25	0.86	0.04	0.61	-0.40	0.04	-0.08	0.41	0.14	0.96	0.19	0.90	-0.11	0.16	-0.02	0.33	-0.03	0.27
	Lewisburg	0.28	0.88	-0.20	0.00	-0.12	0.24	0.63	1.00	0.30	1.00	0.42	0.98	-0.02	0.33	-0.06	0.06	-0.09	0.25
	Newport	-0.37	0.12	-0.23	0.10	-0.15	0.29	0.22	0.88	0.20	1.00	0.41	1.00	-0.06	0.25	-0.02	0.25	-0.09	0.02
	Conestoga	-0.56	0.00	-0.14	0.25	-0.22	0.06	0.19	0.98	0.21	1.00	0.16	0.76	-0.07	0.08	-0.02	0.16	-0.28	0.00

have declined at 10 of the 15 sites, with the largest decline occurred at Conestoga (-0.56) and the largest rise at Danville (+0.36).

Among the 15 sites, four sites show declines at all three discharges, namely, Newport, Conestoga, Pamunkey, and Choptank; one site (Marietta) shows rises at all three discharges; the other sites show mixed trends among the three discharges. Notably, the three largest RIM rivers (*i.e.*, Susquehanna [Conowingo], Potomac, and James; Table 8.1) all show rises at middle and high discharges. In particular, the Conowingo rise at high discharge is much stronger than Marietta or any other site.

#### **8.3.2.2. TP**

At high discharges (red lines in Figure 8.7; Table 8.3), TP coefficients have increased at 11 of the 15 sites. The largest rise occurred at Conowingo ( $\Delta = +0.84$ ), whereas the largest decline occurred at Appomattox (-0.19). Both changes are robust based on the 50 replicate runs. At middle discharges (blue lines), TP coefficients have also increased at 11 of the 15 sites, with the largest decline occurred at Choptank (-0.09) and the largest rise at James (+0.72). At low discharges (black lines), TP coefficients have increased at 11 of the 15 sites, with the largest decline occurred at Choptank (-0.20) and the largest rise at Lewisburg (+0.63).

Among the 15 sites, only two sites (*i.e.*, Appomattox and Choptank) show declines at all three discharges. By contrast, eight sites show rises at all three discharges, including James, Patuxent, and all Susquehanna sites except Towanda. As in the case of SS coefficients, the three largest RIM rivers (*i.e.*, Susquehanna [Conowingo], Potomac, and James) all show rises at middle and high discharges, with Conowingo rise at high discharge being stronger than Marietta or any other site.

### 8.3.2.3. TN

At high discharges (red lines in Figure 8.8; Table 8.3), TN coefficients have declined at 9 of the 15 sites. The largest decline occurred at Conestoga ( $\Delta = -0.28$ ), whereas the largest rise occurred at Conowingo (+0.45). Both changes are robust based on the 50 replicate runs. At middle discharges (blue lines), TN coefficients have declined at 10 of the 15 sites, with the largest decline occurred at Choptank (-0.10) and the largest rise at Patuxent (+0.28). At low discharges (black lines), TN coefficients have declined at 9 of the 15 sites, with the largest decline occurred at James (-0.18) and the largest rise at Patuxent (+0.34).

Among the 15 sites, only two sites (*i.e.*, Potomac and Patuxent) show rises at all three discharges. By contrast, five sites show declines at all three discharges, including Pamunkey and five Susquehanna sites (Towanda, Danville, Lewisburg, Newport, and Conestoga). Again, the three largest RIM rivers (*i.e.*, Susquehanna [Conowingo], Potomac, and James) all show rises at high discharges, with Conowingo showing a stronger rise than any other site.

## 8.4. Discussion

### 8.4.1. Changes in Coefficients with Discharge

For SS and TP,  $C-Q$  patterns are predominantly mobilization under a wide range of discharge conditions. Diffuse (flow-dependent) sources are likely dominant in these watersheds, otherwise dilution patterns would have been prevalent. As discharge increases, source areas across the landscape become more connected hydrologically (Thompson *et al.*, 2012; Wolf *et al.*, 2013; Outram *et al.*, 2016). Therefore, mobilization

patterns observed under low-discharge and high-discharge conditions can possibly reflect different contributions of various source areas. Specifically, low-discharge mobilization would likely arise due to the flushing of nonpoint sources that are near stream (either surface or sub-surface) and/or subject to a rapid transport pathway, whereas the high-discharge mobilization more likely indicates the flushing of nonpoint sources that are far from stream and/or subject to a delayed transport mechanism. Such distinct mechanisms sometimes are referred to as “proximal” and “distal” responses, respectively (Sherriff *et al.*, 2016). The processes are, of course, also influenced by other spatially-varying factors, such as physiography, channel geometry, and flood-plain structure.

At very low discharges, however, chemostasis patterns are observed, including TP at all sites and SS at five sites (Towanda, Conowingo, Rappahannock, Appomattox, and Mattaponi). These patterns likely indicate the existence of flow thresholds for mobilization of particulate constituents, below which watershed surface is still relatively dry and flow paths remain largely hydrologically disconnected with the river channel and/or the source zones (Shanley *et al.*, 2011; Thompson *et al.*, 2012; Wolf *et al.*, 2013). Consequently, constituent concentration is insensitive to flow-generation processes in these cases. Comparatively, such threshold appears to be consistently higher for phosphorus than sediment, which may be because phosphorus is less available or more difficult to remobilize than sediment. At high discharges, chemostasis is rare and only observed with SS at Mattaponi and with TP at Pamunkey and Mattaponi, implying equilibrium between constituent and water fluxes. These cases, as well as several others showing decreased levels of mobilization at the highest discharge interval (*e.g.*, SS for Mattaponi and Patuxent; TP for Pamunkey), may be caused by exhaustion of constituent

supply and/or deposition of constituent fluxes in the flood plains. In this context, Mattaponi behaves very differently from Choptank at the high-discharge end (chemostasis *vs.* mobilization, respectively), although both rivers are dominated by the coastal-plain province. A sensible hypothesis is the marked difference in agricultural intensity in the two watersheds – 19% (Mattaponi) *vs.* 50% (Choptank) in terms of agricultural land fraction (Table 8.1).

For TN, chemostasis is more prevalent under a wide range of discharge conditions, suggesting general equilibrium between nitrogen and water fluxes. Dilution patterns are observed in some cases, particularly Patuxent (most discharges) and Conestoga (high discharges). Such patterns highlight the dominance of point source contributions in these two watersheds, which have the highest fractional areas of urban land (13% and 8%) among all 15 watersheds (Table 8.1). Mobilization patterns are also observed in some cases, including low-discharge conditions (Newport, Potomac, James, and Rappahannock) and high-discharge conditions (James, Rappahannock, and Susquehanna), which indicate the dominance of nonpoint sources in these watersheds and may reflect “proximal” and “distal” responses, respectively, as discussed above. Remarkably, strong mobilization is observed with Conowingo for the highest discharge interval, which is not observed with Marietta (inlet of Conowingo Reservoir). Such contrast highlights the effects of particulate N remobilization within Conowingo Reservoir (Zhang *et al.*, 2016c).

Among all three constituents, TN shows the smallest coefficients and also the weakest variability in coefficients, reflecting distinctions in source and fate of these constituents. (Note that smaller coefficients indicate smaller sensitivities of concentration

to discharge, but not necessarily smaller concentrations.) Specifically, SS and TP, which are more likely dominated by surface transport, are expected to have undergone more spatially-heterogeneous processes, *e.g.*, sources, transport pathways, and perhaps reactions (Heathwaite and Dils, 2000; Brakebill *et al.*, 2010; Dupas *et al.*, 2015; Zhang *et al.*, 2015). By contrast, TN, likely dominated by subsurface transport, is expected to have undergone relatively more homogeneous processes due to subsurface storage and mixing over a range of spatial and temporal scales (Kirchner and Neal, 2013; Sanford and Pope, 2013; Harman, 2015). Being less insensitive to discharge than SS and TP, TN appears to be less dominated by hydrological processes; other factors such as biogeochemical transformations may have played a more important role in flux regulation.

Another related feature of TN is the predominance of dissolved fraction, which partially originates from urban and industrial point sources. Thus,  $\beta_2$  coefficient is expected to be smaller for more urbanized watersheds (*i.e.*, larger point-source dominance). This hypothesis is supported by a statistically significant correlation (p-value < 0.05) between the period-of-record median TN coefficient and the fraction of urban land (or fractional point source contribution) among the 15 sites. A similar relationship, however, is not observed with SS or TP, presumably due to much less dominance of point sources.

Overall, export at the 15 Chesapeake sub-watersheds has been dominated by mobilization patterns for SS and TP (particulate-dominated species) and chemostasis patterns for TN (dissolved-dominated species). The general lack of dilution patterns across most discharge intervals may suggest that none of the three species has been supply-limited in these watersheds. Thus, it is possible that there exists sufficiently large

storage (surface or subsurface) for nutrients and sediment due to long-term legacy inputs, as previously suggested on Chesapeake Bay, Mississippi River, and Lake Erie basins (Meals *et al.*, 2009; Basu *et al.*, 2010; Jarvie *et al.*, 2013; Sharpley *et al.*, 2013).

Consistent with this speculation, the period-of-record decline in TN riverine yield has been generally smaller than that in total source input in these 15 watersheds (Figure 8.9).

While this work was intended to seek common patterns among the 15 watersheds, further research is called for to investigate factors that have driven the unique behavior of each individual watershed, which may include but not limited to, flow-generation processes, flood-plain structure, groundwater and surface-water interactions, river flashiness, watershed input, density of reservoirs and dams, and land uses.

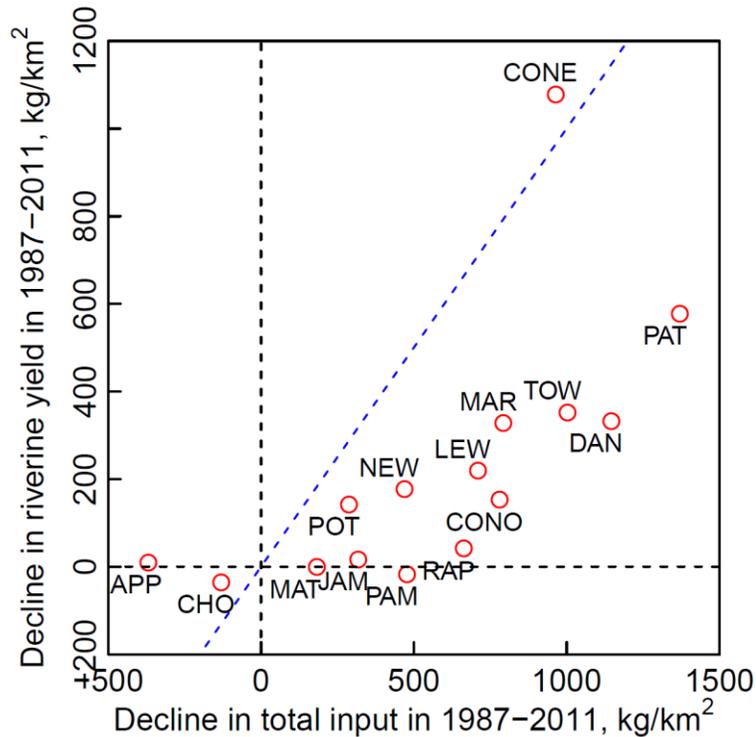


Figure 8.9. Comparison between period-of-record decline in riverine yield and decline in total source input for total nitrogen (TN) across the 15 sites. Blue dashed line is 1:1 reference line.

#### 8.4.2. Temporal Changes in Coefficients for Selected Discharge Conditions

At the selected representative discharges, SS coefficients show mixed temporal trends across sites (Figure 8.6). One of the most remarkable change is Pamunkey at high discharge ( $\Delta = -0.91$ ), which is statistically robust since all 50 replicate runs show negative  $\Delta$  values (*i.e.*,  $P_{\Delta < 0} = 1.0$ ). This decline has caused the export pattern to switch from mobilization to dilution, suggesting decreased dominance of nonpoint sources and possibly significant depletion of such sources over the study period. Another notable pattern is the consistent decline under all three discharge conditions at four sites (*i.e.*, Newport, Conestoga, Pamunkey, and Choptank), possibly reflecting combined reduction in a wide range of nonpoint sources (proximal and distal). By contrast, the three largest RIM rivers (*i.e.*, Susquehanna [Conowingo], Potomac, and James) show rises at middle and high discharges, all with high confidence ( $P_{\Delta > 0} = 0.65\sim 1.0$ ; Table 8.3). The rise at Conowingo ( $\Delta = +0.81$ ) is particularly remarkable, which is also statistically robust (*i.e.*,  $P_{\Delta > 0} = 1.0$ ). Considering that the Marietta coefficient has remained almost unchanged ( $\Delta = +0.06$ ), the dramatic rise at Conowingo can be mainly attributed to the diminished trapping capacity of Conowingo Reservoir, as documented previously using other approaches (Hirsch, 2012; Zhang *et al.*, 2013; Zhang *et al.*, 2016c). The rises at Potomac and James remain less clear but may reflect combined effects of land clearance and urbanization (Brakebill *et al.*, 2010), removal of small mill dams (Walter and Merritts, 2008; Merritts *et al.*, 2011), and altered rainfall and watershed conditions that promote erosion and transport (Karl and Knight, 1998).

For TP, the predominant pattern is positive  $\Delta$  under the three discharge conditions. Among the most significant rises, the Conowingo rise at high discharge ( $\Delta = +0.84$ ) is

most remarkable and statistically robust. This rise is also much larger than that of Marietta ( $\Delta = +0.23$ ) and thus again corroborates the important effect of decreased trapping efficiency of Conowingo Reservoir. Moreover, TP coefficients have increased at high discharges at four sites (*i.e.*, Conowingo, Potomac, James, and Rappahannock), all with high confidence ( $P_{\Delta>0} = 0.84\sim 1.0$ ; Table 8.3). These rises are alarming, since the four sites monitor the four largest RIM tributaries and thus a vast majority of the entire non-tidal Chesapeake watershed. Several sites showed important switches of TP pattern at low discharges: Lewisburg and James switched from dilution to mobilization, whereas Danville, Conowingo, Rappahannock, and Patuxent switched from chemostasis to mobilization. These switches probably suggest depletion of flow-independent sources and/or enrichment of nonpoint (flow-dependent) sources over the study period. Reverse effects have occurred at Appomattox and Choptank, where the low-discharge pattern switched from mobilization to dilution or chemostasis, reflecting decreased relative dominance of nonpoint sources. Among the 15 sites, only two (*i.e.*, Appomattox and Choptank) show declines at all three discharge conditions, *i.e.*, becoming less flow-dependent. By contrast, eight sites show rises at all three discharge conditions, all with high confidence. In particular, the low-discharge rise may reflect decreased point-source contribution due to management actions such as P-detergent ban since 1990 (Litke, 1999) and enhanced nutrient removal at wastewater treatment plants since the 1980s (Sprague *et al.*, 2000; Boynton *et al.*, 2008).

For TN, the predominant pattern is negative  $\Delta$  under the three discharge conditions. Among all sites, the largest decline is observed with Conestoga at high discharge ( $\Delta = -0.28$ ), which is statistically robust based on the 50 replicate runs. Interestingly, the export

pattern has switched from mobilization to strong dilution. Such change may reflect the effectiveness of nonpoint source reduction in this mixed-land-use watershed. By contrast, TN coefficients have increased at high discharges in the four largest RIM tributaries (*i.e.*, Susquehanna [Conowingo], Potomac, James, and Rappahannock), all with high confidence ( $P_{\Delta>0} = 0.69\sim 1.0$ ; Table 8.3). The most significant rise is observed with Conowingo at high discharge ( $\Delta = +0.45$ ), which is statistically robust ( $P_{\Delta>0} = 1.0$ ). This rise has caused the export pattern to switch from chemostasis to strong mobilization. The physical implication is that N concentration in the reservoir effluent has become more sensitive to discharge as it approaches sediment storage capacity (Langland, 2015). Again, considering that Marietta has a negative  $\Delta$  at high discharge ( $\Delta = -0.11$ ), the Conowingo rise highlights diminished trapping of particulate N in Conowingo Reservoir. Six sites show declines at all three discharge conditions, including Pamunkey and five Susquehanna sites, reflecting reduced sensitivity to discharge that was probably attributable to reduction of nonpoint sources and atmospheric sources (Linker *et al.*, 2013). By contrast, Potomac and Patuxent both show rises at all three discharge, which seem to have followed different mechanisms. For Potomac, the coefficients have become more positive, suggesting increased nonpoint source dominance. For Patuxent, the coefficients have become less negative, suggesting decreased point-source dominance due to the upgrade of major wastewater treatment plants (Sprague *et al.*, 2000; Boynton *et al.*, 2008).

Overall, these results demonstrate clear temporal non-stationarity in  $C-Q$  patterns at the 15 Chesapeake sites. The diverse trends under different discharge conditions possibly reflect changes in dominant watershed sources due to management actions. Further

research is needed to systematically investigate factors that have driven the observed functional changes under different discharge conditions for different watershed systems. More generally, these non-stationary relationships further highlight the complexity of watershed function, which should be taken into consideration when riverine  $C$ - $Q$  data are used to infer transport processes or estimate concentrations and fluxes on unsampled dates using regression-based approaches (Horowitz, 2003; Crowder *et al.*, 2007; Hirsch *et al.*, 2010; Hirsch, 2014).

## 8.5. Conclusions

This paper presents a top-down synthesis of  $C$ - $Q$  patterns for multiple constituents in nine major tributaries (15 sites) of Chesapeake Bay, which has offered several new insights on the complexity of watershed function (*i.e.*, concentration sensitivity to discharge) in the Chesapeake sub-watersheds. First of all, constituent export at the 15 sites has been dominated by mobilization patterns for SS and TP (particulate-dominated species) and chemostasis patterns for TN (dissolved-dominated species) under many discharge conditions. The general lack of dilution patterns may suggest that none of the three constituents has been supply-limited in these watersheds. In addition, SS and TP coefficients have followed a clear positive monotonic pattern with respect to discharge, exhibiting threshold behaviors (*i.e.*, chemostasis at low discharge *vs.* mobilization at high discharge). Moreover, for many site-constituent combinations, coefficients show clear temporal non-stationarity in  $C$ - $Q$  relationships under selected fixed discharges, possibly reflecting changes in dominant watershed sources due to management actions. Further research is needed to systematically investigate factors that have driven the observed functional changes under different discharge conditions for different watershed systems.

These results also highlight the potential pitfalls of assuming fixed  $C$ - $Q$  relationships over discharge or time, which violates the underlying behavior of the monitoring data, and if unaccounted for, can cause doubtful estimation of constituent concentration and flux.

Methodologically, this work has made three new contributions. First, it has demonstrated the utility of the WRTDS method to make robust and informative interpretation of  $C$ - $Q$  relationships in traditional (discretely sampled) monitoring data. Unlike many previous  $C$ - $Q$  studies that focused on stormflow conditions, this approach allows simultaneous examination of various discharge conditions in a consistent manner. Second, this work has adopted a top-down approach that has provided a synthesis of temporal and spatial patterns of  $C$ - $Q$  relationships across multiple sites. These results may be useful to infer the status and changes of constituent export from different watersheds as well as the relative dominance of sources under different flow classes. With a focus on the sensitivity of concentration to discharge, the WRTDS coefficients provide an alternative approach for interpretation of water-quality data and for generation of sensible hypothesis on dominant processes in different watersheds. Third, such synthesis was conducted within an uncertainty framework to provide sound conclusions, which is believed to be the first of its kind under this topic. More generally, this approach is adaptable to other river systems and our synthesis effectively illustrates the value of long-term data collection and maintenance for multiple locations in the watersheds.

## **8.6. Supporting Information**

Supporting information to this chapter is provided in Appendix F.

## 8.7. Acknowledgements

This work was supported by the Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090), Maryland Water Resources Research Center (2015MD329B), and National Science Foundation (CBET-1360415). We thank Bob Hirsch, Doug Moyer, and Joel Blomquist (USGS) and Dano Wilusz (JHU) for insightful discussions. We thank Gary Shenk, Guido Yactayo, and Gopal Bhatt (Chesapeake Bay Program Office) for providing the source input data and Andrew Sekellick for his assistance on the watershed map production.

## 8.8. Literature Cited

- Basu, N. B., G. Destouni, J. W. Jawitz, S. E. Thompson, N. V. Loukinova, A. Darracq, S. Zanardo, M. Yaeger, M. Sivapalan, A. Rinaldo and P. S. C. Rao, 2010. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophys. Res. Lett.* 37:L23404, DOI: 10.1029/2010gl045168.
- Basu, N. B., S. E. Thompson and P. S. C. Rao, 2011. Hydrologic and biogeochemical functioning of intensively managed catchments: A synthesis of top-down analyses. *Water Resour. Res.* 47:W00J15, DOI: 10.1029/2011WR010800.
- Bieroza, M. Z. and A. L. Heathwaite, 2015. Seasonal variation in phosphorus concentration–discharge hysteresis inferred from high-frequency in situ monitoring. *Journal of Hydrology* 524:333-347, DOI: 10.1016/j.jhydrol.2015.02.036.
- Bowes, M. J., J. T. Smith, H. P. Jarvie and C. Neal, 2008. Modelling of phosphorus inputs to rivers from diffuse and point sources. *Sci. Total Environ.* 395:125-138, DOI:

10.1016/j.scitotenv.2008.01.054.

Bowes, M. J., J. T. Smith, H. P. Jarvie, C. Neal and R. Barden, 2009. Changes in point and diffuse source phosphorus inputs to the River Frome (Dorset, UK) from 1966 to 2006. *Sci. Total Environ.* 407:1954-1966, DOI:

10.1016/j.scitotenv.2008.11.026.

Boynton, W. R., J. D. Hagy, J. C. Cornwell, W. M. Kemp, S. M. Greene, M. S. Owens, J. E. Baker and R. K. Larsen, 2008. Nutrient Budgets and Management Actions in the Patuxent River Estuary, Maryland. *Estuaries Coasts* 31:623-651, DOI:

10.1007/s12237-008-9052-9.

Brakebill, J. W., S. W. Ator and G. E. Schwarz, 2010. Sources of suspended-sediment flux in streams of the Chesapeake Bay Watershed: A regional application of the SPARROW model. *J. Am. Water Resour. Assoc.* 46:757-776, DOI:

10.1111/j.1752-1688.2010.00450.x.

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

Chanat, J. G., D. L. Moyer, J. D. Blomquist, K. E. Hyer and M. J. Langland, 2016. Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year 2012. U.S. Geological Survey Scientific Investigations Report 2015-5133, Reston, VA, p. 76. <http://dx.doi.org/10.3133/sir20155133>.

Chanat, J. G., K. C. Rice and G. M. Hornberger, 2002. Consistency of patterns in

- concentration-discharge plots. *Water Resour. Res.* 38:1-10, DOI: 10.1029/2001WR000971.
- Crowder, D. W., M. Demissie and M. Markus, 2007. The accuracy of sediment loads when log-transformation produces nonlinear sediment load–discharge relationships. *Journal of Hydrology* 336:250-268, DOI: 10.1016/j.jhydrol.2006.12.024.
- Dupas, R., C. Gascuel-Oudou, N. Gilliet, C. Grimaldi and G. Gruau, 2015. Distinct export dynamics for dissolved and particulate phosphorus reveal independent transport mechanisms in an arable headwater catchment. *Hydrol. Process.* 29:3162–3178, DOI: 10.1002/hyp.10432.
- Evans, C. and T. D. Davies, 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. *Water Resour. Res.* 34:129-137, DOI: 10.1029/97WR01881.
- Godsey, S. E., J. W. Kirchner and D. W. Clow, 2009. Concentration-discharge relationships reflect chemostatic characteristics of US catchments. *Hydrol. Process.* 23:1844-1864, DOI: 10.1002/hyp.7315.
- Gray, A. B., G. B. Pasternack, E. B. Watson, J. A. Warrick and M. A. Goñi, 2015. Effects of antecedent hydrologic conditions, time dependence, and climate cycles on the suspended sediment load of the Salinas River, California. *Journal of Hydrology* 525:632-649, DOI: 10.1016/j.jhydrol.2015.04.025.
- Harman, C. J., 2015. Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed. *Water Resour. Res.* 51:1–30, DOI: 10.1002/2014WR015707.

- Heathwaite, A. and R. Dils, 2000. Characterising phosphorus loss in surface and subsurface hydrological pathways. *Sci. Total Environ.* 251-252:523-538, DOI: 10.1016/S0048-9697(00)00393-4.
- Herndon, E. M., A. L. Dere, P. L. Sullivan, D. Norris, B. Reynolds and S. L. Brantley, 2015. Landscape heterogeneity drives contrasting concentration–discharge relationships in shale headwater catchments. *Hydrol. Earth Syst. Sci.* 19:3333-3347, DOI: 10.5194/hess-19-3333-2015.
- Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.
- Hirsch, R. M., 2014. Large Biases in Regression-Based Constituent Flux Estimates: Causes and Diagnostic Tools. *J. Am. Water Resour. Assoc.* 50:1401-1424, DOI: 10.1111/jawr.12195.
- Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.
- Hirsch, R. M. and L. De Cicco, 2015. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). U.S. Geological Survey Techniques and Methods Book 4, Chapter A10, Reston, VA, p. 93. <http://dx.doi.org/10.3133/tm4A10>.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time,

- discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Horowitz, A. J., 2003. An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrol. Process.* 17:3387-3409, DOI: 10.1002/hyp.1299.
- House, W. A. and M. S. Warwick, 1998. Hysteresis of the solute concentration/discharge relationship in rivers during storms. *Water Res.* 32:2279-2290, DOI: 10.1016/S0043-1354(97)00473-9.
- Jarvie, H. P., A. N. Sharpley, B. Spears, A. R. Buda, L. May and P. J. A. Kleinman, 2013. Water Quality Remediation Faces Unprecedented Challenges from “Legacy Phosphorus”. *Environ. Sci. Technol.* 47:8997-8998, DOI: 10.1021/es403160a.
- Karl, T. R. and R. W. Knight, 1998. Secular Trends of Precipitation Amount, Frequency, and Intensity in the United States. *Bull. Am. Meteorol. Soc.* 79:231-241, DOI: 10.1175/1520-0477(1998)079<0231:STOPAF>2.0.CO;2.
- Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C. Stevenson, 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* 303:1-29, DOI: 10.3354/meps303001.
- Kirchner, J. W. and C. Neal, 2013. Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc. Natl.*

*Acad. Sci. U. S. A.* 110:12213-12218, DOI: 10.1073/pnas.1304328110.

Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18.  
<http://dx.doi.org/10.3133/ofr20141235>.

Linker, L. C., R. Dennis, G. W. Shenk, R. A. Batiuk, J. Grimm and P. Wang, 2013. Computing Atmospheric Nutrient Loads to the Chesapeake Bay Watershed and Tidal Waters. *J. Am. Water Resour. Assoc.* 49:1025-1041, DOI: 10.1111/jawr.12112.

Litke, D. W., 1999. Review of phosphorus control measures in the United States and their effects on water quality. U.S. Geological Survey Water-Resources Investigations Report 99-4007, Denver, CO, p. 43. <http://pubs.usgs.gov/wri/wri994007/>.

Meals, D. W., S. A. Dressing and T. E. Davenport, 2009. Lag time in water quality response to best management practices: a review. *J. Environ. Qual.* 39:85-96, DOI: 10.2134/jeq2009.0108.

Merritts, D., R. Walter, M. Rahnis, J. Hartranft, S. Cox, A. Gellis, N. Potter, W. Hilgartner, M. Langland, L. Manion, C. Lippincott, S. Siddiqui, Z. Rehman, C. Scheid, L. Kratz, A. Shilling, M. Jenschke, K. Datin, E. Cranmer, A. Reed, D. Matuszewski, M. Voli, E. Ohlson, A. Neugebauer, A. Ahamed, C. Neal, A. Winter and S. Becker, 2011. Anthropocene streams and base-level controls from historic dams in the unglaciated mid-Atlantic region, USA. *Philosophical Transactions of the Royal Society A* 369:976-1009, DOI: 10.1098/rsta.2010.0335.

Meybeck, M. and F. Moatar, 2012. Daily variability of river concentrations and fluxes:

- Indicators based on the segmentation of the rating curve. *Hydrol. Process.* 26:1188-1207, DOI: 10.1002/hyp.8211.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Murphy, R. R., W. M. Kemp and W. P. Ball, 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* 34:1293-1309, DOI: 10.1007/s12237-011-9413-7.
- Outram, F. N., R. J. Cooper, G. Sünnerberg, K. M. Hiscock and A. A. Lovett, 2016. Antecedent conditions , hydrological connectivity and anthropogenic inputs : Factors affecting nitrate and phosphorus transfers to agricultural headwater streams. *Sci. Total Environ.* 546:184-199, DOI: 10.1016/j.scitotenv.2015.12.025.
- Outram, F. N., C. E. M. Lloyd, J. Jonczyk, C. M. H. Benskin, F. Grant, M. T. Perks, C. Deasy, S. P. Burke, A. L. Collins, J. Freer, P. M. Haygarth, K. M. Hiscock, P. J. Johnes and A. L. Lovett, 2014. High-frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of the 2011–2012 drought in England. *Hydrol. Earth Syst. Sci.* 18:3429-3448, DOI: 10.5194/hess-18-3429-2014.
- R Development Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3900051070. <http://www.r-project.org>.
- Sanford, W. E. and J. P. Pope, 2013. Quantifying Groundwater's Role in Delaying

- Improvements to Chesapeake Bay Water Quality. *Environ. Sci. Technol.* 47:13330-13338, DOI: 10.1021/es401334k.
- Shanley, J. B., W. H. McDowell and R. F. Stallard, 2011. Long-term patterns and short-term dynamics of stream solutes and suspended sediment in a rapidly weathering tropical watershed. *Water Resour. Res.* 47:W07515, DOI: 10.1029/2010WR009788.
- Sharpley, A., H. P. Jarvie, A. Buda, L. May, B. Spears and P. Kleinman, 2013. Phosphorus Legacy: Overcoming the Effects of Past Management Practices to Mitigate Future Water Quality Impairment. *J. Environ. Qual.* 42:1308–1326, DOI: 10.2134/jeq2013.03.0098.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- Sherriff, S. C., J. S. Rowan, O. Fenton, P. Jordan, A. R. Melland, P.-E. Mellander and D. Ó. hUallacháin, 2016. Storm Event Suspended Sediment-Discharge Hysteresis and Controls in Agricultural Watersheds: Implications for Watershed Scale Sediment Management. *Environ. Sci. Technol.* 50:1769-1778, DOI: 10.1021/acs.est.5b04573.
- Sprague, L. A., M. J. Langland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. U.S. Geological Survey Water-Resources Investigations Report 00-4218, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).

- Stallard, R. F. and S. F. Murphy, 2014. A Unified Assessment of Hydrologic and Biogeochemical Responses in Research Watersheds in Eastern Puerto Rico Using Runoff-Concentration Relations. *Aquat. Geochem.* 20:115-139, DOI: 10.1007/s10498-013-9216-5.
- Susquehanna River Basin Commission, 2014. Sediment and nutrient assessment program. <http://www.srbc.net/programs/cbp/nutrientprogram.htm>.
- Thompson, J. J. D., D. G. Doody, R. Flynn and C. J. Watson, 2012. Dynamics of critical source areas: does connectivity explain chemistry? *Sci. Total Environ.* 435-436:499-508, DOI: 10.1016/j.scitotenv.2012.06.104.
- Thompson, S. E., N. B. Basu, J. Lascurain, A. Aubeneau and P. S. C. Rao, 2011. Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients. *Water Resour. Res.* 47:W00J05, DOI: 10.1029/2010WR009605.
- U.S. Geological Survey, 2014. Surface-water data for the nation.
- Walter, R. C. and D. J. Merritts, 2008. Natural streams and the legacy of water-powered mills. *Science* 319:299-304, DOI: 10.1126/science.1151716.
- Wolf, K. L., G. B. Noe and C. Ahn, 2013. Hydrologic Connectivity to Streams Increases Nitrogen and Phosphorus Inputs and Cycling in Soils of Created and Natural Floodplain Wetlands. *J. Environ. Qual.* 42:1245-1255, DOI: 10.2134/jeq2012.0466.
- Zhang, Q. and W. P. Ball, 2016. Data associated with Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-Quality Data. Johns Hopkins University Data Archive, Baltimore, MD. <http://dx.doi.org/10.7281/T18G8HM0>, DOI: 10.7281/T18G8HM0.

- Zhang, Q., W. P. Ball and D. L. Moyer, 2016a. Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns. *Sci. Total Environ.* 563-564: 1016-1029, DOI: 10.1016/j.scitotenv.2016.03.104.
- Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI: 10.1016/j.scitotenv.2013.02.012.
- Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.
- Zhang, Q., C. J. Harman and W. P. Ball, 2016b. An Improved Method for Interpretation of Riverine Concentration-Discharge Relationships Indicates Long-Term Shifts in Reservoir Sediment Trapping, manuscript under review.
- Zhang, Q., R. M. Hirsch and W. P. Ball, 2016c. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* 50:1877-1886, DOI: 10.1021/acs.est.5b04073.

*Page intentionally left blank*

## **Chapter 9. Improving Riverine Constituent Concentration and Flux Estimation by Accounting for Antecedent Discharge Conditions<sup>21</sup>**

### **Abstract**

Regression-based approaches are often employed to estimate riverine constituent concentrations and fluxes based on typically sparse concentration observations. One such approach is the recently-developed WRTDS (“Weighted Regressions on Time, Discharge, and Season”) method, which has been shown to provide more accurate estimates than prior approaches in a wide range of applications. Centered on WRTDS, this work was aimed at developing improved models for constituent concentration and flux estimation by accounting for antecedent discharge conditions. Twelve modified models were developed and tested, each of which contains one additional flow variable to represent antecedent conditions and which can be directly derived from the daily discharge record. High-resolution (~ daily) data at nine diverse monitoring sites were used to evaluate the relative merits of the models for estimation of six constituents – chloride (Cl), nitrate-plus-nitrite (NO<sub>x</sub>), total Kjeldahl nitrogen (TKN), total phosphorus (TP), soluble reactive phosphorus (SRP), and suspended sediment (SS). For each site-constituent combination, 30 concentration subsets were generated from the original data through Monte Carlo sub-sampling and then used to evaluate model performance. For the sub-sampling, three sampling strategies were adopted: (A) 1 random sample each month

---

<sup>21</sup> This chapter will be submitted for publication in a peer-reviewed journal. Bill Ball was involved in hypothesis development, study design, results interpretation, and editing. All figures, tables, and data were created by Qian Zhang.

(12/year), (B) 12 random monthly samples plus additional 8 random samples per year (20/year), and (C) flow-stratified sampling with 12 regular (non-storm) and 8 storm samples per year (20/year). Simulation results reveal that estimation performance varies with both model choice and sampling strategy. In terms of model choice, the modified models show general improvement over the original model under all three sampling strategies. Major improvements were achieved for NO<sub>x</sub> by the long-term flow-anomaly model and for Cl by the ADF (average discounted flow) model and the short-term flow-anomaly model. Moderate improvements were achieved for SS, TP, and TKN by the ADF model. By contrast, no such achievement was achieved for SRP by any proposed model. In terms of sampling strategy, performance of all models (including the original) was generally best using strategy C and worst using strategy A, and especially so for SS, TP, and SRP, confirming the value of routinely collecting storm-flow samples. Overall, this work provides a comprehensive set of statistical evidence for supporting the incorporation of antecedent discharge conditions into the WRTDS model for estimation of constituent concentration and flux, thereby combining the advantages of two recent developments in water quality modeling.

### **9.1. Introduction**

In many river monitoring programs, flow discharge is usually recorded daily but water-quality constituent concentration is sampled at most only a few times each month. For example, for some major sites in the Chesapeake Bay watershed (USA), a minimum of 20 water-quality samples have been collected annually since the 1980s, usually comprising at least 12 regular monthly samples and 8 storm-flow samples (Chanat *et al.*, 2016; Zhang *et al.*, 2016a). In using such sparse data toward load estimation, it is

important to make best possible estimates of concentration for each day in the record, thus enabling calculation of daily flux (“load”) as the product of daily concentration and daily discharge. From a management perspective, such daily estimates and their temporal aggregates (monthly, seasonal, or annual averages) play an essential role toward assessment against regulatory standards (*e.g.*, maximum concentration level), establishment and assessment of restoration targets (*e.g.*, total maximum daily load), calibration of watershed models, and evaluation of long-term trends in river quality (Stow and Borsuk, 2003; Bowes *et al.*, 2008; Shenk and Linker, 2013; Zhang *et al.*, 2015).

Over the last few decades, alternative regression-based approaches have been developed by various investigators, including but not limited to, Dolan *et al.* (1981); Cohn *et al.* (1989); Cohn *et al.* (1992); Kronvang and Bruhn (1996); Crowder *et al.* (2007); Johnes (2007); Stenback *et al.* (2011); and Park and Engel (2015). These approaches generally estimate daily concentrations and fluxes based on modeled relations between observed concentration and some set of explanatory variables, which include typically time, discharge, and season. With these covariates, a single regression model is usually established for the entire record based on common assumptions of homoscedasticity of model errors and fixed relations between concentration and each covariate. These assumptions, however, can be frequently violated in reality – see examples in Hirsch *et al.* (2010).

With consideration of the above limitations, Hirsch *et al.* (2010) developed a method called “Weighted Regressions on Time, Discharge, and Season (WRTDS)”. Similar to the prior approaches, WRTDS makes use of time, discharge, and season as explanatory variables but has a major distinction in that it develops a separate regression model for

each day in the record. For each day, it re-evaluates the dependencies of concentration on time, discharge, and season based on data deemed to be most relevant to the day of estimation (Hirsch *et al.*, 2010; Hirsch and De Cicco, 2015). Consequently, WRTDS can better represent the temporally-varying seasonal and discharge-related patterns and is more resistant to problems of flux-estimation bias than prior methods (Moyer *et al.*, 2012; Hirsch, 2014). Since its publication in 2010, WRTDS has been applied in a suite of large-scale water-quality studies (Sprague *et al.*, 2011; Hirsch, 2012; Medalie *et al.*, 2012; Moyer *et al.*, 2012; Zhang *et al.*, 2013; Murphy *et al.*, 2014; Pellerin *et al.*, 2014; Corsi *et al.*, 2015; Zhang *et al.*, 2015; Chanut *et al.*, 2016; Zhang *et al.*, 2016a; Zhang *et al.*, 2016b).

Although WRTDS represents an advance relative to prior approaches, it is of course eligible for further improvement. In this context, the original authors suggested that additional explanatory variables, perhaps especially those that can account for watershed wetness conditions, may further enhance the model (Hirsch *et al.*, 2010). Mechanistically, antecedent conditions can influence both the momentum and pathways of constituent transport from watershed to stream. In a relatively dry watershed, particulate constituents tend to be kept locally and dissolved constituents may enter the subsurface near their point of application or generation. When the watershed becomes wetter, particulate constituents can be more easily mobilized with surface runoff and dissolved constituents can be more efficiently flushed to stream due to saturation of subsurface storage (Cirino and McDonnell, 1997; Inamdar *et al.*, 2004; Dupas *et al.*, 2015). In the last two decades, an extensive set of literature has documented the effects of antecedent condition on riverine export of constituents, including nitrogen (Biron *et al.*, 1999; Turgeon and

Courchesne, 2008; Macrae *et al.*, 2010; Davis *et al.*, 2014; Murphy *et al.*, 2014), phosphorus (McDowell and Sharpley, 2002; Ide *et al.*, 2008; Warner *et al.*, 2009; Macrae *et al.*, 2010), suspended sediment (McDowell and Sharpley, 2002; Gray *et al.*, 2015), and dissolved organic carbon (Biron *et al.*, 1999; Turgeon and Courchesne, 2008). Of special relevance to the current work, Vecchia (2003) has developed the flow anomaly concept for quantification of antecedent discharge conditions, which has been subsequently used and refined in numerous analyses of surface-water quality (Vecchia, 2005; Vecchia *et al.*, 2008; Vecchia *et al.*, 2009; Ryberg and Vecchia, 2014). In the context of WRTDS, Murphy *et al.* (2014) have previously considered Vecchia's approach, but that work was a preliminary effort of limited scope and depth, with evaluation of only a limited set of antecedent condition variables and with a particular focus on nitrate transport.

The overall objective of the study described here was to substantially expand upon the work of Murphy *et al.* (2014) by developing and rigorously testing several alternative statistical models that build on the WRTDS model by also accounting for antecedent discharge conditions, thereby providing improved estimation of riverine concentration and flux. For method transferability, we focused on variables that can be directly derived from the daily river discharge record, since such record is typically available regardless of watershed characteristics or geographic regions. These variables include flow anomalies on various temporal scales (Vecchia *et al.*, 2009; Ryberg and Vecchia, 2014), average discounted flow (Wang *et al.*, 2011; Wang and Tian, 2013), base-flow index (Lyne and Hollick, 1979; Nathan and McMahon, 1990), and flow gradient. Six constituents were evaluated, namely, chloride (Cl), nitrate-plus-nitrite (NO<sub>x</sub>), total Kjeldahl nitrogen (TKN), total phosphorus (TP), soluble reactive phosphorus (SRP), and

suspended sediment (SS), using high-resolution (nearly daily) data at nine diverse monitoring sites. For each site-constituent combination, concentration subsets were generated through Monte Carlo sub-sampling and then used to evaluate the performance of the original and modified models. The sub-sampling was conducted based on three different sampling strategies to mimic common monitoring practice.

## 9.2. Statistical Models: Original and Modified WRTDS

The original WRTDS model uses the following principle equation (Hirsch *et al.*, 2010):

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (9.1)$$

where  $C$  is constituent concentration,  $t$  is time in decimal years,  $Q$  is daily discharge,  $\beta_i$  are fitted coefficients, and  $\varepsilon$  is the error term. On the right side of the equation, the 2<sup>nd</sup> and 3<sup>rd</sup> terms represent time and discharge effects, respectively, while the 4<sup>th</sup> and 5<sup>th</sup> terms represent seasonal effects. The log-log relationship is adopted for three reasons. First, it provides a generally better fit than a linear relationship. Second, the residuals from this model are more nearly normal and homoscedastic. Third, it avoids the possibility of obtaining negative values for estimated concentration (Hirsch and De Cicco, 2015). For each day in the record, WRTDS establishes one separate regression model to estimate concentration and flux. First, it pre-screens the entire concentration data and selects at least 100 samples that are sufficiently “close” to the estimation day in three dimensions, *i.e.*, time, season, and discharge. For these three dimensions, default half-window widths are 7 years, 0.5 year, and 2 natural log units, respectively. Second, the selected samples are used to build a weighted regression model based on Equation 9.1. Third, the fitted coefficients are used to estimate  $\ln(C)$  on the estimation day by

substituting known values of  $t$  and  $Q$ , which is then transformed to concentration in real space using a proper bias correction factor. (Note that the set of fitted coefficients is unique for each day of estimation.) Finally, the estimated  $C$  is multiplied by daily  $Q$  to obtain daily flux. Full details of the above process are described in Hirsch and De Cicco (2015).

To account for antecedent discharge conditions, we modified Equation 9.1 by adding one of several alternative flow variables. For model parsimony, only one new variable was added in each modified model, as represented by  $X$  in Equation 9.2:

$$\ln C = \beta_0 + \beta_1 t + \beta_2 \ln Q + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \beta_5 X + \varepsilon \quad (9.2)$$

The variables for  $X$  include flow anomalies (FA) on various temporal scales, average discounted flow (ADF), base-flow index (BFI), and flow gradient ( $dQ/dt$ ), all of which can be directly derived from the daily discharge record. These terms are defined below in Equations 9.3-9.15. In general, antecedent discharge variables quantified at a river monitoring site (*i.e.*, the outlet of a watershed) can be used as a proxy for antecedent moisture conditions across the watershed by amalgamating highly-variable soil water and groundwater storage conditions caused by spatially-varying climatic, hydrogeologic, and land use conditions across the watershed (Murphy *et al.*, 2014). Depending upon the characteristics of the watershed and the water-quality variable of interest, one might expect different time-scales of antecedent condition to be relevant, such that different model selections may be more appropriate for different constituent-watershed combinations.

The FA term of Equation 9.2 quantifies how different (how “anomalous”) the antecedent discharge conditions actually are for the selected temporal scale of the FA

term and for the given day on which the analysis is conducted. For any given day, the FA term can be defined with the following general form (Vecchia *et al.*, 2009; Ryberg and Vecchia, 2014):

$$FA(t) = X_{T_1}(t) - X_{T_2}(t) \quad (9.3)$$

where  $X_{T_1}(t)$  and  $X_{T_2}(t)$  are averages of log-transformed daily discharge for periods of  $T_1$  and  $T_2$  preceding a given day  $t$ , with  $T_1$  as the shorter of the two periods. With this definition, a positive value of FA corresponds to a situation in which the shorter-term (more recent) period has been “wetter” (has experienced more average rainfall and riverflow) than the longer-term (precedent) period. In this work, flow anomalies were quantified for varying groups of temporal scales using the *R* package “*waterData*” (Ryberg and Vecchia, 2014):

$$\text{Long-term FA: } LTF A(t) = X_{1 \text{ year}}(t) - X_{\text{entire period}}(t) \quad (9.4)$$

$$\text{Mid-term FA: } MTF A(t) = X_{1 \text{ month}}(t) - X_{1 \text{ year}}(t) \quad (9.5)$$

$$\text{Short-term FA: } STF A(t) = X_{\text{current day}}(t) - X_{1 \text{ month}}(t) \quad (9.6)$$

$$\text{Annual FA: } AnnualFA(t) = X_{1 \text{ year}}(t) - X_{5 \text{ year}}(t) \quad (9.7)$$

$$\text{Seasonal FA: } SeasonFA(t) = X_{1 \text{ season}}(t) - X_{1 \text{ year}}(t) \quad (9.8)$$

$$\text{Daily FA: } DailyFA(t) = X_{\text{current day}}(t) - X_{1 \text{ season}}(t) \quad (9.9)$$

$$\text{100-day FA: } FA100(t) = X_{100 \text{ day}}(t) - X_{\text{entire period}}(t) \quad (9.10)$$

$$\text{10-day FA: } FA10(t) = X_{10 \text{ day}}(t) - X_{100 \text{ day}}(t) \quad (9.11)$$

$$\text{1-day FA: } FA1(t) = X_{\text{current day}}(t) - X_{10 \text{ day}}(t) \quad (9.12)$$

These flow anomalies were quantified for each day in the period of record, as demonstrated with Maumee River time series in Figure 9.1b-9.1j. For this work, the Maumee is one of several Ohio sites (described further in the following section) for

which concentration observations are available on almost all days – such sites provide excellent test cases for the work at hand because WRTDS model estimates (based on systematically generated sub-samples of original data) can be tested against actual observations. For the work described herein, all of the alternative FA terms of Equations 9.4 to 9.12 were considered in order to test which would most appropriately represent antecedent conditions for the various constituents in a given watershed.

The ADF term was originally proposed by Wang *et al.* (2011) to incorporate the influence of historical flows into flux estimation:

$$ADF(t) = \frac{\sum_{j=1}^J d^{t-t_j} Q_{t_j}}{\sum_{j=1}^J d^{t-t_j}} \quad (9.13)$$

Where  $J$  is the maximum number of daily discharge observations in the record,  $t$  is the day of estimation,  $t_j$  is a historical day in relative to  $t$  (*i.e.*,  $t_j < t$ ) and  $t - t_j$  is the lag time between  $t$  and  $t_j$  ( $1 \leq j \leq J$ ), and  $d$  is the discounting rate. In this work,  $d$  was taken to be 0.95 per day (~0.5 per fortnight). An example time series of ADF is given in Figure 9.1k. So far, the development and application of ADF has been limited to sediment (Wang *et al.*, 2011; Wang and Tian, 2013; Wang *et al.*, 2015). In this work, we seek to compare the effectiveness of ADF and the other flow terms for six common constituents.

Operationally, the logarithmic form of ADF was taken as the  $X$  term in Equation 9.2.

The BFI is simply the ratio of baseflow to total stream flow:

$$BFI(t) = BF(t)/Q(t) \quad (9.14)$$

where  $BF(t)$  is the baseflow on day  $t$ , which was calculated using the Lyne-Hollick filter method (Lyne and Hollick, 1979; Nathan and McMahon, 1990) through the *R* package “*EcoHydRology*” (Fuka *et al.*, 2014). An example time series of BFI is given in Figure 9.1l. Operationally, the logarithmic form of BFI was taken as the  $X$  term in Equation 9.2.

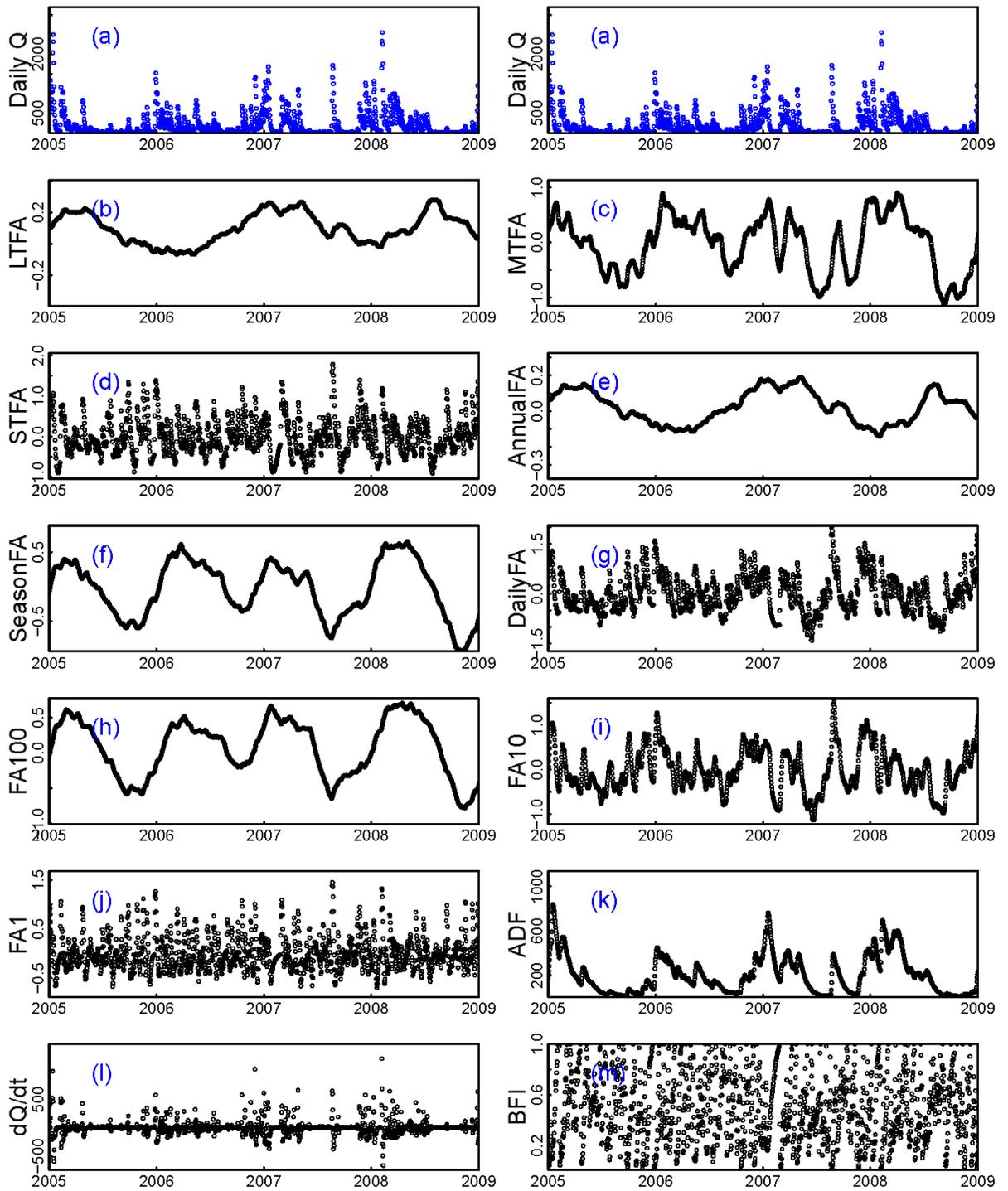


Figure 9.1. Time series of (a) daily discharge and (b-m) the derived flow variables for Maume River at Waterville, OH between 2005 and 2009.

As a measure of the slow continuous contribution of groundwater, BFI may be effective in capturing dynamics of subsurface solute transport, particularly Cl and NO<sub>x</sub> (Böhlke and Denver, 1995; Kirchner *et al.*, 2001; Sanford and Pope, 2013; Harman, 2015).

The last alternative variable for  $X$  in equation 9.2 is  $dQ/dt$ , which is simply the first derivative of the daily discharge series at the day of interest:

$$dQ/dt(t) = Q_t - Q_{t-1} \quad (9.15)$$

An example time series of  $dQ/dt$  is given in Figure 9.1m. This term may be effective in capturing the effect of hysteresis, *i.e.*, concentration may vary significantly depending on whether the sampling is occurring on the rising or falling limb of the hydrograph (Evans and Davies, 1998; Bierzoza and Heathwaite, 2015).

### **9.3. Data and Analysis**

#### **9.3.1. Testing Sites and Data**

The main body of our work used the above described approaches with high-resolution (nearly-daily) water-quality constituent concentration data for several tributaries to Lake Erie and Ohio River, as obtained from the National Center for Water Quality Research at Heidelberg University, Ohio, USA (National Center for Water Quality Research, 2015). Data were obtained for nine sites with at least 15 years of samples: Maumee River, Scioto River, Great Miami River, Sandusky River, River Raisin, Cuyahoga River, Grand River, Honey Creek, and Rock Creek (Table 9.1). These sites have all been frequently sampled and cover a wide range of drainage area (34.6 ~ 6,330 mi<sup>2</sup>), agriculture land (30.4% ~ 89.9%), wooded land (7.3% ~ 50.1%), and urban land (0.6% ~ 9.6%), thus allowing a test of the extent to which each model's results are site-specific or generally applicable and for an initial consideration of the effect of watershed size and characteristics on model

Table 9.1. Details of the nine river monitoring sites.

Station Number	River Site	Drainage area, mi <sup>2</sup>	Upstream land use (percent)			
			Agricultural	Urban	Wooded	Other
04193500	Maumee River at Waterville, OH	6,330	89.9	1.2	7.3	1.6
03231500	Scioto River at Chillicothe, OH	3,854	80.2	4.6	12.9	2.3
03271601	Great Miami River below Miamisburg, OH	2,685	82.1	4.7	10.3	2.9
04198000	Sandusky River near Fremont, OH	1,251	84.1	0.9	13.0	2.0
04176500	River Raisin near Monroe, MI	1,042	79.0	2.3	14.0	4.7
04208000	Cuyahoga River at Independence, OH	707	30.4	9.6	50.1	9.9
04212100	Grand River near Painesville, OH	685	40.0	0.9	45.2	13.1
04197100	Honey Creek at Melmore, OH	149	85.6	0.6	12.5	1.3
04197170	Rock Creek at Tiffin, OH	34.6	82.0	0.9	16.1	1.0

selection. For each site, daily river discharge data were directly imported from the U.S. Geological Survey National Water Information System (U.S. Geological Survey, 2015) using the *R* package “*dataRetrieval*” (Hirsch and De Cicco, 2015). Time series of the proposed flow variables were then obtained using the respective definitions and associated *R* packages described in Section 9.2.

In addition, concentration data were compiled for each site for six constituents, *i.e.*, Cl, NO<sub>x</sub>, TKN, TP, SRP, and SS. For each site-constituent combination, concentration data were imported from pre-prepared files using the *R* package “*dataRetrieval*” (Hirsch and De Cicco, 2015) and removed any non-positive values. For days with multiple concentration samples, their median value is used. The average number of sampled days ranges between 256 and 349 per year. (See details in Table 9.2.)

Table 9.2. Temporal coverages of water-quality data at the nine study sites.

River Site	First Sampled Day	Last Sampled Day	Total Number of Sampled Days (Average Number of Sampled Days Per Year)					
			Cl	NO <sub>x</sub>	TKN	SRP	TP	SS
			Maumee River at Waterville, OH	1975-01-10	2014-12-31	11,447 (345)	10,898 (328)	11,397 (343)
Scioto River at Chillicothe, OH	1996-04-23	2015-01-03	4828 (258)	4816 (258)	4812 (257)	4785 (256)	4825 (258)	4803 (257)
Great Miami River below Miamisburg, OH	1996-04-22	2015-04-04	6356 (335)	6357 (335)	6309 (333)	6317 (333)	6349 (335)	6333 (334)
Sandusky River near Fremont, OH	1974-10-02	2014-12-31	12,055 (301)	12,054 (300)	10,719 (275)	11,319 (281)	12,363 (307)	12,315 (306)
River Raisin near Monroe, MI	1982-03-06	2014-12-31	8,268 (281)	8,213 (279)	8,228 (280)	7,621 (259)	8,269 (281)	8,236 (280)
Cuyahoga River at Independence, OH	1981-11-04	2014-12-31	10,856 (327)	10,890 (329)	10,838 (327)	10,627 (321)	10,896 (329)	10,862 (328)
Grand River near Painesville, OH	1988-02-16	2006-08-10	5,621 (304)	5,435 (294)	5,545 (300)	4,977 (269)	5,598 (303)	5,565 (301)
Honey Creek at Melmore, OH	1976-01-28	2014-12-31	11,604 (342)	11,617 (343)	11,353 (342)	10,647 (314)	11,623 (343)	11,499 (339)
Rock Creek at Tiffin, OH	1982-10-11	2014-12-31	9,249 (349)	8,911 (336)	9,202 (347)	8,697 (328)	9,251 (349)	9,191 (347)

### 9.3.2. Monte Carlo Sub-sampling

To mimic the roughly monthly sampling frequency in many monitoring programs, I performed Monte Carlo sub-sampling of the original concentration data to produce concentration subsets for subsequent model evaluation. Three common sampling strategies were considered:

- (A) One concentration sample was randomly selected for each calendar month for which samples were available (*i.e.*, 12 samples per year),
- (B) 12 random monthly samples plus additional 8 random samples per year (*i.e.*, 20

samples per year), and

(C) Flow-stratified sampling with 12 regular samples and 8 storm samples per year (20 samples per year). Within each specific year of the record, samples with daily discharge exceeding the 90<sup>th</sup> percentile of daily discharge distribution of the year are marked as “storm sample”; all other samples are considered as “regular samples”. This sampling strategy has been adopted for many river sites within the Chesapeake Bay watershed; see details in Chanat *et al.* (2016).

With each sampling strategy, 30 concentration subsets were independently generated for each constituent and for each site. The original concentration data and the three types of sub-samples are illustrated in Figure 9.2.

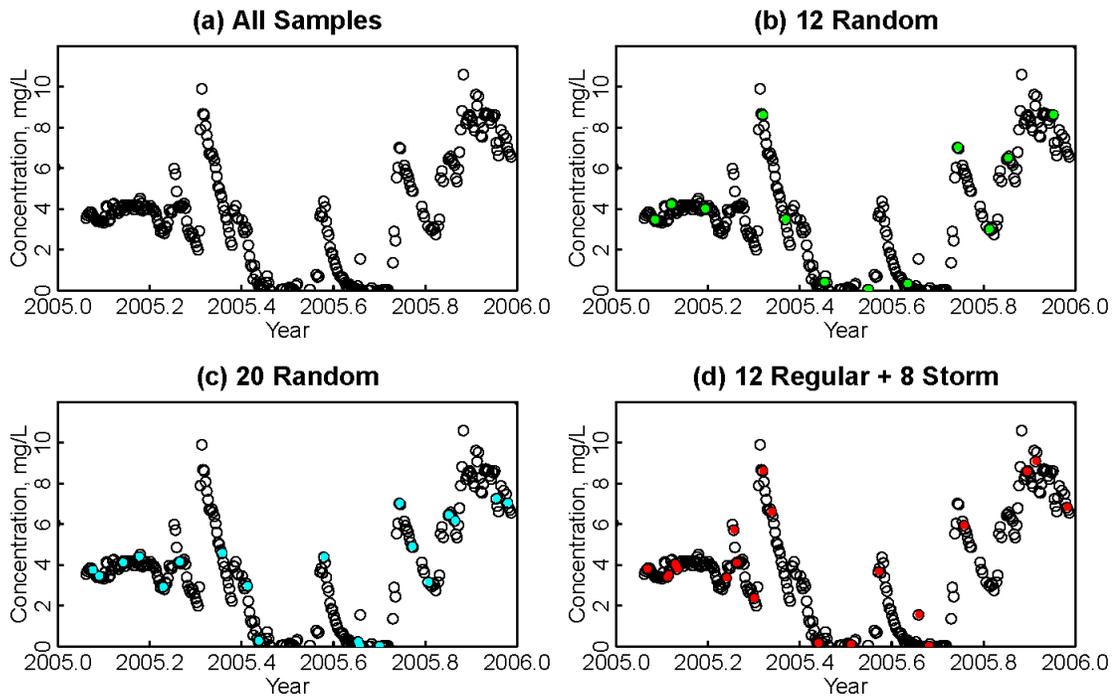


Figure 9.2. Example time series of (a) the original NO<sub>x</sub> concentration data (black open circle) and (b-d) three types of concentration subsets (non-black solid circle) in Maumee River at Waterville, OH between 2005 and 2006. See text for details of the three sampling strategies.

To formally assign uncertainties on estimates from the original WRTDS models, one approach is to use the bootstrap method of Hirsch *et al.* (2015), which involves resampling (with replacement) of the original concentration data to obtain many realizations of representative sets and associated estimates. (See Zhang *et al.* (2016b) for examples.) In this work, we used 30 subsets in order to reduce the required total computational time, whilst providing a sufficient number of replicates to reasonably describe overall uncertainties in each model’s performance.

### 9.3.3. Model Evaluation

With each concentration subset, the original and modified models (13 in total) were run to estimate daily concentration and flux, which was implemented in *R* using codes modified after the existing WRTDS package “*EGRET*” (Hirsch and De Cicco, 2015).

These estimates were then compared against observations for calculation of model residuals ( $Y_i^{residual}$ ):

$$Y_i^{residual} = Y_i^{obs} - Y_i^{est} \quad (9.16)$$

where  $Y_i^{obs}$  and  $Y_i^{est}$  are the observed and modeled values for the  $i^{\text{th}}$  day, respectively.

The residuals can be interpreted as the portion of concentration (or flux) not accounted for by a specific model. Positive and negative residuals indicate under- and over-estimation, respectively. Spearman’s correlation coefficient ( $\gamma$ ) was also quantified to examine structural relationships between the residuals from the original WRTDS model and each of the proposed flow variables (Helsel and Hirsch, 2002). Intuitively, a high correlation coefficient may imply that the flow variable has a high potential for improving the estimation performance.

To facilitate performance comparison among the 13 models, three common numerical measures were considered – namely, Nash-Sutcliffe efficiency (NSE), modified NSE (E), and percent bias (PBIAS), as defined below (Krause *et al.*, 2005; Moriasi *et al.*, 2007):

$$NSE = 1 - \left[ \frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{est})^2}{\sum_{i=1}^n (Y_i^{obs} - Y_i^{mean})^2} \right] \quad (9.17)$$

$$E = 1 - \left[ \frac{\sum_{i=1}^n |Y_i^{obs} - Y_i^{est}|}{\sum_{i=1}^n |Y_i^{obs} - Y_i^{mean}|} \right] \quad (9.18)$$

$$PBIAS = \left[ \frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{est})}{\sum_{i=1}^n (Y_i^{obs})} \right] \times 100 \quad (9.19)$$

where  $Y_i^{obs}$  and  $Y_i^{est}$  are defined as above,  $Y_i^{mean}$  is the average of observed data in the record, and  $n$  is the total number of observations. NSE quantifies the squared difference between the estimated and observed values normalized by the variance of observations. A model that perfectly matches available observations has  $NSE = 1$ , with values  $< 0$  representing inferior levels of fit (Krause *et al.*, 2005; Moriasi *et al.*, 2007). Despite being the most common measure, NSE values are very sensitive to large errors (*i.e.*, large differences between data and model fits) due to the square operation. To deal with this issue, the second measure E uses absolute differences instead of squared differences. As with NSE, a perfectly accurate model has  $E = 1$ . Unless otherwise noted, results documented herein are based on E. Finally, PBIAS measures the average tendency of the estimates to be larger or smaller than observations. For model(s) that precisely match available data,  $PBIAS = 0$ . For each constituent (6 in total) at each site (9 in total), the three performance measures were calculated for each model (13 in total) in each replication runs (30 for each of three sampling strategies).

#### **9.4. Results and Discussion**

This section is organized as follows. In Section 9.4.1, I describe two cases where modified models show major improvement over the original model. In Section 9.4.2, I

present the overall performance of all models based on simulation results at all nine sites. Focus here is on the effects of model choices on estimation performance. For brevity, the presented simulation results are associated with sampling strategy “A” unless otherwise noted. In Section 9.4.3, I discuss the effects of sampling strategies on model performance.

#### **9.4.1. Selected Cases with Major Improvement in Performance**

##### ***9.4.1.1. Case 1: LTFA Model for NO<sub>x</sub> Estimation in Grand River***

The first case considered here is NO<sub>x</sub> in Grand River. For this parameter, concentration residuals from the original WRTDS model appear to fluctuate around zero over time, indicating low bias for the original model (Figure 9.3a). There exists, however, a clear negative correlation (Spearman’s  $\gamma = -0.27$ ) between the original model residuals and values of the LTFA (Figure 9.3b). In other words, large LTFA values correspond to negative residuals from the original model, and vice versa. Such correlation is likely reflective of a “storage-and-then-release” effect. More specifically, for days with small LTFAs, the watershed is relatively dry in the preceding 1-year and NO<sub>x</sub> is more likely to have accumulated during the dry period (*i.e.*, less regular “flushing”). Consequently, observed NO<sub>x</sub> concentrations on such days (with small LTFAs) tend to be higher for a given flow than those predicted by the original model. Conversely, for days with large LTFAs, relatively wet watershed conditions existed in the preceding 1-year (relative to the entire record), and this likely corresponds with more regular “flushing” of contaminants and less storage. Hence, observed concentrations on these days are likely to be lower than those predicted by the original model.

Given the above considerations, we hypothesized that the original WRTDS model can be improved by adding LTFA per Equation 9.2. This hypothesis was validated by

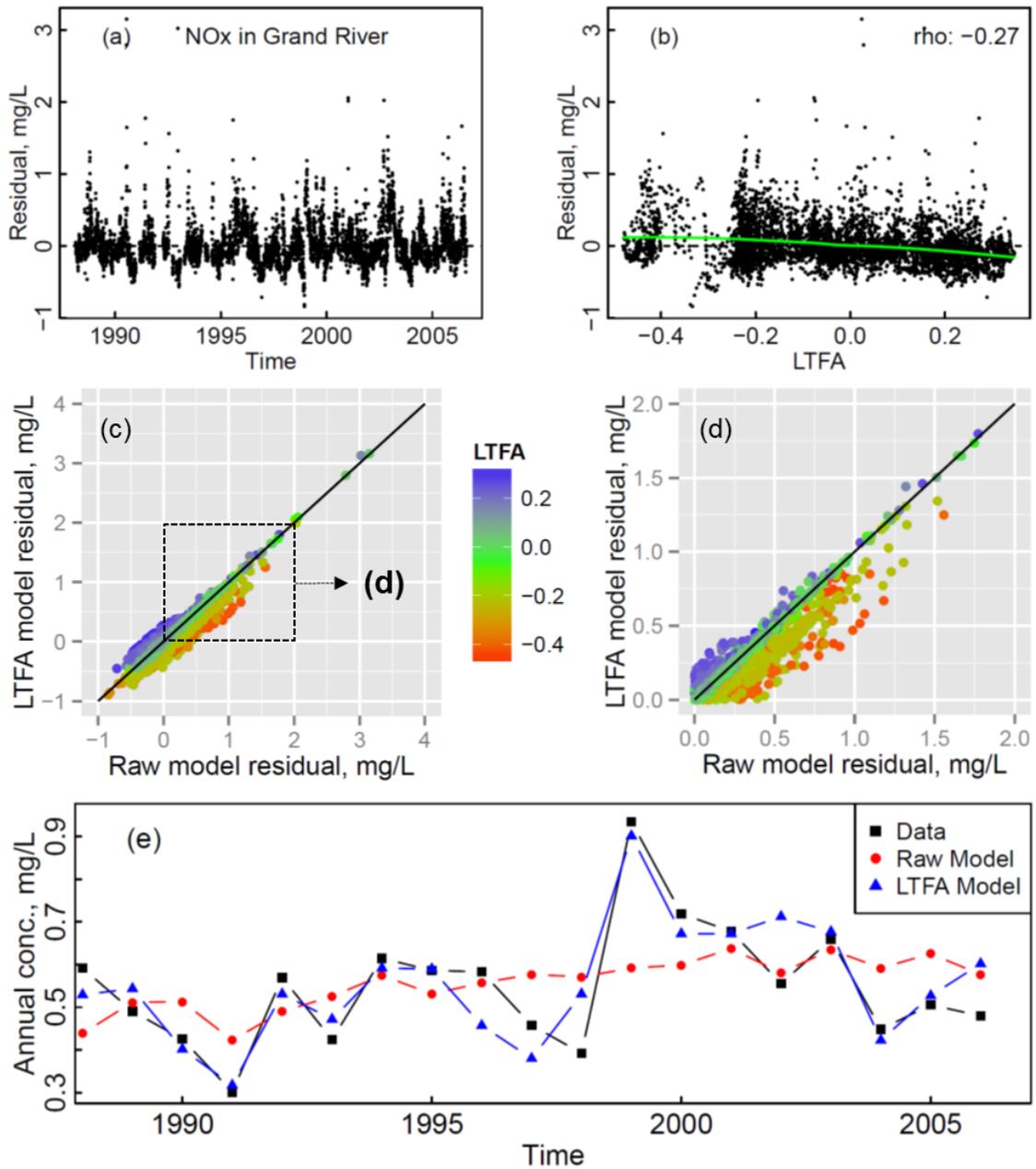


Figure 9.3. Comparison of the original WRTDS model and the *LTFA* (*long-term flow anomaly*) model for NO<sub>x</sub> in Grand River near Painesville, OH. (a) Time series of residuals from the original model. (b) Relationship between the original model residuals and LTFA, as modeled with LOWESS curve (green line). Spearman's correlation ( $\gamma$ ) is shown on the plot. (c) Comparison between the LTFA model residuals and the original model residuals. (d) Enlarged view of the first quadrant of plot (c). (e) Time series of annual concentration based on observed data, the original model, and the LTFA model.

comparing residuals of the LTFA model and those of the original model (Figure 9.3c). A more focused look at the majority of data with residual values in the range of 0 – 2.0 mg/L (Figure 9.3d) clearly reveals more points below the 1:1 reference line, reflecting smaller positive residuals (*i.e.*, less under-estimation) for the LFTA model. Such improvements are achieved primarily for days with highly negative LTFAs, as shown by red (as opposed to green) colors in Figure 9.3c and 9.3d. This is consistent with the discussion above that the original model tends to under-estimate concentrations for days with strongly negative LTFAs. By taking into account the antecedent discharge effect, the LTFA model effectively alleviates that issue. Conversely, for days with positive LTFAs, the LTFA model appears to exacerbate the issue of over-estimation; however, such undesirable cases are much less frequent. Moreover, improvement with the LTFA model is reflected by the time series of annual flow-weighted concentration – compared with the original model, the LTFA model shows annual NO<sub>x</sub> concentration estimates that follow observed concentrations much more closely, with markedly improved modeling of extreme concentrations (Figure 9.3e).

Overall, the LTFA model provides major improvement over the original model for NO<sub>x</sub> estimation in Grand River, as quantified using numerical measures in Table 9.3. Compared with the original model, the LTFA model has larger NSE and E for both concentration and flux. In addition, the LTFA model has much a smaller absolute PBIAS for flux.

#### ***9.4.1.2. Case 2: STFA Model for Cl Estimation in Maumee River***

The second case considered here is that of Cl in Maumee River. Similar to the case above, concentration residuals from the original model appear to fluctuate around zero

Table 9.3. Summary of cases with major improvements in estimation performance. Each case shows results that are based on one selected model run.

Cases	Models	Concentration			Flux		
		NSE	E	PBIAS	NSE	E	PBIAS
Grand River NO <sub>x</sub>	Raw Model	0.28	0.18	2.3	0.73	0.63	-2.2
	LTFA Model	0.41	0.27	1.7	0.83	0.71	0.83
Maumee River Cl	Raw Model	0.54	0.36	0.46	0.83	0.72	1.7
	STFA Model	0.77	0.54	0.43	0.89	0.77	1.3

(Figure 9.4a). In this case, however, there exists a clear positive correlation ( $\gamma = 0.48$ ) between the original model residuals and STFAs (Figure 9.4b). In other words, small STFAs correspond to small residuals from the original model. Note that in the case of STFA, the comparison is between the current day (as  $X_{T_1}$  in Equation 9.3) and the monthly antecedent discharge (as  $X_{T_2}$  in Equation 9.3), which can be contrasted with the LTFA, where the annual antecedent discharge (as  $X_{T_1}$ ) is compared to the entire discharge record (as  $X_{T_2}$ ). Hence, with STFA, a positive (rather than negative) residual corresponds to relatively wet day but a drier antecedent month, whereas with LTFA, negative residual corresponds to a dry antecedent year relative to the full record. As a result, the positive STFA correlation is what now suggests the dry antecedent conditions of a “storage-and then-release” effect, as opposed to the negative LTFA correlation for the NO<sub>x</sub> case above. In particular, a small STFA now represents a relatively wet condition in the preceding 1-month and is believed to reflect a period with sustained “flushing” or loss of storage. Consequently, observed Cl concentrations on such days tend to be lower than those predicted by the original model (*i.e.*, negative residuals). These interpretations are consistent with the case of LTFA above. Other analyses (results summarized subsequently) revealed that STFA was less effective than LTFA for

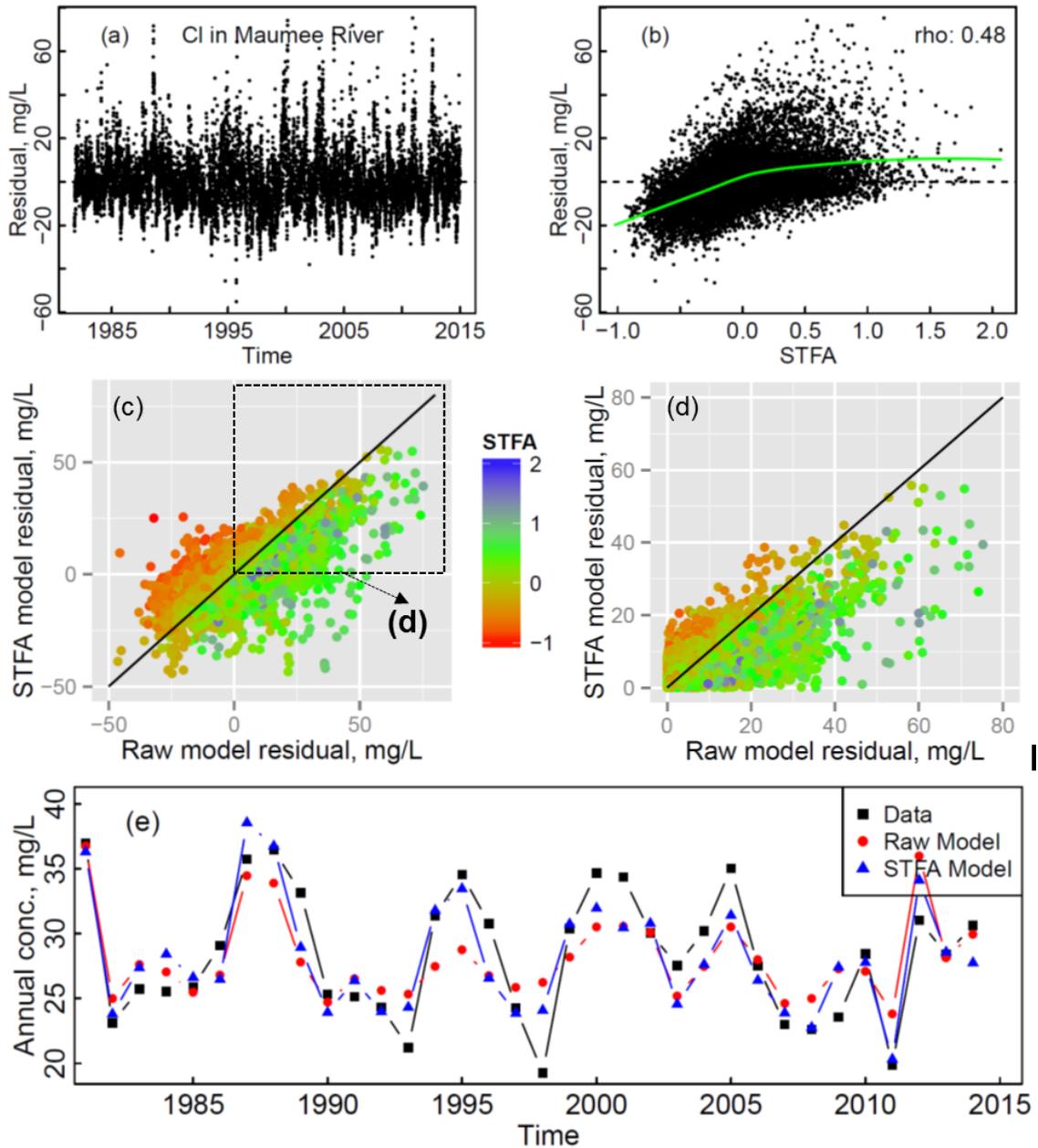


Figure 9.4. Comparison of the original WRTDS model and the *STFA* (*short-term flow anomaly*) model for Cl in Maumee River at Waterville, OH. (a) Time series of residuals from the original model. (b) Relationship between the original model residuals and STFA, as modeled with LOWESS curve (green line). Spearman's correlation ( $\gamma$ ) is shown on the plot. (c) Comparison between the STFA model residuals and the original model residuals. (d) Enlarged view of the first quadrant of plot (c). (e) Time series of annual concentration based on observed data, the original model, and the STFA model.

estimation of NO<sub>x</sub> at Grand River but that STFA was more effective than LTFA for estimation of Cl at Maumee. We hypothesize that these differences reflect differences in response times of the two different constituents in the two different watershed systems. In terms of watershed size alone, the finding might seem somewhat counter-intuitive, because Grand River is the smaller basin (Table 9.1) yet is better understood in terms of LTFA. These results suggest, perhaps not surprisingly, that parameters other than watershed size will control the time scale at which antecedent discharge is relevant toward regulating downstream concentrations. In this case, the constituent of interest (NO<sub>x</sub> vs. Cl) is also important, and a reasonable explanation might be that Cl response is more sensitive to short-term antecedent relative to the present discharge, whereas NO<sub>x</sub> processing is more sensitive to the past year conditions relative to the long-term record. The mechanistic underpinnings of these statistical results are a topic deserving of further investigation.

As with the Grand River LTFA NO<sub>x</sub> case, the residual-STFA correlation for Cl at Maumee led us to hypothesize that the original model could be improved by specifically including the flow anomaly (in this case, the STFA) into Equation 9.2. Again, the hypothesis was validated by the comparison – see Figure 9.4c-9.4d – and as before, there are evidently more points below the 1:1 reference line, which represent cases where the STFA model has smaller positive residuals (*i.e.*, less model under-estimation) than the original model. Such improvements are generally achieved for days with positive STFAs, as shown with green color. Thus, the STFA model effectively improves the estimation by taking into account the antecedent discharge effect. Conversely, the STFA model appears

to exacerbate the issue of over-estimation for days with negative STFAs, but again such undesirable cases are less frequent.

Moreover, one can again turn to a comparison of the time series of modeled versus observed values of annual flow-weighted Cl concentration as a good overall point of comparison. In this regard, Figure 9.4e reveals that the STFA model estimates follow observed annual average concentrations much more closely than do the original model estimates, with markedly improved modeling of extreme concentrations. Also, our quantitative measures of fit for the full daily record (shown in Table 9.3) again reveal that the modified model (now using STFA) has larger NSE and E for both concentration and flux. The STFA model also shows moderately smaller absolute PBIAS for both concentration and flux.

#### **9.4.2. Performance Comparison for All Simulations: Effects of Model Choices**

Here I present the overall performance of all models based on simulation results for all nine sites. For brevity, the presented simulation results are associated with sampling strategy A (unless otherwise noted). Conclusions are similar based on the other two strategies (see Figures G1-G2). For quantitative comparison with the original model, it is useful to calculate the ratio of E-values between each modified model and the original model for each replicate run and for each site (*i.e.*, 30 runs/sites x 9 sites = 270 cases). These 270 ratios are shown with boxplots in Figure 9.5. Simple guidelines for model evaluation were developed. Specifically, the modified models are considered to show (1) major improvement, (2) moderate improvement, and (3) inferior performance, respectively, if the E-value ratio falls to the following regions: (1) ratio > 1.2 (indicated by blue bars in Figure 9.5), (2)  $1 < \text{ratio} < 1.2$  (grey bars), and (3) ratio < 1.0 (red bars).

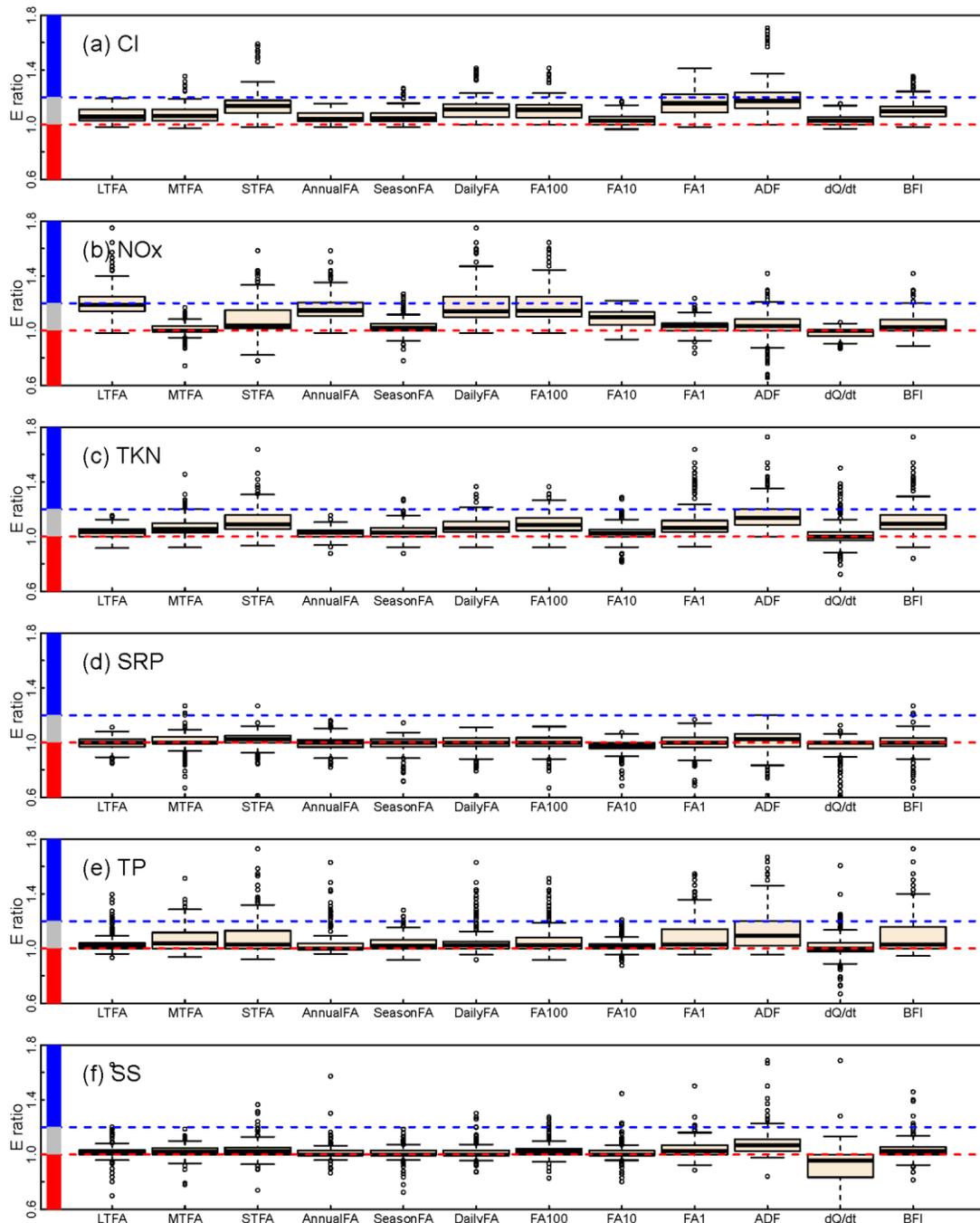


Figure 9.5. Boxplots of ratio between the modified Nash-Sutcliffe efficiency ( $E$ ) of each modified model and  $E$ -value of the original model for the nine sites for (a) Cl, (b)  $\text{NO}_x$ , (c) TKN, (d) SRP, (e) TP, and (f) SS. The ratios are divided into three regions: (1) major improvement (ratio  $> 1.2$ ; blue bars), (2) minor improvement ( $1 < \text{ratio} < 1.2$ ; grey bars), and (3) inferior performance (ratio  $< 1.0$ ; red bars). Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy A*.

In addition, Spearman's  $\gamma$  was calculated to quantify the correlation between original model residuals and each proposed flow variable for each constituent and for each site, which can provide some indication with respect to if a certain flow variable may improve the original model (see discussion in Section 9.4.1). For better comparison of correlation strengths, we use boxplots (Figure 9.6) to summarize the absolute correlations (*i.e.*,  $|\gamma|$ ) for all 270 simulation runs (*i.e.*, 30 runs/sites x 9 sites). In general, a flow variable with a high  $|\gamma|$  may have a high potential for model improvement.

#### **9.4.2.1. Constituents with Major Improvement: Cl and NO<sub>x</sub>**

The modified models show generally major improvements in the estimation of Cl and NO<sub>x</sub>. For Cl (Figure 9.5a), most of the modified models have E-value ratio > 1.0, indicating that the modified models are generally effective in improving Cl estimation. Among all modified models, ADF, STFA, and FA1 show the highest E-value ratios, with about 25%-50% falling into the category of major improvement. For NO<sub>x</sub> (Figure 9.5b), the LTFA model shows the best overall performance, followed by AnnualFA, DailyFA, and FA100. These four models generally have E-value ratio > 1.0. Among these models, LTFA falls into the category of major improvement most frequently. Many other models, however, show some occasions of inferior performance relative to the original model (*i.e.*, E-value ratio < 1.0), including MTFA, STFA, SeasonFA, FA1, ADF, BFI, and particularly  $dQ/dt$ .

These results are consistent with those obtained from correlation analysis. For Cl, highest  $|\gamma|$  values are associated with STFA, FA1, and ADF (Figure 9.6a). In terms of the correlation sign, STFA consistently shows positive  $\gamma$  with the original model residuals (Figure G3a). For NO<sub>x</sub>, LTFA shows the highest  $|\gamma|$  among all variables (Figure 9.6b) and

the  $\gamma$  sign is always negative (Figure G3b). These  $\gamma$  signs are consistent with the two cases discussed in Section 9.4.1.

Comparing LTFA with STFA, we observe a phenomenon previously discussed in the context of our two case study examples – *i.e.*, that LTFA is more effective for NO<sub>x</sub> estimation, whereas STFA is more effective for Cl estimation. The key difference between the two terms lies in the temporal scale for quantification of antecedent conditions. Specifically, LTFA compares average discharges in the preceding 1-year and in the entire record (Equation 9.4), whereas STFA compares discharge on the day of estimation and average discharge in the preceding 1-month (Equation 9.6). Thus, “wet antecedence” is a wet month relative the estimation day for STFA, but a wet year relative to the entire record for LTFA. The relatively longer temporal scale of antecedent relevance for NO<sub>x</sub> than Cl may reflect the relevance of long-term biogeochemical processes that convert ammonia and organic nitrogen to NO<sub>x</sub>, but will likely also reflect that NO<sub>x</sub> is dominantly applied to pervious land through atmospheric deposition and fertilizer application, such that it is more likely to reflect transport through deep and long pathways in the groundwater system (Böhlke and Denver, 1995; Sanford and Pope, 2013). By contrast, Cl is heavily influenced by winter road salt applications in urban and sub-urban areas, which are either rapidly flushed to the river channel or temporally stored in the shallow groundwater system and later slowly released via baseflow (Corsi *et al.*, 2015). To verify these or other possible explanations, further research is required to better understand and compare the source and fate of the two constituents in these watersheds as well as watersheds elsewhere that may have very different sizes and land uses. Additionally, it might be useful to consider alternative definitions of FA terms – *i.e.*,

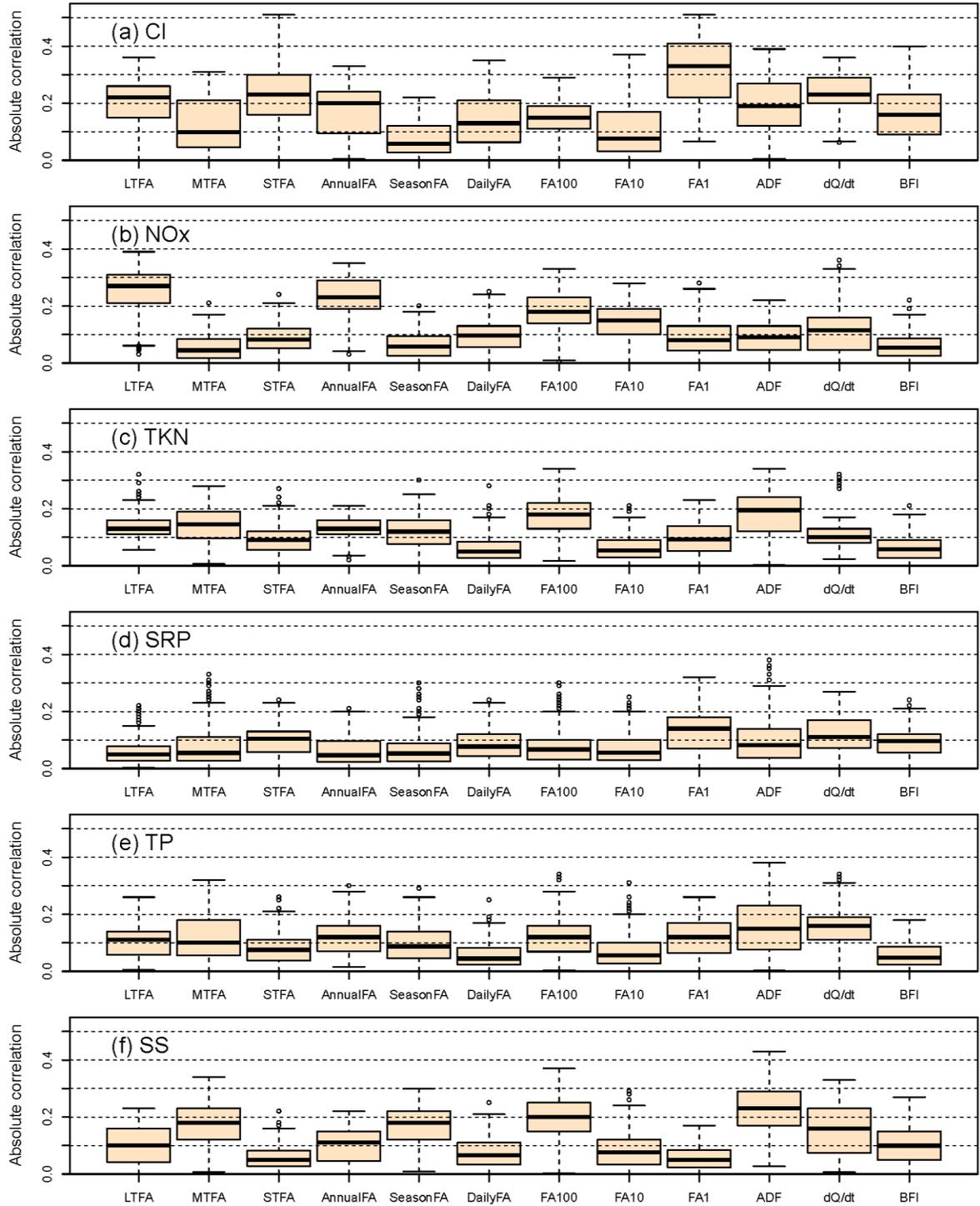


Figure 9.6. Boxplots of absolute values of Spearman's correlation coefficient ( $|\gamma|$ ) between residuals from the original WRTDS model and each of the twelve proposed flow variables. Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy A*.

comparisons among several FAs that consider antecedent conditions of different duration relative to the entire record.

#### **9.4.2.2. Constituents with Moderate Improvement: TKN, TP, and SS**

For TKN (Figure 9.5c), TP (Figure 9.5e), and SS (Figure 9.5f), the ADF model shows the best performance among all models, with E-value ratios (modified model versus original model) that are generally  $> 1.0$ , although a large proportion of the ratios show only moderate improvement. BFI and STFA also show improved estimation of TKN and TP. The other models, however, show some occasions of inferior performance relative to the original model. In particular, the  $dQ/dt$  model is inferior to the original model more often than not, with some E-value ratios far less than 1.0.

Again, these results are consistent with those from correlation analysis. For all three constituents (TKN, TP, and SS), ADF shows the highest  $|\gamma|$  with the original model residuals (Figure 9.6c, 9.6e, 9.6f) and the signs are generally negative (Figure G3c, Figure G3e, Figure G3f), which is similar to the case of LTFA discussed in Section 9.4.1.2 and reflective of the “storage-and-then-release” effect.

#### **9.4.2.3. Constituent without Improvement: SRP**

For soluble reactive phosphorus (SRP), none of the modified models shows consistent improvement (Figure 9.5d). In fact, these models have approximately equal probability of being better or worse than the original model. As elsewhere, the results are consistent with those of the correlation analysis: these proposed variables show generally small  $|\gamma|$  (Figure 9.6d) and the correlations generally have mixed signs (Figure G3d).

### 9.4.3. Performance Comparison for All Simulations: Effects of Sampling Strategies

#### 9.4.3.1. Original Model: Performance under Different Sampling Strategies

Estimation performance of the original model varies with sampling strategies, as shown in Figure 9.7. Based on E-values for concentration, estimation performance generally follows the order of strategy C > strategy B > strategy A, especially for SS and SRP (Figure 9.7a). Based on E-values for flux, estimation performance follows the same order, with SS, TP, and SRP showing the strongest improvement under strategy C (Figure 9.7b). Similarly, based on absolute percent bias ( $|PBIAS|$ ), estimation performance for both concentration and flux also follows the order of C > B > A, especially for SS, TP, and SRP (Figure 9.7c-9.7d). The improvement observed with strategy C is least strong for the two dissolved constituents, *i.e.*, Cl and NO<sub>x</sub>, particularly for concentration estimation.

These results provide several useful insights with respect to the effects of sampling strategy on model performance. First, the overall better performance with strategy B relative to strategy A illustrates the value of collecting additional samples (*i.e.*, 20 vs. 12 samples per year). Second, the overall better performance with strategy C than strategy B indicates the value of routinely collecting high-flow samples, assuming that the number of total samples per year is fixed (*i.e.*, 20 per year). Consequently, strategy C always showed the best estimation performance (*i.e.*, largest E-value and smallest  $|PBIAS|$ ). Clearly, increased sampling frequency and special emphasis on storm sampling should continue to be recommended for river monitoring implementation. An area where more detailed further investigation may be needed, however, is in regard to finding an optimum ratio of storm sampling versus random sampling, and to searching for possible

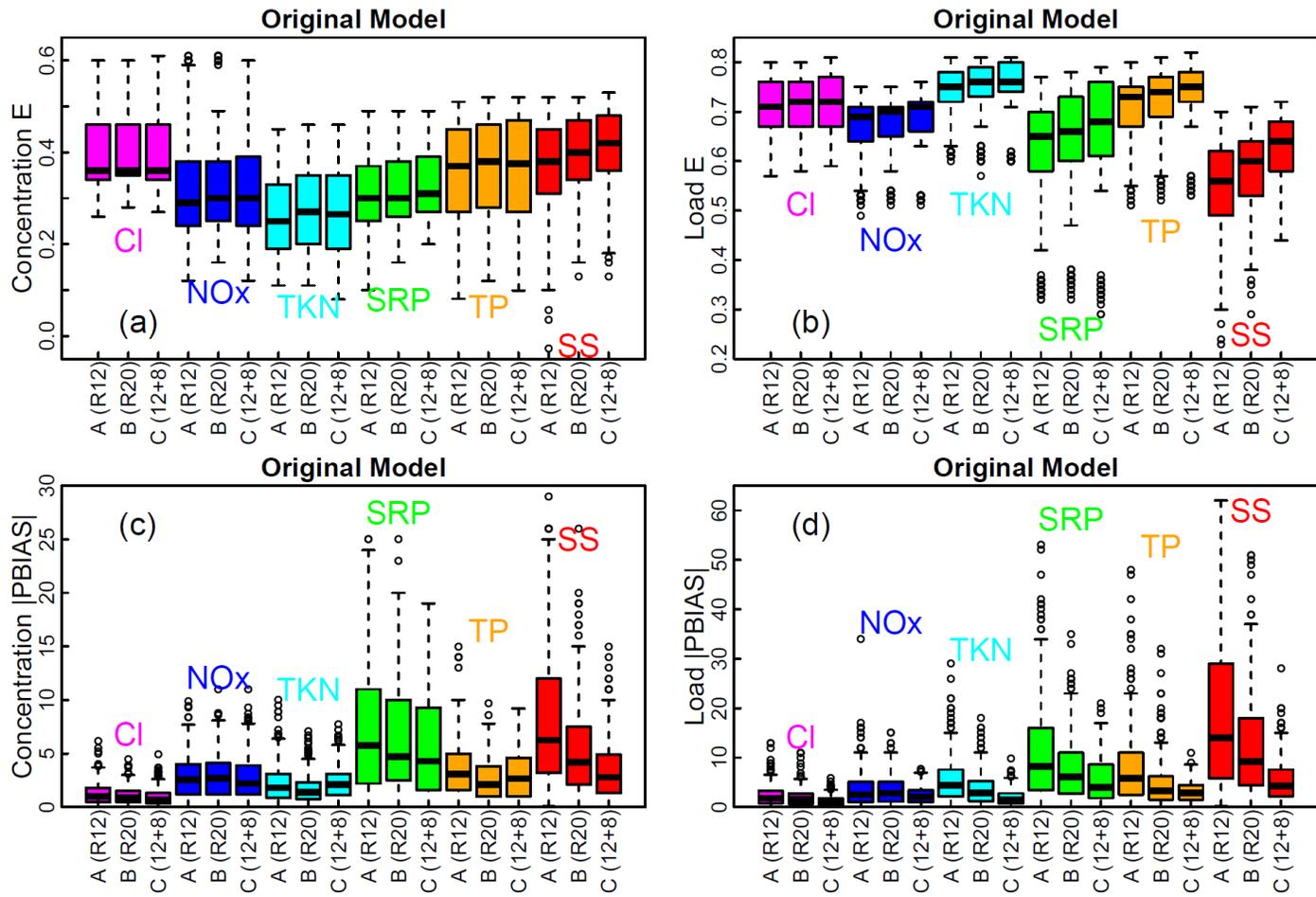


Figure 9.7. Performance of the original WRTDS model under each of the three sampling strategies. The performance is quantified with modified Nash-Sutcliffe efficiency (E) for (a) concentration and (b) flux and with absolute percent bias (|PBIAS|) for (c) concentration and (d) flux. Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy.

dependence of results on parameter of interest and watershed characteristics.

The results of the work presented here demonstrate the effects of sampling strategy on estimation performance vary considerably with constituent of interest. Among all six constituents considered, SS generally showed overall the least satisfactory performance (lowest E-value and highest |PBIAS| for flux) and under all three sampling strategies. Following SS, TP and SRP also show large |PBIAS| values. Interestingly, these three constituents also show strong *differences* in estimation performance among the three sampling strategies. By contrast, Cl, NO<sub>x</sub>, and TKN show the smallest |PBIAS| for both concentration and flux and also show very weak differences among the three sampling strategies in terms of estimation performance. These patterns undoubtedly relate to the different mechanisms of release, transformation, and transport of these constituents. One useful measure to consider in this regard is the coefficient of variation (CV) for observed concentrations. Based on monitoring data at the nine sites, SS has the largest CV among all constituents, followed by TP and SRP (Figure 9.8). From statistical theory, a larger

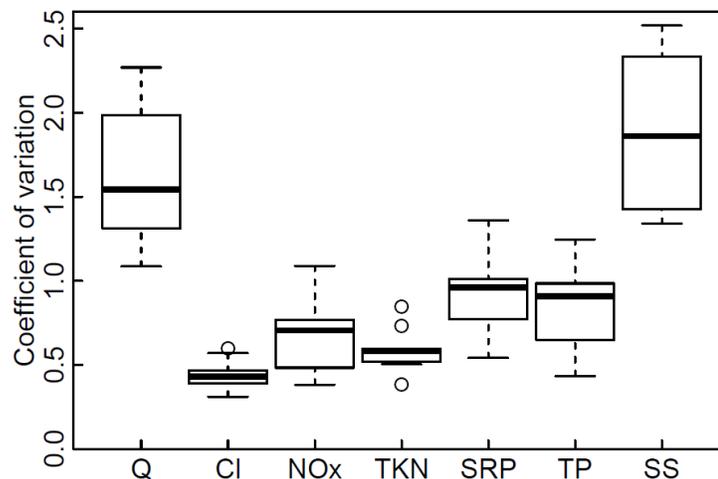


Figure 9.8. Coefficient of variation in daily discharge (Q) data and constituent concentration data. Each boxplot summarizes nine coefficients, each of which was based on data at one individual monitoring site.

CV means that accurate estimation will require a larger number of concentration samples (Haan, 2002). In addition, a larger CV may also require better coverage of the full hydrograph (*i.e.*, both regular flow events and episodic storm-flow events). This interpretation further highlights the benefits of adopting strategy C for river water-quality monitoring, especially for SS, TP, and SRP.

#### ***9.4.3.2. All Models: Performance under Different Sampling Strategies***

Estimation performance of the modified statistical models also varies with sampling strategies, as shown in Figures 9.9-9.10 and G4. For brevity, these figures only compare the median values of each performance measure based on all 270 simulation runs (*i.e.*, 30 runs/sites x 9 sites) for each sampling strategy. The performance measures include E-values for concentration (Figure 9.9), E-values for flux (Figure G4), and |PBIAS| for flux (Figure 9.10).

In terms of E-values for concentration, the above-documented improvements with the modified models are observed under all three sampling strategies (Figure 9.9a-9.9c). Specifically, the ADF, STFA, and FA1 models show major improvement in Cl estimation. The LTFA model shows major improvement in NO<sub>x</sub> estimation. The ADF model shows moderate improvement for SS, TP, and TKN. None of the modified models, however, can effectively improve SRP estimation. Comparing the three sampling strategies directly, median E-values for concentration are generally higher for strategy B than A (Figure 9.9d), generally higher for strategy C than B (Figure 9.9f), and of course, generally higher for strategy C than A (Figure 9.9e), and these are most pronounced for SS and SRP. In terms of E-values for flux, the results are similar, with SS, SRP, and TP showing considerable improvement with strategy C (Figure G4).

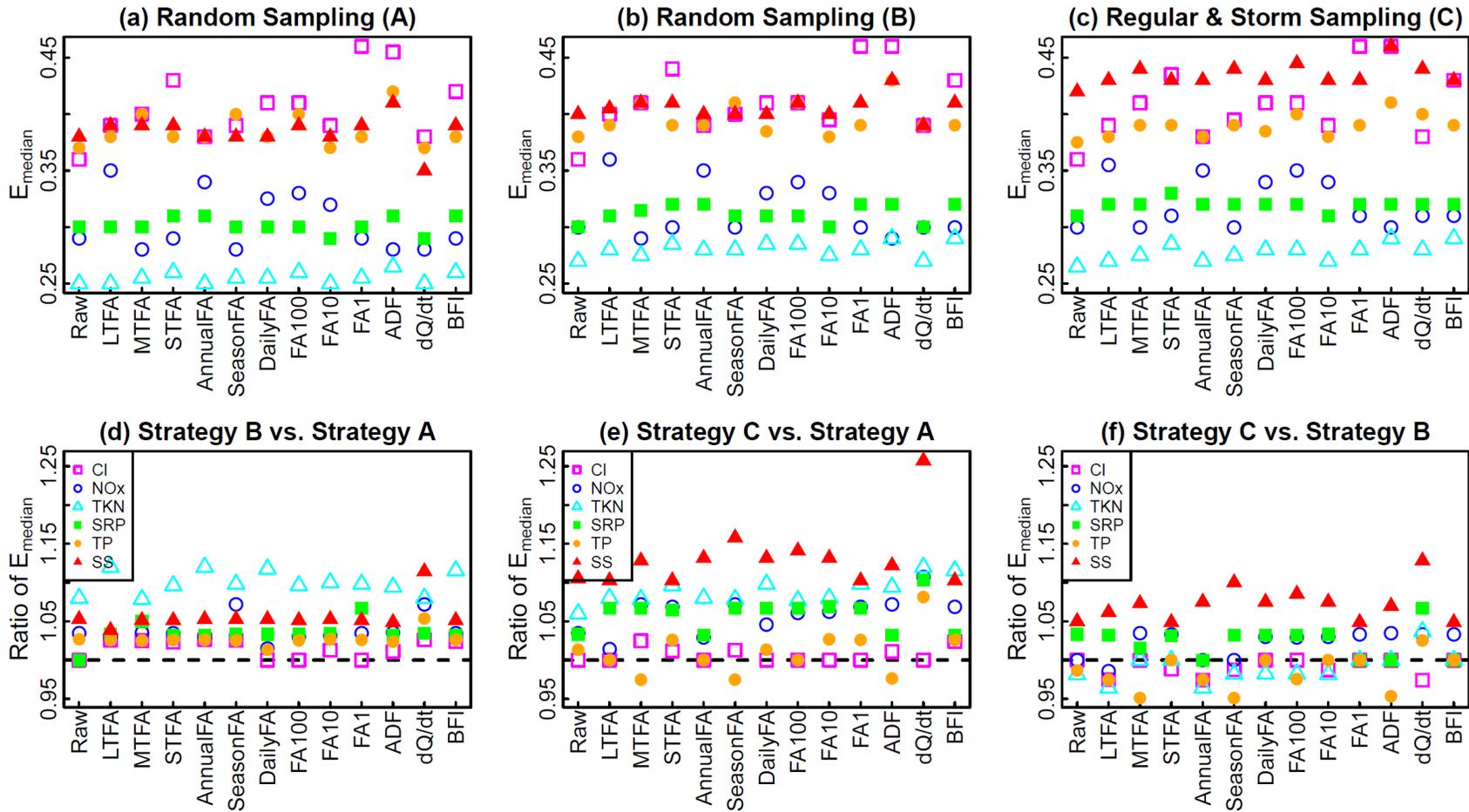


Figure 9.9. Performance of the original and modified models under the three sampling strategies. Plots (a)-(c) summarizes the performance with medians of modified Nash-Sutcliffe efficiency ( $E$ ) for *concentration* based on 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy. Plots (d)-(f) compares the three sampling strategies directly using ratios of  $E$  median.

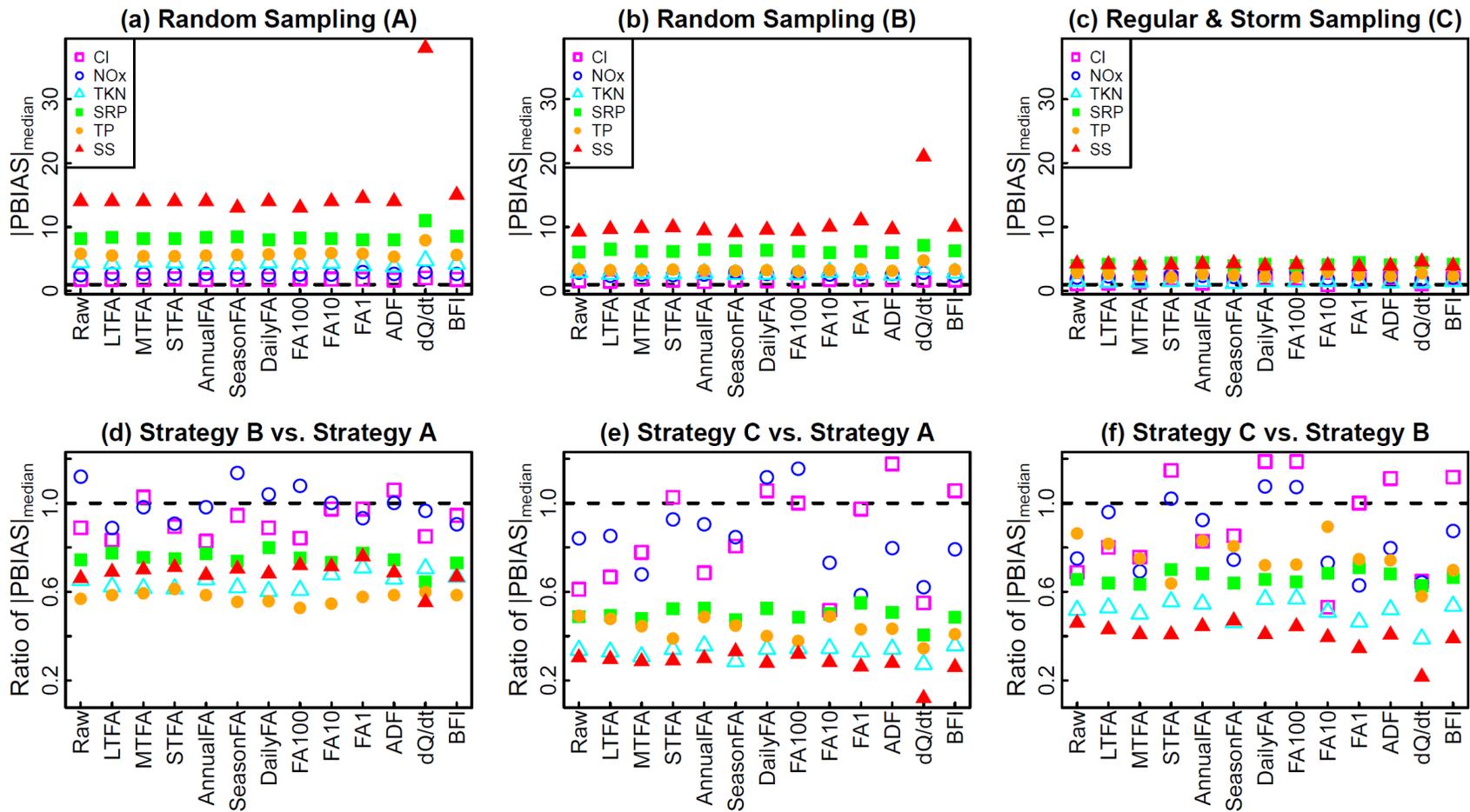


Figure 9.10. Performance of the original and modified models under the three sampling strategies. Plots (a)-(c) summarizes the performance with medians of absolute value of percent bias ( $|PBIAS|$ ) for *flux* based on 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy. Plots (d)-(f) compares the three sampling strategies directly using ratios of  $|PBIAS|$  median.

In terms of |PBIAS| for flux, estimation performance is also generally the best under strategy C, with absolute bias consistently < 5% (Figure 9.10c). Notably, the  $dQ/dt$  model performs much worse than the original model under sampling strategies A and B (Figure 9.10a-9.10b). (This candidate variable seems to be a “confounding” variable in the sense that it compromises the power of the original model’s independent variables (*i.e.*, time, discharge, and season) in capturing the variability of the independent variable (*i.e.*, concentration.) For SS in particular, the absolute biases are about 40% and 20% under A and B, respectively, which are much higher than those of any of the other models. Such inefficacy (high bias) disappears only under strategy C, when storm-flow samples are routinely collected (Figure 9.10c). Comparing the three sampling strategies directly, |PBIAS| medians for flux are generally smaller for strategy B than A (Figure 9.10d), smaller for strategy C than A (Figure 9.10e), and smaller for strategy C than B (Figure 9.10f), and these patterns are most pronounced for SS, SRP, TP, and TKN (Figure G4).

## **9.5. Summary and Final Remarks**

In this work, we developed and evaluated alternative statistical models for estimation of riverine concentration and flux by accounting for antecedent discharge conditions. Simulation results reveal that the modified models show general improvement over the original model under all three sampling strategies. Major improvements were achieved for  $\text{NO}_x$  by the LTFA model and for Cl by the ADF model and the STFA model. In terms of sampling strategy, performance of all models (including the original) was generally best using strategy C and worst using strategy A, and especially so for SS, TP, and SRP, confirming the value of routinely collecting storm-flow samples.

Accurate estimation of riverine constituent concentrations and fluxes has been an active research area for decades and will remain a critical challenge to the hydrological community (Cohn *et al.*, 1989; Cohn *et al.*, 1992; Hirsch *et al.*, 2010; Hirsch, 2014). The major contribution of this work is to provide a comprehensive set of statistical evidence for supporting the incorporation of antecedent discharge into the WRTDS model for improved concentration and flux estimation, thereby combining the advantages of these two recent developments in water quality modeling. In recognition of the need for additional future work, several limitations of this study are acknowledged. First, this study has been limited to only six constituents for nine watersheds within a specific geographical location. Model performance may vary for constituents not covered here. In addition, more studies need to be carried in different types of watershed in other geographical regions to place these results into a more generalizable context to inform further improvement of estimation methods. (In this regard, however, there are unfortunately few if any data of comparable resolution and duration as the Heidelberg set studied here.) Second, all modified models contain only one additional flow variable for the sake of model parsimony. It is possible that a mixture of these terms can better incorporate export dynamics at different temporal scales and thus provide additional improvement. Third, this work examined three sampling strategies that are representative of common monitoring practice, but not strategies with different sampling frequencies (*e.g.*, weekly *vs.* monthly), record lengths (*e.g.*, 10 *vs.* 30 years), or proportions of routine versus storm sampling. Further understanding of these aspects can help inform the allocation of valuable (and often limited) resources toward river quality monitoring. Fourth, in addition to the time, discharge, and season windows in the current WRTDS

model, an additional window may be introduced for the new variable on antecedent discharge, such that higher weights are assigned to concentration samples that have similar antecedent discharge conditions to the estimation day. Fifth, antecedent discharge variables may also be considered toward use with the flow-normalization algorithm in WRTDS for water-quality trend analysis. Last but not least, this work has focused entirely on traditional monitoring data that derive from infrequently measured water-quality samples. With increasing use of new technology that allows very high-frequency or even continuous data collection from in-situ probes, a practical challenge facing the hydrological community is how to combine the use of both traditional low-frequency data and in-situ high-frequency data toward riverine water quality analysis and load estimation. Moving forward, there is need for additional research to advance understandings on these aspects and to incorporate all available understanding into the development and upgrade of software that can be broadly applied.

## **9.6. Supporting Information**

Supporting information to this chapter is provided in Appendix G.

## **9.7. Acknowledgements**

This work was supported by the U.S. Geological Survey (G15AC00067), Maryland Water Resources Research Center (2015MD329B), Maryland Sea Grant (NA10OAR4170072 and NA14OAR1470090), and National Science Foundation (CBET-1360415). We would like to thank Bob Hirsch, Lori Sprague, and Skip Vecchia (USGS) for their advices, reviews, and encouragements. We thank Jeffery Chant (USGS) for sharing the R codes for generating flow-stratified sub-samples. The National Center for

Water Quality Research is acknowledged for collecting and maintaining the high-quality data sets.

## 9.8. Literature Cited

Böhlke, J. K. and J. M. Denver, 1995. Combined Use of Groundwater Dating, Chemical, and Isotopic Analyses to Resolve the History and Fate of Nitrate Contamination in Two Agricultural Watersheds, Atlantic Coastal Plain, Maryland. *Water Resour. Res.* 31:2319-2339, DOI: 10.1029/95WR01584.

Bieroza, M. Z. and A. L. Heathwaite, 2015. Seasonal variation in phosphorus concentration–discharge hysteresis inferred from high-frequency in situ monitoring. *Journal of Hydrology* 524:333-347, DOI: 10.1016/j.jhydrol.2015.02.036.

Biron, P. M., A. G. Roy, F. Courschesne, W. H. Hendershot, B. Côté and J. Fyles, 1999. The effects of antecedent moisture conditions on the relationship of hydrology to hydrochemistry in a small forested watershed. *Hydrol. Process.* 13:1541-1555, DOI: 10.1002/(SICI)1099-1085(19990815)13:11<1541::AID-HYP832>3.0.CO;2-J.

Bowes, M. J., J. T. Smith, H. P. Jarvie and C. Neal, 2008. Modelling of phosphorus inputs to rivers from diffuse and point sources. *Sci. Total Environ.* 395:125-138, DOI: 10.1016/j.scitotenv.2008.01.054.

Chanat, J. G., D. L. Moyer, J. D. Blomquist, K. E. Hyer and M. J. Langland, 2016. Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year

2012. U.S. Geological Survey Scientific Investigations Report 2015-5133, Reston, VA, p. 76. <http://dx.doi.org/10.3133/sir20155133>.
- Cirno, C. P. and J. J. McDonnell, 1997. Linking the hydrologic and biogeochemical controls of nitrogen transport in near-stream zones of temperate-forested catchments: a review. *Journal of Hydrology* 199:88-120, DOI: 10.1016/S0022-1694(96)03286-6.
- Cohn, T. A., D. L. Calder, E. J. Gilroy, L. D. Zynjuk and R. M. Summers, 1992. The validity of a simple statistical model for estimating fluvial constituent loads: An Empirical study involving nutrient loads entering Chesapeake Bay. *Water Resour. Res.* 28:2353-2353, DOI: 10.1029/92wr01008.
- Cohn, T. A., L. L. DeLong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.
- Corsi, S. R., L. A. De Cicco, M. A. Lutz and R. M. Hirsch, 2015. River chloride trends in snow-affected urban watersheds: increasing concentrations outpace urban growth rate and are common among all seasons. *Sci. Total Environ.* 508:488-497, DOI: 10.1016/j.scitotenv.2014.12.012.
- Crowder, D. W., M. Demissie and M. Markus, 2007. The accuracy of sediment loads when log-transformation produces nonlinear sediment load–discharge relationships. *Journal of Hydrology* 336:250-268, DOI: 10.1016/j.jhydrol.2006.12.024.
- Davis, C. A., A. S. Ward, A. J. Burgin, T. D. Loecke, D. A. Riveros-Iregui, D. J. Schnoebelen, C. L. Just, S. A. Thomas, L. J. Weber and M. A. St. Clair, 2014.

- Antecedent Moisture Controls on Stream Nitrate Flux in an Agricultural Watershed. *J. Environ. Qual.* 43:1494-1503, DOI: 10.2134/jeq2013.11.0438.
- Dolan, D. M., A. K. Yui and R. D. Geist, 1981. Evaluation of River Load Estimation Methods for Total Phosphorus. *J. Great Lakes Res.* 7:207-214, DOI: 10.1016/s0380-1330(81)72047-1.
- Dupas, R., C. Gascuel-Oudou, N. Gilliet, C. Grimaldi and G. Gruau, 2015. Distinct export dynamics for dissolved and particulate phosphorus reveal independent transport mechanisms in an arable headwater catchment. *Hydrol. Process.* 29:3162–3178, DOI: 10.1002/hyp.10432.
- Evans, C. and T. D. Davies, 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. *Water Resour. Res.* 34:129-137, DOI: 10.1029/97WR01881.
- Fuka, D., M. Walter, J. Archibald, T. Steenhuis and Z. Easton, 2014. EcoHydRology: A community modeling foundation for Eco-Hydrology. R package version 0.4.12.
- Gray, A. B., G. B. Pasternack, E. B. Watson, J. A. Warrick and M. A. Goñi, 2015. Effects of antecedent hydrologic conditions, time dependence, and climate cycles on the suspended sediment load of the Salinas River, California. *Journal of Hydrology* 525:632-649, DOI: 10.1016/j.jhydrol.2015.04.025.
- Haan, C. T., 2002. *Statistical Methods in Hydrology*. Ames, Iowa, Iowa State University Press,
- Harman, C. J., 2015. Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed. *Water Resour. Res.* 51:1–30, DOI: 10.1002/2014WR015707.

- Helsel, D. R. and R. M. Hirsch, 2002. Statistical Methods in Water Resources. *U.S. Geological Survey Techniques of Water-Resources Investigations Book 4, Chapter A3*. U.S. Geological Survey, Reston, VA, p. 522.  
<http://pubs.usgs.gov/twri/twri4a3/>.
- Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.
- Hirsch, R. M., 2014. Large Biases in Regression-Based Constituent Flux Estimates: Causes and Diagnostic Tools. *J. Am. Water Resour. Assoc.* 50:1401-1424, DOI: 10.1111/jawr.12195.
- Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.
- Hirsch, R. M. and L. De Cicco, 2015. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). U.S. Geological Survey Techniques and Methods Book 4, Chapter A10, Reston, VA, p. 93. <http://dx.doi.org/10.3133/tm4A10>.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.

- Ide, J., H. Haga, M. Chiwa and K. Otsuki, 2008. Effects of antecedent rain history on particulate phosphorus loss from a small forested watershed of Japanese cypress (*Chamaecyparis obtusa*). *Journal of Hydrology* 352:322-335, DOI: 10.1016/j.jhydrol.2008.01.012.
- Inamdar, S. P., S. F. Christopher and M. J. Mitchell, 2004. Export mechanisms for dissolved organic carbon and nitrate during summer storm events in a glaciated forested catchment in New York, USA. *Hydrol. Process.* 18:2651-2661, DOI: 10.1002/hyp.5572.
- Johnes, P. J., 2007. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology* 332:241-258, DOI: 10.1016/j.jhydrol.2006.07.006.
- Kirchner, J. W., X. Feng and C. Neal, 2001. Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations. *Journal of Hydrology* 254:82-101, DOI: 10.1016/S0022-1694(01)00487-5.
- Krause, P., D. P. Boyle and F. Bäse, 2005. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* 5:89-97, DOI: 10.5194/adgeo-5-89-2005.
- Kronvang, B. and A. J. Bruhn, 1996. Choice of sampling strategy and estimation method for calculating nitrogen and phosphorus transport in small lowland streams. *Hydrol. Process.* 10:1483-1501, DOI: 10.1002/(SICI)1099-1085(199611)10:11<1483::AID-HYP386>3.0.CO;2-Y.
- Lyne, V. D. and M. Hollick, 1979. Stochastic Time-Variable Rainfall-Runoff Modeling.

*In: Hydrology and Water Resources Symposium.* Institution of Engineers  
Australia, Perth, Australia, pp. 89-92.

Macrae, M. L., M. C. English, S. L. Schiff and M. Stone, 2010. Influence of antecedent hydrologic conditions on patterns of hydrochemical export from a first-order agricultural watershed in Southern Ontario, Canada. *Journal of Hydrology* 389:101-110, DOI: 10.1016/j.jhydrol.2010.05.034.

McDowell, R. W. and A. N. Sharpley, 2002. The effect of antecedent moisture conditions on sediment and phosphorus loss during overland flow: Mahantango Creek catchment, Pennsylvania, USA. *Hydrol. Process.* 16:3037-3050, DOI: 10.1002/hyp.1087.

Medalie, L., R. M. Hirsch and S. A. Archfield, 2012. Use of flow-normalization to evaluate nutrient concentration and flux changes in Lake Champlain tributaries, 1990–2009. *J. Great Lakes Res.* 38:58-67, DOI: 10.1016/j.jglr.2011.10.002.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel and T. L. Veith, 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASAE* 50:885-900, DOI: 10.13031/2013.23153.

Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.

Murphy, J. C., R. M. Hirsch and L. A. Sprague, 2014. Antecedent flow conditions and nitrate concentrations in the Mississippi River basin. *Hydrol. Earth Syst. Sci.*

18:967-979, DOI: 10.5194/hess-18-967-2014.

Nathan, R. J. and T. A. McMahon, 1990. Evaluation of automated techniques for base flow and recession analyses. *Water Resour. Res.* 26:1465-1473, DOI: 10.1029/WR026i007p01465.

National Center for Water Quality Research, 2015. User's Guide to the River Data Sets. <http://www.heidelberg.edu/academiclife/distinctive/newqr/data>. Accessed July 23, 2015.

Park, Y. S. and B. A. Engel, 2015. Analysis for Regression Model Behavior by Sampling Strategy for Annual Pollutant Load Estimation. *J. Environ. Qual.* 44:1843-1851, DOI: 10.2134/jeq2015.03.0137.

Pellerin, B. A., B. A. Bergamaschi, R. J. Gilliom, C. G. Crawford, J. Saraceno, C. P. Frederick, B. D. Downing and J. C. Murphy, 2014. High frequency measurement of nitrate concentration in the Lower Mississippi River, USA. *Environ. Sci. Technol.* 48:12612-12619, DOI: 10.1021/es504029c.

Ryberg, K. R. and A. V. Vecchia, 2014. waterData: An R Package for Retrieval, Analysis, and Anomaly Calculation of Daily Hydrologic Time Series Data. R package version 1.0.4.

Sanford, W. E. and J. P. Pope, 2013. Quantifying Groundwater's Role in Delaying Improvements to Chesapeake Bay Water Quality. *Environ. Sci. Technol.* 47:13330-13338, DOI: 10.1021/es401334k.

Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.

- Sprague, L. A., R. M. Hirsch and B. T. Aulenbach, 2011. Nitrate in the Mississippi River and its tributaries, 1980 to 2008: are we making progress? *Environ. Sci. Technol.* 45:7209-7216, DOI: 10.1021/es201221s.
- Stenback, G. A., W. G. Crumpton, K. E. Schilling and M. J. Helmers, 2011. Rating curve estimation of nutrient loads in Iowa rivers. *Journal of Hydrology* 396:158-169, DOI: 10.1016/j.jhydrol.2010.11.006.
- Stow, C. A. and M. E. Borsuk, 2003. Assessing TMDL effectiveness using flow-adjusted concentrations: a case study of the Neuse River, North Carolina. *Environ. Sci. Technol.* 37:2043-2050, DOI: 10.1021/es020802p.
- Turgeon, J. M. L. and F. Courchesne, 2008. Hydrochemical behaviour of dissolved nitrogen and carbon in a headwater stream of the Canadian Shield: relevance of antecedent soil moisture conditions. *Hydrol. Process.* 22:327-339, DOI: 10.1002/hyp.6613.
- U.S. Geological Survey, 2015. Surface-water data for the nation. <http://dx.doi.org/10.5066/F7P55KJN>. Accessed July 23, 2015.
- Vecchia, A. V., 2003. Relation between climate variability and stream water quality in the continental United States. *Hydrological Science and Technology* 19:77-98.
- Vecchia, A. V., 2005. Water-quality trend analysis and sampling design for streams in the Red River of the North Basin, Minnesota, North Dakota, and South Dakota, 1970-2001. U.S. Geological Survey Scientific Investigations Report 2005-5224, Reston, VA, p. 54. <http://pubs.usgs.gov/sir/2005/5224/>.
- Vecchia, A. V., R. J. Gilliom, D. J. Sullivan, D. L. Lorenz and J. D. Martin, 2009. Trends in concentrations and use of agricultural herbicides for Corn Belt rivers, 1996-

2006. *Environ. Sci. Technol.* 43:9096-9102, DOI: 10.1021/es902122j.
- Vecchia, A. V., J. D. Martin and R. J. Gilliom, 2008. Modeling Variability and Trends in Pesticide Concentrations in Streams. *J. Am. Water Resour. Assoc.* 44:1308–1324, DOI: 10.1111/j.1752-1688.2008.00225.x.
- Wang, Y.-G., P. Kuhnert and B. Henderson, 2011. Load estimation with uncertainties from opportunistic sampling data – A semiparametric approach. *Journal of Hydrology* 396:148-157, DOI: 10.1016/j.jhydrol.2010.11.003.
- Wang, Y.-G. and T. Tian, 2013. Sediment concentration prediction and statistical evaluation for annual load estimation. *Journal of Hydrology* 482:69-78, DOI: 10.1016/j.jhydrol.2012.12.043.
- Wang, Y.-G., S. S. J. Wang and J. Dunlop, 2015. Statistical modelling and power analysis for detecting trends in total suspended sediment loads. *Journal of Hydrology* 520:439-447, DOI: 10.1016/j.jhydrol.2014.10.062.
- Warner, S., G. Kiely, G. Morgan and J. O'Halloran, 2009. Does quantifying antecedent flow conditions improve stream phosphorus export estimation? *Journal of Hydrology* 378:97-104, DOI: 10.1016/j.jhydrol.2009.09.009.
- Zhang, Q., W. P. Ball and D. L. Moyer, 2016a. Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns. *Sci. Total Environ.* 563-564: 1016-1029, DOI: 10.1016/j.scitotenv.2016.03.104.
- Zhang, Q., D. C. Brady and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.* 452-453:208-221, DOI:

10.1016/j.scitotenv.2013.02.012.

Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.

Zhang, Q., R. M. Hirsch and W. P. Ball, 2016b. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* 50:1877-1886, DOI: 10.1021/acs.est.5b04073.

*Page intentionally left blank*

## Chapter 10. Evaluation of Methods for Estimating Long-Range Dependence in Water Quality Time Series with Irregular Sampling<sup>22</sup>

### Abstract

River water-quality time series often exhibit long-range dependence (LRD), which is to say the autocorrelation between observations decays more slowly than exponential. LRD presents challenges to the identification of deterministic trends, but traditional methods for estimating LRD are generally inapplicable to irregularly sampled data. Here we consider two types of estimation approaches for irregularly sampled data and evaluate their performance using synthetic time series. These time series were generated such that (1) their sampling gap intervals mimic the sampling irregularity in real water-quality data, and (2) their data values exhibit a wide range of prescribed LRD behaviors, ranging from white noise (spectral slope = 0) to Brownian noise (spectral slope = 2). The results suggest that none of the existing methods fully account for the effects of sampling irregularity on LRD estimation. First, the results illustrate the danger of using interpolation for gap-filling when examining auto-correlation, as the interpolation methods consistently under-estimate or over-estimate LRDs under a wide range of LRDs and gap distributions. Second, the long-established Lomb-Scargle spectral method also consistently under-estimates LRD. A modified form, using only the lowest 5% of the frequencies for spectral slope estimation, has very poor precision, although the overall

---

<sup>22</sup> This chapter will be submitted for publication in a peer-reviewed journal. Ciaran Harman and James Kirchner were involved in hypothesis development, study design, results interpretation, and editing. All figures, tables, and data were created by Qian Zhang.

bias is small. Third, a wavelet-based method, coupled with an aliasing filter, generally has the smallest bias and root-mean-squared error among all methods for a wide range of LRDs and gap distributions. The aliasing method, however, does not itself account for sampling irregularity, and this introduces some bias in the result. Nonetheless, the wavelet method is recommended for estimating LRD in irregular time series until improved methods are developed. Finally, all methods' performances depend strongly on the sampling irregularity (as quantified by both the skewness and mean of gap-interval lengths), highlighting that the accuracy and precision of each method are data-specific. Accurately quantifying LRD in irregular water-quality time series remains an unresolved challenge for the hydrologic community and for other disciplines that must grapple with irregular sampling.

## **10.1. Introduction**

### **10.1.1. Autocorrelations in Time Series**

It is well known that time series from natural systems often exhibit auto-correlation, that is, observations at each time step are correlated with observations one or more time steps in the past. This property is usually characterized by the autocorrelation function (ACF), which is defined as follows for a process  $X_t$  at lag  $k$ :

$$\gamma(k) = cov(X_t, X_{t+k}) \quad (10.1)$$

In practice, auto-correlation has been frequently modeled with classical techniques such as auto-regressive (AR) or auto-regressive moving-average (ARMA) models (Darken *et al.*, 2002; Yue *et al.*, 2002; Box *et al.*, 2008). These models assume that the underlying

process has short-term memory or short-range dependence (SRD), *i.e.*, the ACF decays exponentially with lag  $k$ , which implies that the ACF is summable (Box *et al.*, 2008).

Although the SRD assumption holds sometimes, it cannot adequately describe many time series whose ACFs decay hyperbolically (thus much slower than exponentially) toward zero and may not reach zero even for large lags, which implies that the ACF is non-summable. This property is commonly referred to as “long-term persistence” or “long-range dependence (LRD)”, as opposed to SRD (Beran, 2010).

### 10.1.2. Overview of Approaches for LRD Quantification

Several equivalent metrics can be used to quantify LRD. Here I provide a review of the definitions of LRD and the typical approaches for LRD modeling, including both time-domain and frequency-domain techniques, with special attention to their reconciliation. For a more comprehensive review, readers are referred to Beran *et al.* (2013), Boutahar *et al.* (2007), and Witt and Malamud (2013).

Strictly speaking,  $X_t$  is called a stationary LRD process if the condition

$$\lim_{k \rightarrow \infty} k^\alpha \gamma(k) = C_1 > 0 \quad (10.2)$$

where  $C_1$  is a constant, is satisfied by some  $\alpha \in (0,1)$  (Boutahar *et al.*, 2007; Beran *et al.*, 2013). Equivalently,  $X_t$  is a LRD process if, in the spectral domain, the condition

$$\lim_{\omega \rightarrow 0} |\omega|^\beta f(\omega) = C_2 > 0 \quad (10.3)$$

is satisfied by some  $\beta \in (0,1)$ , where  $C_2$  is a constant and  $f(\omega)$  is the spectral density function (SDF) of  $X_t$ , which is related to the ACF as follows (which is also known as the Wiener-Khinchin theorem):

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-ik\omega} \quad (10.4)$$

where  $\omega$  is angular frequency (Boutahar *et al.*, 2007).

One popular model for describing LRD processes is the so-called fractional autoregressive integrated moving-average model, or ARFIMA  $(p, q, d)$ , which is an extension of ARMA models and is defined as follows:

$$(1 - B)^d \varphi(B) X_t = \psi(B) \varepsilon_t \quad (10.5)$$

where  $\varepsilon_t$  is a series of i.i.d. Gaussian random numbers  $\sim (0, \sigma_\varepsilon^2)$ ,  $B$  is the backshift operator (*i.e.*,  $BX_t = X_{t-1}$ ), and functions  $\varphi(\cdot)$  and  $\psi(\cdot)$  are polynomials of order  $p$  and  $q$ , respectively. The fractional differencing parameter  $d$  is related to the parameter  $\alpha$  in Equation (10.2) as follows:

$$d = \frac{1 - \alpha}{2} \in (-0.5, 0.5) \quad (10.6)$$

(Beran *et al.*, 2013; Witt and Malamud, 2013).

In addition to a slowly decaying ACF, LRD manifests itself in two other equivalent fashions. One is the so-called ‘‘Hurst effect’’, which states that, on a log-log scale, the range of variability of a process changes linearly with the length of time period under consideration. This power-law slope is often referred to as the ‘‘Hurst exponent’’ or ‘‘Hurst coefficient’’  $H$  (Hurst, 1951), which is related to  $d$  as follows:

$$H = d + 0.5 \quad (10.7)$$

(Beran *et al.*, 2013; Witt and Malamud, 2013). The second equivalent description of LRD, this time from a frequency-domain perspective, is ‘‘fractal scaling’’, which describes a power-law decrease in spectral power with increasing frequency, yielding

power spectra that are linear on log-log axes (Lomb, 1976; Scargle, 1982; Kirchner, 2005). Mathematically, this inverse proportionality can be expressed as:

$$f(\omega) = C_3 |\omega|^{-\beta} \quad (10.8)$$

where  $C_3$  is a constant and the scaling exponent  $\beta$  is termed the “spectral slope”. For spectral slopes of zero, one, and two, the underlying processes are termed as “white”, “pink” (or “flicker”), and “brown” (or “red”) noises, respectively (Witt and Malamud, 2013). Illustrative examples of these three noises are shown in Figure 10.1a-10.1c.

In addition, it can be shown that that the spectral density function for ARFIMA  $(p,d,q)$  is

$$f(\omega) = \frac{\sigma_\varepsilon^2 |\psi(e^{-i\omega})|^2}{2\pi |\varphi(e^{-i\omega})|^2} |1 - e^{-i\omega}|^{-2d} \quad (10.9)$$

for  $-\pi < \omega < \pi$  (Boutahar *et al.*, 2007; Beran *et al.*, 2013). For  $|\omega| \ll 1$ , Equation (10.9) can be approximated by:

$$f(\omega) = C_4 |\omega|^{-2d} \quad (10.10)$$

with

$$C_4 = \frac{\sigma_\varepsilon^2 |\psi(1)|^2}{2\pi |\varphi(1)|^2} \quad (10.11)$$

Equation (10) thus exhibits the asymptotic behavior required for a LRD process given by Equation (3). In addition, a comparison of Equations (10.10) and (10.8) reveals that,

$$\beta = 2d \quad (10.12)$$

Overall, these derivations indicate that these different types of scaling parameters (*i.e.*,  $\alpha$ ,  $d$ , and  $H$  and  $\beta$ ) can be used equivalently to describe the LRD strength of a process.

Specifically, their equivalency can be summarized as follows:

$$\beta = 2d = 1 - \alpha = 2H - 1 \quad (10.13)$$

It should be noted that  $d$ ,  $\alpha$ , and  $H$  are only applicable over a fixed range of LRD strengths, which is equivalent to  $(-1, 1)$  in terms of  $\beta$ .

LRD has been increasingly recognized in studies of hydrological time series, particularly for the common task of trend identification. Such hydrological series include

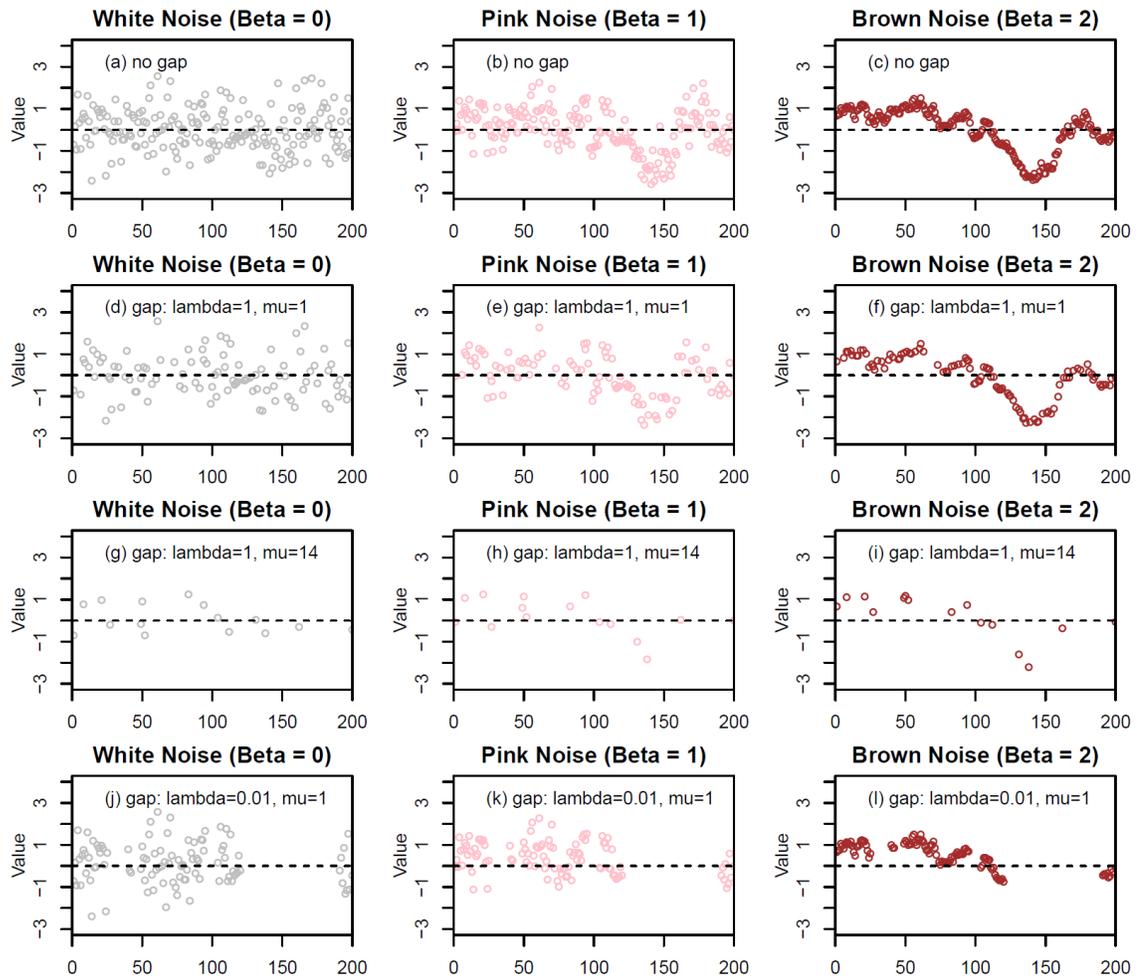


Figure 10.1. Synthetic time series with 200 time steps for three representative LRD strengths that correspond to white noise ( $\beta = 0$ ), pink noise ( $\beta = 1$ ), and brown noise ( $\beta = 2$ ). The 1<sup>st</sup> row shows the simulated time series without any gap. The three rows below show the same time series as in the 1<sup>st</sup> row but with data gaps that were simulated using three different negative binomial (NB) distributions -- 2<sup>nd</sup> row:  $NB(\lambda = 1, \mu = 1)$ ; 3<sup>rd</sup> row:  $NB(\lambda = 1, \mu = 14)$ ; 4<sup>th</sup> row:  $NB(\lambda = 0.01, \mu = 1)$ .

riverflow (Montanari *et al.*, 2000; Khaliq *et al.*, 2008; Khaliq *et al.*, 2009; Ehsanzadeh and Adamowski, 2010), air and sea temperature (Fatichi *et al.*, 2009; Lennartz and Bunde, 2009; Franzke, 2012b; Franzke, 2012a), conservative tracers (Kirchner *et al.*, 2000; Kirchner *et al.*, 2001; Godsey *et al.*, 2010), and non-conservative chemical species (Kirchner and Neal, 2013). Because for LRD processes the variance of the sample mean converges to zero much slower than the rate of  $n^{-1}$  ( $n$ : sample size), the LRD property of a time series must be taken into account to avoid "false positives" (Type I errors) when inferring the statistical significance of trends (Cohn and Lins, 2005; Fatichi *et al.*, 2009; Ehsanzadeh and Adamowski, 2010; Franzke, 2012a). Unfortunately, as stressed by Cohn and Lins (2005), it is “surprising that nearly every assessment of trend significance in geophysical variables published during the past few decades has failed [to do so]”, and a similar tendency is evident in the decade following that statement as well.

### **10.1.3. Motivation and Objective of this Work**

To account for LRD in trend analysis, one must be able to first quantify the strength of LRD for a given time series. Numerous estimation methods have been developed for this purpose, including Hurst rescaled range analysis, Higuchi’s method, Geweke and Porter-Hudak’s method, Whittle’s maximum likelihood estimator, detrended fluctuation analysis, and others (Taqqu *et al.*, 1995; Montanari *et al.*, 1997; Montanari *et al.*, 1999; Rea *et al.*, 2009; Stroe-Kunold *et al.*, 2009). For brevity, these methods are not elaborated here; readers are referred to Beran (2010) and Witt and Malamud (2013) for details. While these estimation methods have been extensively evaluated, they are unfortunately only applicable to regular (*i.e.*, evenly-spaced) data, *e.g.*, daily streamflow discharge, monthly temperature. In practice, many types of hydrological data, including

river water-quality data, are often sampled irregularly or have missing values, and hence their LRD strengths cannot be readily estimated with the above estimation methods.

Thus LRD estimation in irregularly sampled data is an important challenge for hydrologists and practitioners. Many data analysts may be tempted to interpolate the time series to make it regular and hence analyzable (Graham, 2009). Although technically convenient, interpolation can be problematic if it distorts the series' autocorrelation structure (Kirchner and Weil, 1998). In this regard, it is critically important to evaluate various types of interpolation methods using carefully designed benchmark tests, to identify the scenarios under which the interpolated data can yield reliable (or, alternatively, biased) estimates of LRD.

Moreover, quantification of LRD in real-world water-quality data is subject several common complexities. First, many water-quality data are rarely normally distributed; instead, they are typically characterized by log-normal or other skewed distributions (Hirsch *et al.*, 1991; Helsel and Hirsch, 2002), with potential consequences for estimates of LRD. Moreover, water-quality data also tend to exhibit long-term trend, seasonality, and flow-dependence (Hirsch *et al.*, 1991; Helsel and Hirsch, 2002), which can also affect the accuracy of LRD estimates. Thus, it may be more plausible to quantify LRD in transformed time series after accounting for the seasonal patterns and discharge-driven variations in the original time series. For the trend aspect, however, it remains a puzzle whether the data set should be de-trended before conducting LRD estimation. Such de-trending treatment can certainly affect the estimated value of LRD and hence the validity of (or confidence in) any inference made regarding the statistical significance of temporal trends in the time series. This somewhat circular issue is beyond the scope of our current

work -- it has been previously discussed in the context of short-range dependence (Zetterqvist, 1991; Darken *et al.*, 2002; Yue *et al.*, 2002; Noguchi *et al.*, 2011; Clarke, 2013; Sang *et al.*, 2014), but it is not well understood in the context of LRD and hence presents an important area for future research.

In the above context, the main objective of this work was to use Monte Carlo simulation to systematically evaluate and compare two broad types of approaches for estimating LRD in irregularly sampled river water-quality time series. Specific aims of this work include the following:

- (1) To examine the sampling irregularity of typical river water-quality monitoring data and to simulate time series that contain such irregularity; and
- (2) To evaluate two broad types of approaches for estimating LRD in simulated irregularly sampled time series.

The first type of approach adopts several forms of interpolation techniques for gap-filling, thus making the data regular and analyzable by traditional LRD estimation methods. The second type of approach includes the well-known Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982) and a wavelet method combined with a spectral aliasing filter (Kirchner and Neal, 2013). The latter two methods can be directly applied to irregularly spaced data. While this study was intended to help hydrologists and practitioners choose appropriate methods to estimate LRD in water-quality time series, the findings and approaches may be broadly applicable to irregularly sampled data in other scientific fields.

The rest of the paper is organized as follows. I propose an approach to model sampling irregularity in typical river water-quality data and discuss our approach on

simulation of irregularly-sampled data (Section 10.2). I then introduce the various types of method for estimating LRD in irregular time series and compare their estimation performance (Section 10.3). I close with a discussion of the results (Section 10.4).

## **10.2. Quantification of Sampling Irregularity in Typical River Water-Quality Data**

### **10.2.1. Modeling of Sampling Irregularity**

In practice, river water-quality data are usually sampled irregularly. In some cases, samples are taken more frequently during particular periods of interest, such as high flows or drought periods; here we will address the implications of the irregularity, but not the (intentional) bias, inherent in such a sampling strategy. In other cases, the sampling is planned with a fixed sampling interval (*e.g.*, 1 day) but samples are missed (or lost, or fail quality-control checks) at some time steps during implementation. In still other cases, the sampling is intrinsically irregular because, for example, one cannot measure the chemistry of rainfall on rainless days or the chemistry of a stream that has dried up. Theoretically, any deviation from fixed-interval sampling can affect the subsequent analysis of the time series.

To quantify the sampling irregularity, we propose a simple and general approach that can be applied to any time series of monitoring data. Specifically, for a given time series with  $N$  points, the time intervals between adjacent samples are calculated; these intervals themselves make up a time series of  $N-1$  points that we call  $\Delta t$ . In addition, the following parameters are calculated to quantify its sampling irregularity:

- $L$  = the length of the period of record,
- $N$  = the number of samples in the record,

- $\Delta t_{nominal}$  = the nominal sampling interval under regular sampling (e.g.,  $\Delta t_{nominal}$  = 1 day for daily samples),
- $\Delta t^* = \Delta t / \Delta t_{nominal}$ , the sample intervals non-dimensionalized by the nominal sampling interval,
- $\Delta t_{average} = L / (N-1)$  the average of all the entries in  $\Delta t$ .

The quantification is illustrated with two simple examples. The first example contains data sampled every hour from 1:00 am to 11:00 am on one day. In this case,  $L = 10$  hours,  $N = 11$  samples,  $\Delta t = \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$  hours, and  $\Delta t_{nominal} = \Delta t_{average} = 1$  hour. The second example contains data sampled at 1:00 am, 3:00 am, 4:00 am, 8:00 am, and 11:00 am. In this case,  $L = 10$  hours,  $N = 5$  samples,  $\Delta t = \{2, 1, 4, 3\}$  hours,  $\Delta t_{nominal} = 1$  hour, and  $\Delta t_{average} = 2.5$  hours. It is readily evident that the first case corresponds to fixed-interval (regular) sampling that has the property of  $\Delta t_{average} / \Delta t_{nominal} = 1$  (dimensionless), whereas the second case corresponds to irregular sampling for which  $\Delta t_{average} / \Delta t_{nominal} > 1$ .

The dimensionless set  $\Delta t^*$  contains essential information for determining sampling irregularity. This set is modeled as independent, identically distributed values drawn from a negative binomial (NB) distribution. This distribution has two dimensionless parameters, the shape parameter ( $\lambda$ ) and the mean parameter ( $\mu$ ), which collectively represent the irregularity of the samples. The NB distribution is a flexible distribution that provides a discrete analogue of a gamma distribution. The geometric distribution, itself the discrete analogue of the exponential distribution, is a special case of the NB when  $\lambda = 1$ .

The parameters  $\mu$  and  $\lambda$  represent different aspects of sampling irregularity, as illustrated by the examples shown in Figure 10.2. The mean parameter  $\mu$  represents the fractional increase in the average interval between samples due to gaps:  $\mu = \text{mean}(\Delta t^*) - 1 = (\Delta t_{\text{average}} - \Delta t_{\text{nominal}})/\Delta t_{\text{nominal}}$ . Thus the special case of  $\mu = 0$  corresponds to regular sampling (*i.e.*,  $\Delta t_{\text{average}} = \Delta t_{\text{nominal}}$ ), whereas any larger value of  $\mu$  corresponds to irregular sampling (*i.e.*,  $\Delta t_{\text{average}} > \Delta t_{\text{nominal}}$ ) (Figure 10.2c). The shape parameter  $\lambda$  characterizes the

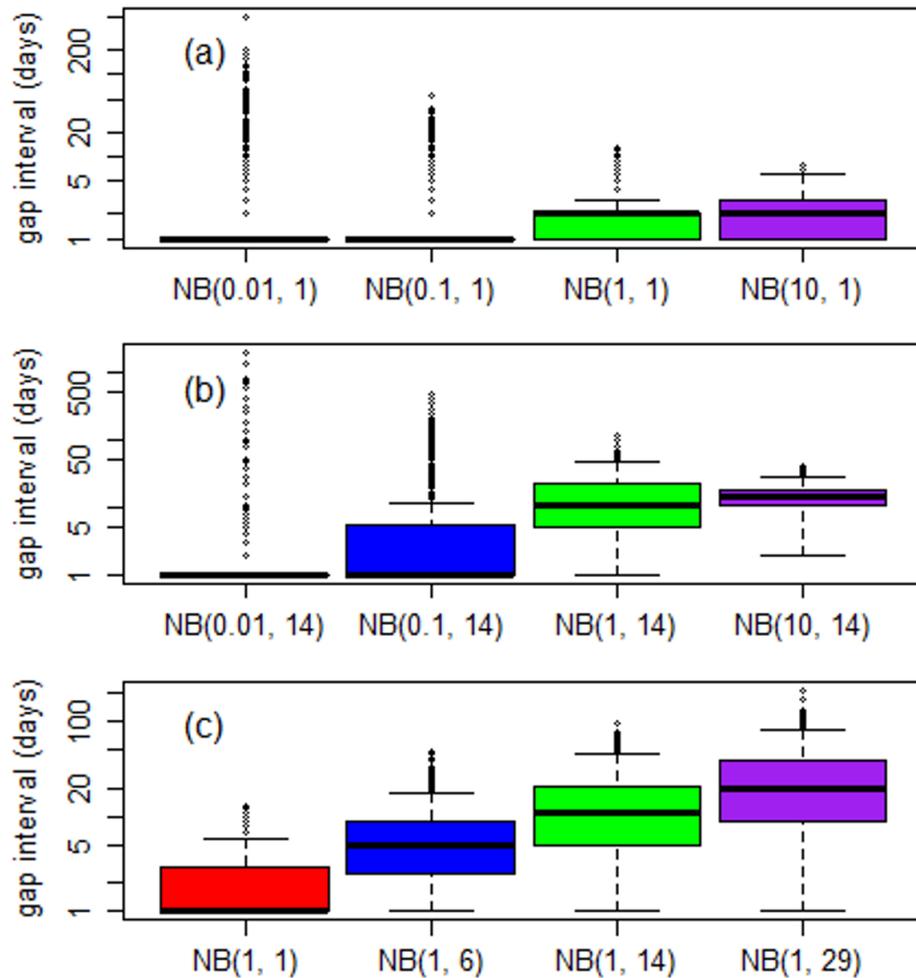


Figure 10.2. Examples of gap interval simulation using binomial distributions, NB (shape  $\lambda$ , mean  $\mu$ ). Simulation parameters:  $L = 9125$  days,  $\Delta t_{\text{nominal}} = 1$  day. The three panels show simulation with fixed (a)  $\mu = 1$ , (b)  $\mu = 14$ , and (c)  $\lambda = 1$ . Note that  $\Delta t_{\text{average}}/\Delta t_{\text{nominal}} = \mu + 1$ .

similarity of gaps to each other; that is, a small  $\lambda$  indicates that the samples contain only a few gaps of varying lengths, whereas a large  $\lambda$  indicates that the samples contain many gaps of similar lengths (Figure 10.2a-10.2b).

To visually illustrate these gap distributions, representative samples of gappy time series are presented in Figure 10.1 for the three special processes described above (Section 10.1.2), *i.e.*, white noise, pink noise, and brown noise. Specifically, three different gap distributions, namely,  $\text{NB}(\lambda = 1, \mu = 1)$ ,  $\text{NB}(\lambda = 1, \mu = 14)$ , and  $\text{NB}(\lambda = 0.01, \mu = 1)$ , were simulated and each was applied to convert the original (complete) time series (Figure 10.1a-10.1c) to gappy series (Figure 10.1d-10.1f). These simulations clearly illustrate the effects of the two parameters  $\lambda$  and  $\mu$ . In particular, compared with  $\text{NB}(\lambda = 1, \mu = 1)$ ,  $\text{NB}(\lambda = 1, \mu = 14)$  shows a similar level of sampling irregularity (same  $\lambda$ ) but a much longer average gap interval (larger  $\mu$ ). Again compared with  $\text{NB}(\lambda = 1, \mu = 1)$ ,  $\text{NB}(\lambda = 0.01, \mu = 1)$  shows the same average interval (same  $\mu$ ) but a much more irregular (skewed) gap distribution that contains a few very long gaps (smaller  $\lambda$ ).

### **10.2.2. Examination of Sampling Irregularity in Real River Water-Quality Data**

The above approach was applied to real water-quality data from two large river monitoring networks in the United States. One network is the Chesapeake Bay River Input Monitoring program, which typically samples streams bi-monthly to monthly, accompanied with additional sampling during stormflows (Langland *et al.*, 2012; Zhang *et al.*, 2015). The other network is the Lake Erie and Ohio tributary monitoring program, which typically samples streams at a daily resolution (National Center for Water Quality Research, 2015). For each site, we have examined concentration data of nitrogen and phosphorus and determined the above-mentioned parameters to quantify sampling

irregularity. The shape parameter  $\lambda$  of a NB distribution can be calculated directly from the mean and variance of  $\Delta t^*$  as follows:  $\lambda = \mu^2 / [\text{var}(\Delta t^*) - \mu] = (\text{mean}(\Delta t^*) - 1)^2 / [\text{var}(\Delta t^*) - \text{mean}(\Delta t^*) + 1]$ . Alternatively, a maximum likelihood approach can be used, which employs the “*fitdist*” function in the “*fitdistrplus*” R package (Delignette-Muller and Dutang, 2015). In general, the two approaches have yielded in similar results, which are summarized in Table 10.1, with two examples of NB distributions shown in Figure 10.3.

For the Chesapeake Bay River Input Monitoring program (9 sites), total nitrogen (TN) and total phosphorus (TP) are taken as representatives of water-quality constituents. According to the maximum likelihood approach, the shape parameter  $\lambda$  varies between

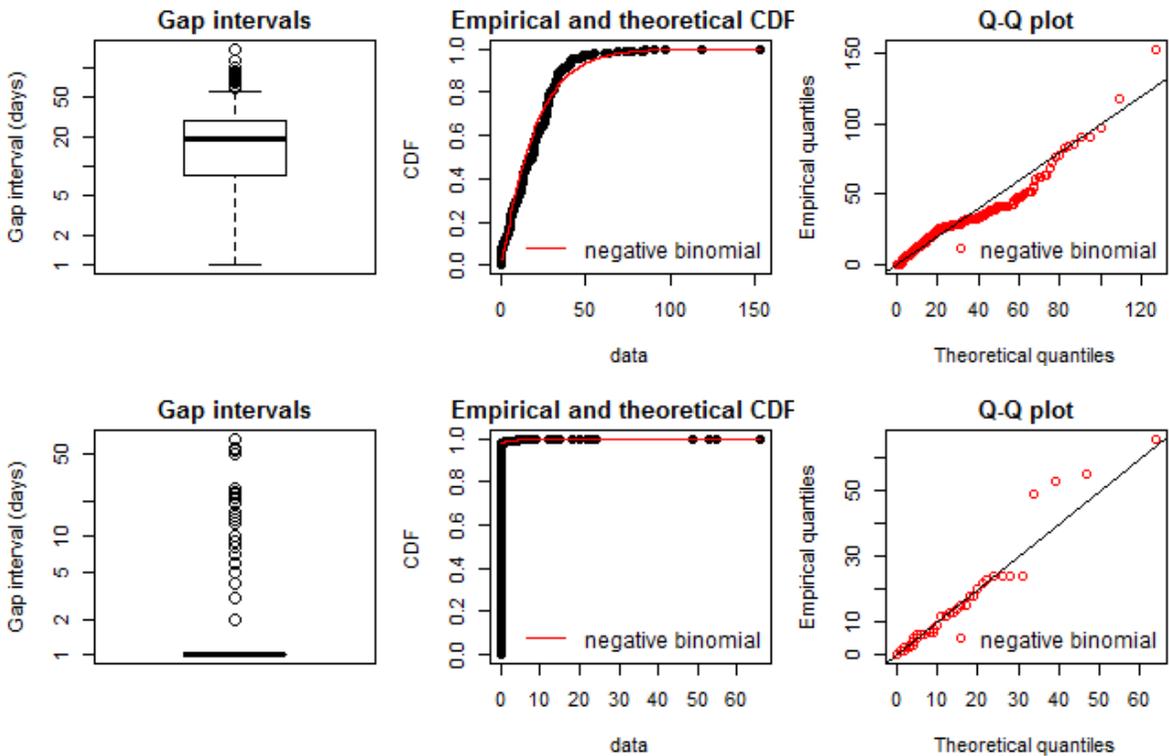


Figure 10.3. Examples of quantified sampling irregularity with negative binomial (NB) distributions: total nitrogen in Choptank River (top) and total phosphorus in Cuyahoga River (bottom). Theoretical CDF and quantiles are based on the fitted NB distributions. See Table 10.1 for estimated mean and shape parameters.

Table 10.1. Quantification of sampling irregularity for selected water-quality constituents at nine sites of the Chesapeake Bay River Input Monitoring program and six sites of the Lake Erie and Ohio tributary monitoring program. ( $\mu$ : mean parameter;  $\lambda$ : shape parameter estimated using maximum likelihood;  $\lambda'$ : shape parameter estimated using the direct approach (see Section 10.2.2).  $\Delta t_{average}$ : average gap interval;  $N$ : total number of samples.)

*I. Chesapeake Bay River Input Monitoring program*

Site ID	River and station name	Drainage area (mi <sup>2</sup> )	Total nitrogen (TN)					Total phosphorus (TP)				
			$\lambda$	$\lambda'$	$\mu$	$\Delta t_{average}$ (days)	$N$	$\lambda$	$\lambda'$	$\mu$	$\Delta t_{average}$ (days)	$N$
01578310	Susquehanna River at Conowingo, MD	27100	0.8	1.1	13.5	14.5	876	0.8	1.0	13.4	14.4	881
01646580	Potomac River at Chain Bridge, Washington D.C.	11600	0.9	0.6	9.5	10.5	1385	1.1	1.0	24.4	25.4	579
02035000	James River at Cartersville, VA	6260	0.8	1.0	13.9	14.9	960	0.8	1.1	13.7	14.7	974
01668000	Rappahannock River near Fredericksburg, VA	1600	0.8	0.6	15.6	16.6	776	0.8	0.6	15.2	16.2	796
02041650	Appomattox River at Matoaca, VA	1340	0.8	0.8	15.1	16.1	798	0.8	0.8	14.9	15.9	810
01673000	Pamunkey River near Hanover, VA	1071	0.8	0.9	15.1	16.1	873	0.8	1.0	14.7	15.7	894
01674500	Mattaponi River near Beulahville, VA	601	0.7	0.9	14.3	15.3	810	0.8	0.9	14.2	15.2	820
01594440	Patuxent River at Bowie, MD	348	0.9	1.1	15.3	16.3	787	0.8	0.8	14.0	15.0	861
01491000	Choptank River near Greensboro, MD	113	1.2	1.5	19.6	20.6	680	1.1	1.0	20.5	21.5	690

*II. Lake Erie and Ohio tributary monitoring program*

Site ID	River and station name	Drainage area (mi <sup>2</sup> )	Nitrate-plus-nitrite (NO <sub>x</sub> )					Total phosphorus (TP)				
			$\lambda$	$\lambda'$	$\mu$	$\Delta t_{average}$ (days)	$N$	$\lambda$	$\lambda'$	$\mu$	$\Delta t_{average}$ (days)	$N$
04193500	Maumee River at Waterville, OH	6330	0.005	0.0003	0.19	1.19	9101	0.005	0.0003	0.19	1.19	9101
04198000	Sandusky River near Fremont, OH	1253	0.01	0.003	0.22	1.22	9641	0.01	0.003	0.22	1.22	9655
04208000	Cuyahoga River at Independence, OH	708	0.007	0.006	0.13	1.13	7421	0.007	0.006	0.13	1.13	7426
04212100	Grand River near Painesville, OH	686	0.01	0.005	0.21	1.21	5023	0.01	0.005	0.22	1.22	4994
04197100	Honey Creek at Melmore, OH	149	0.007	0.005	0.06	1.06	9914	0.007	0.005	0.06	1.06	9914
04197170	Rock Creek at Tiffin, OH	34.6	0.007	0.008	0.06	1.06	8422	0.007	0.008	0.06	1.06	8440

0.7 and 1.2 for TN and between 0.8 and 1.1 for TP (Table 10.1). These  $\lambda$  values are around 1.0, reflecting the fact that these sites have relatively even gap distributions (*i.e.*, relatively balanced counts of large and small gaps). The mean parameter  $\mu$  varies between 9.5 and 19.6 for TN and between 13.4 and 24.4 for TP in the Chesapeake monitoring network, corresponding to  $\Delta t_{average}$  of 10.5–20.6 days for TN and 14.4–25.4 days for TP, respectively. This is consistent with the fact that these sites have typically been sampled bi-monthly to monthly, along with additional sampling during stormflows (Langland *et al.*, 2012; Zhang *et al.*, 2015).

For the Lake Erie and Ohio tributary monitoring program (6 sites), the record of nitrate-plus-nitrite ( $\text{NO}_x$ ) and TP were examined. According to the maximum likelihood approach, the shape parameter  $\lambda$  is approximately 0.01 for both constituents (Table 10.1). These very low  $\lambda$  values occur because these time series contain a limited number of very large gaps, ranging from 35 days to 1109 days ( $\sim 3$  years). The mean parameter  $\mu$  varies between 0.06 and 0.22, corresponding to  $\Delta t_{average}$  of 1.06 and 1.22 days, respectively. This is consistent with fact that these sites have been sampled at a daily resolution with occasional missing values.

### **10.2.3. Simulation of Time Series with Irregular Sampling**

To evaluate the various LRD estimation methods, Monte Carlo simulation was used to produce time series that mimic the sampling irregularity observed in real water-quality monitoring data. We began by simulating regular (gap-free) time series using the fractional noise simulation method of Witt and Malamud (2013), which is based on inverse Fourier filtering of white noises. Our analysis showed this method performed well compared to other simulation methods for  $\beta$  values between 0 and 1 (data not

shown). In addition, it has the advantage of being able to simulate  $\beta$  values beyond this range. The noises simulated by the Witt and Malamud method, however, are band-limited to the Nyquist frequency (half of the sampling frequency) of the underlying white-noise time series, whereas true fractional noises would contain spectral power at all frequencies, extending well above the Nyquist frequency for any sampling. Thus these band-limited noises will be less susceptible to spectral aliasing than true fractional noises would be; see Kirchner (2005) for detailed discussions of the aliasing issue.

30 replicates of regular time series were produced for nine prescribed LRD strengths. In terms of spectral slope ( $\beta$ ), these LRD strengths vary from zero (white noise) to two (Brownian motion or “random walk”) with an increment of 0.25 (*i.e.*, 0, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, and 2). These regular time series each have a length ( $N$ ) of 9125, which can be interpreted as 25 years of regular daily samples (that is,  $\Delta t_{nominal} = 1$  day).

Each of the simulated regular time series was converted to irregular time series using gap intervals that were simulated with NB distributions. To make these gap intervals mimic those in typical river water-quality monitoring time series, representative NB parameters were chosen based on results from Section 10.2.2. Specifically,  $\mu$  was set at 1 and 14, corresponding to  $\Delta t_{average}$  of 2 days and 15 days respectively. For the shape parameter  $\lambda$ , we chose four values that span three orders of magnitude, *i.e.*, 0.001, 0.1, 1, and 10. Note that when  $\lambda = 1$  the generated time series corresponds to a Bernoulli process. With the chosen values, a total of eight scenarios were generated, which were implemented using the “*rnbinom*” function in the “*stats*” R package (R Development Core Team, 2014):

a.1)  $\mu = 1$  (*i.e.*,  $\Delta t_{average}/\Delta t_{nominal} = 2$ ),  $\lambda = 0.01$ ,

- a.2)  $\mu = 1, \lambda = 0.1,$
- a.3)  $\mu = 1, \lambda = 1,$
- a.4)  $\mu = 1, \lambda = 10,$
- b.1)  $\mu = 14$  (i.e.,  $\Delta t_{average} / \Delta t_{nominal} = 15$ ),  $\lambda = 0.01,$
- b.2)  $\mu = 14, \lambda = 0.1,$
- b.3)  $\mu = 14, \lambda = 1,$
- b.4)  $\mu = 14, \lambda = 10.$

Examples of these simulations are shown with boxplots in Figure 10.2.

### 10.3. Evaluation of Various LRD Estimation Methods for Irregular Time Series

#### 10.3.1. Summary of Estimation Methods

For the simulated data, LRD was estimated using two types of approaches. The first type uses 11 different interpolation methods (designated as B1-B11 below) to fill the data gaps, thus making the data regular and analyzable by traditional estimation methods:

- B1) Global mean: all missing values replaced with the mean of all observations;
- B2) Global median: all missing values replaced with the median of all observations;
- B3) Random replacement: all missing values replaced with observations randomly drawn (with replacement) from the time series;
- B4) Next observation carried backward: each missing value replaced with the next available observation;
- B5) Last observation carried forward: each missing value replaced with the preceding available observation;

- B6) Average of the two nearest samples: it replaces each missing value with the mean of its next and preceding available observations;
- B7) Lowess (locally weighted scatterplot smoothing) with smoothing span of 1: missing values replaced using fitted values from a lowess model determined using all available observations (Cleveland, 1981);
- B8) Lowess with smoothing span of 0.75: same as B7 except that the smoothing span is 75% of the available data (similar distinction follows for B9-B11);
- B9) Lowess with smoothing span of 50%;
- B10) Lowess with smoothing span of 30%;
- B11) Lowess with smoothing span of 10%;

B4 and B5 were implemented using the “*na.locf*” function in the “*zoo*” R package (Zeileis and Grothendieck, 2005). B7-B11 were implemented using the “*loess*” function in the “*stats*” R package (R Development Core Team, 2014). An illustration of the interpolation methods is provided in Figure 10.4. The interpolated data, along with the original regular data (designated as A1) were analyzed using the Whittle’s maximum likelihood method for LRD estimation, which was implemented using the “*FDWhittle*” function in the “*fractal*” R package (Constantine and Percival, 2014).

The second type of approaches estimates LRD in the irregularly sampled data directly, using several variants of the Lomb-Scargle periodogram (designated as C1a-C1c below), and a recently proposed wavelet-based method (designated as C2 below).

Specifically, these approaches are:

- C1a) Lomb-Scargle periodogram: the spectral density of the time series (with gaps) is estimated directly and the spectral slope is fit using all frequencies (Lomb, 1976;

Scargle, 1982). This is a classic method for examining periodicity in irregularly-sampled data, which is analogous to the more familiar fast Fourier transform method often used for regularly-sampled data;

C1b) Lomb-Scargle periodogram with 5% data: same as C1a except that the fitting of the spectral slope considers only the lowest 5% frequencies (Montanari *et al.*, 1999);

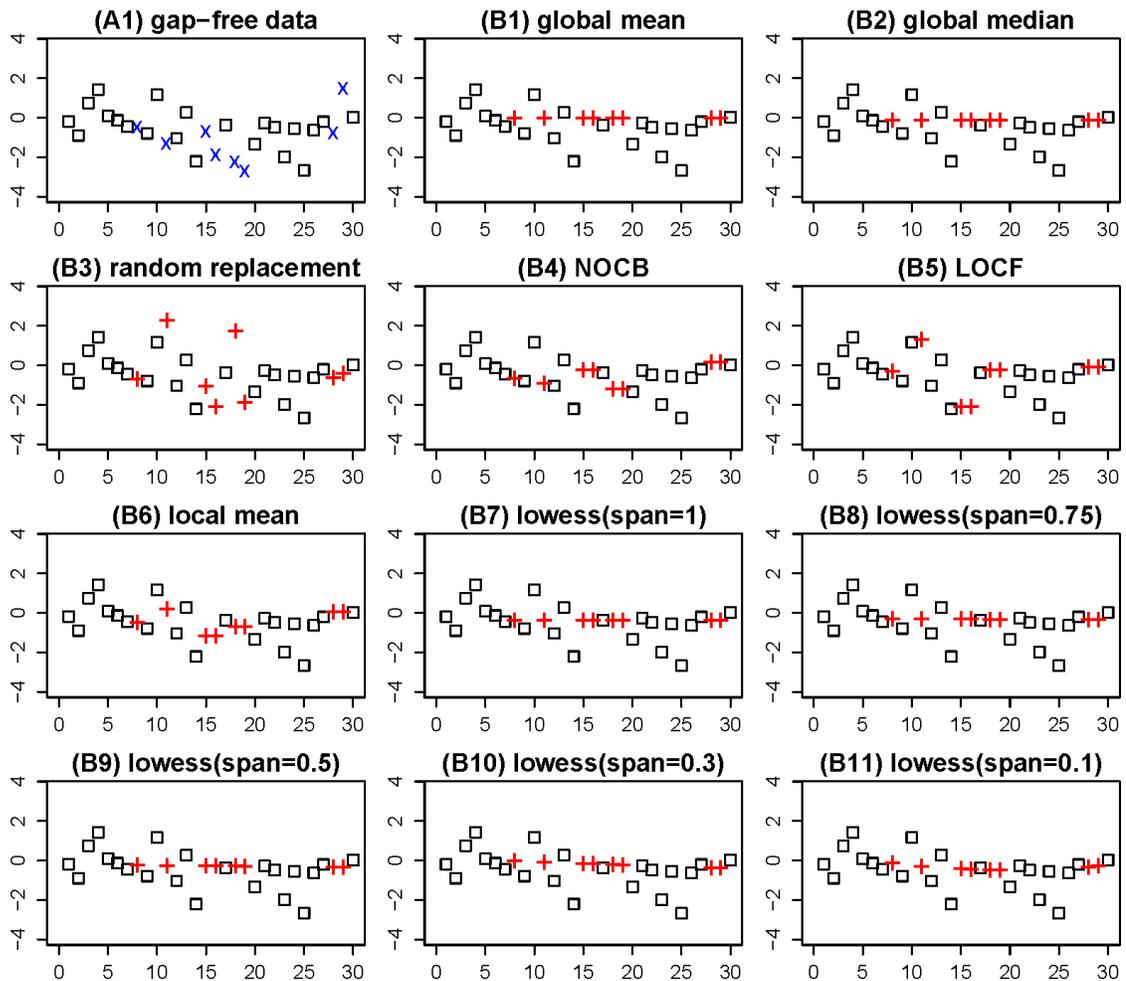


Figure 10.4. Illustration of the interpolation methods for gap-filling. The gap-free data (A1) was simulated with a series length of 500, with the first 30 data shown. (x: omitted data for gap-filling; +: interpolated data; NOCB: next observation carried backward; LOCF: last observation carried forward; lowess: locally weighted scatterplot smoothing.)

- C1c) Lomb-Scargle periodogram with “binned” data: same as C1a except that the fitting of the spectral slope is performed on binned data in three steps: (1) The entire range of frequency is divided into 100 equal-interval bins on logarithmic scale. (2) The respective medians of frequency and power spectral density are calculated for each of the 100 bins. (3) The 100 pairs of median frequency and median spectral density are used to estimate the spectral slope on a log-log scale.
- C2) Kirchner and Neal (2013)’s wavelet method: uses a modified version of Foster’s weighted wavelet spectrum (Foster, 1996) to suppress spectral leakage from low frequencies and applies an aliasing filter (Kirchner, 2005) to remove spectral aliasing artifacts at high frequencies.

C1a was implemented using the “*spec.ls*” function in the “*cts*” R package (Wang, 2013).

C2 was run in *C*, using codes modified from those in Kirchner and Neal (2013).

### 10.3.2. Evaluation of Method Performance

Each estimation method listed above was applied to the simulated irregular data (Section 10.2.3) to estimate LRD, which were then compared with the prescribed (“true”) LRD to quantify the performance of each method. Plots of method evaluation for all simulations are provided as Figures H1-H10 in Appendix H.

Close inspections of these plots reveal some general patterns of the methods’ performance. For brevity, these patterns are presented with a subset of the plots, which correspond to the simulation cases where true LRD  $\beta = 1$  and shape parameter  $\lambda = 0.01, 0.1, 1, \text{ and } 10$  (Figure 10.5). First of all, LRD values estimated using the regular data (A1) are generally very close to 1.0, which indicates that the adopted fractional noise

generation method and the Whittle's maximum likelihood estimator have small combined simulation and estimation bias. This is perhaps unsurprising, since the LRD estimator is based on the Fourier transform and the noise generator is based on an inverse

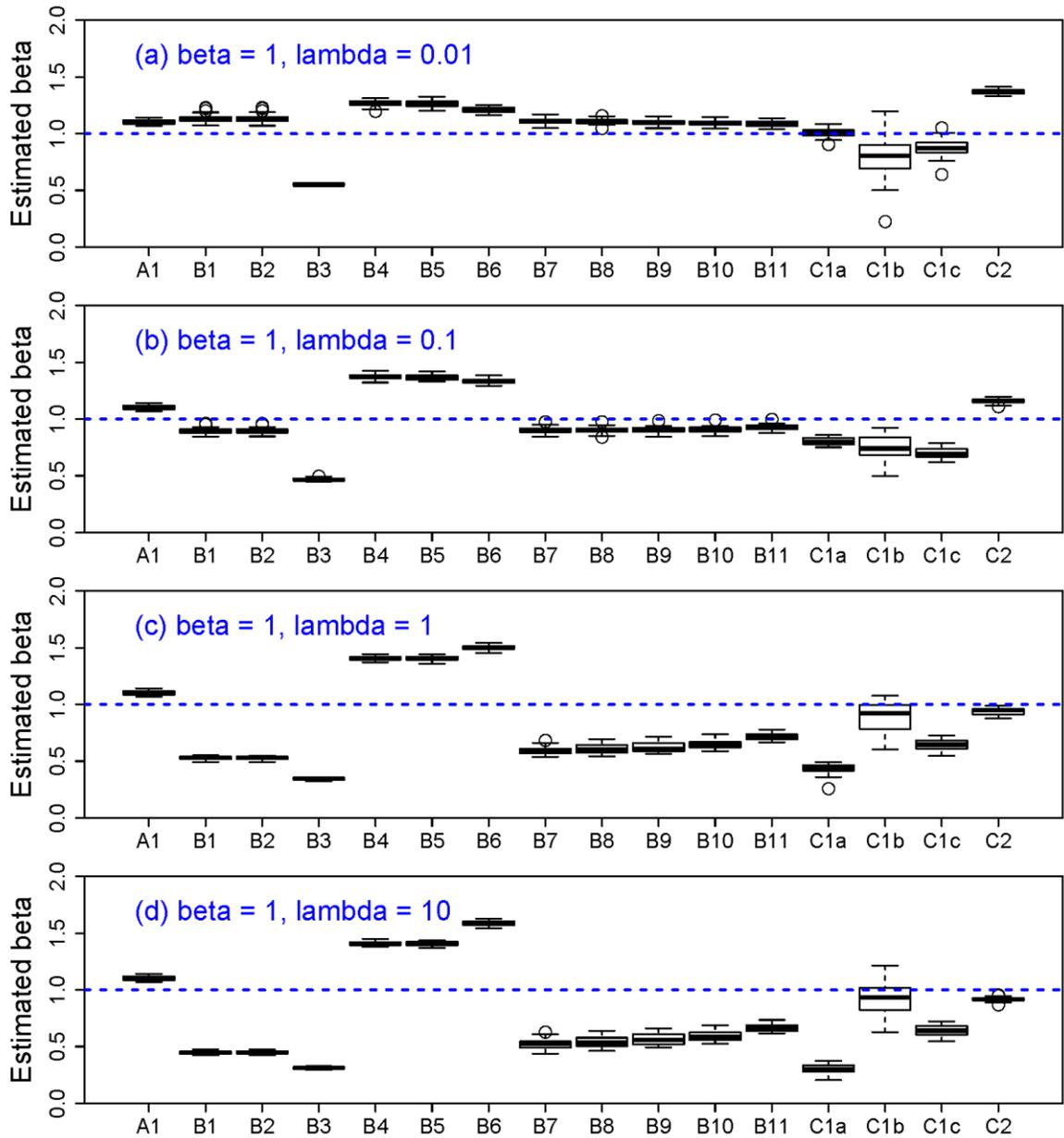


Figure 10.5. Comparison of bias in estimated LRD strengths in irregular data that are simulated with  $\beta = 1$  (30 replicates), series length of 9125, and gap intervals simulated with (a) NB ( $\lambda = 0.01, \mu = 1$ ), (b) NB ( $\lambda = 0.1, \mu = 1$ ), (c) NB ( $\lambda = 1, \mu = 1$ ), and (d) NB ( $\lambda = 10, \mu = 1$ ).

Fourier transform; thus one method is essentially just the inverse of the other. One should also note that when fractional noises are not arbitrarily band-limited at the Nyquist frequency (as they inherently are with the noise generator that is used here), spectral aliasing should lead to spectral slopes that are flatter than expected (Kirchner, 2005), and thus to underestimates of LRD.

For irregular data, the proposed estimation methods differ widely in their performance. Specifically, three interpolation methods (*i.e.*, B4-B6) consistently over-estimate LRD, indicating that they introduce additional correlations into the time series, reducing its short-timescale variability. In contrast, the other eight interpolation methods (*i.e.*, B1-B3 and B7-B11) consistently under-estimate LRD, indicating that the interpolated points are less correlated than the original time series, thus introducing additional variability on short time scales. As expected, results from the lowess methods (B7-B11) depend strongly on the size of smoothing window, that is, more severe under-estimation is produced as the smoothing window becomes wider. In fact, when the smoothing window is 1.0 (*i.e.*, method B7), lowess performs the interpolation using all data available and thus behaves similarly to interpolations based on global means (B1) or global medians (B2), except that lowess fits a polynomial surface instead of constant values. However, whenever a sampling gap is much shorter than the smoothing window, the in-filled lowess value will be close to the local mean or median, and the abrupt jumps produced by these in-filled values will artificially increase the variance in the time series at high frequencies, leading to an artificially reduced spectral slope  $\beta$  and an underestimate of LRD. This mechanism explains why lowess interpolation distorts  $\beta$

more when there are many small gaps (large  $\lambda$ ), and therefore more jumps to, and away from, the infilled values, than when there are only a few large gaps (small  $\lambda$ ).

Among the direct methods (C1a, C1b, C1c, and C2), the Lomb-Scargle method, with original data (C1a) or binned data (C1c) tends to under-estimate LRD, though the underestimation by C1c is generally less severe. The modified Lomb-Scargle method (C1b), using only the lowest 5% of frequencies, yields estimates that are centered around 1.0. However, C1b has the highest variability (*i.e.*, least precision) in LRD estimates among all methods evaluated. Compared with all the above methods, the wavelet method (C2) has much better performance in terms of both accuracy and precision when  $\lambda$  is 1 or 10 and similar performance when  $\lambda$  is 0.1.

The shape parameter  $\lambda$  affects the performance of the estimation methods. All the interpolation methods that under-estimate LRD (*i.e.*, B1-B3 and B7-B11) perform worse as  $\lambda$  increases from 0.01 to 10. This effect can be interpreted as follows: when the time series contains a large number of relatively small gaps (*e.g.*,  $\lambda = 1$  or 10), there are many jumps (which, as noted above, contain mostly high-frequency variance) between the original data and the in-filled values, resulting in more severe under-estimation. In contrast, when the data contain only a small number of very large gaps (*e.g.*,  $\lambda = 0.01$  or 0.1), there are fewer of these jumps, resulting in minimal under-estimation. Similar effects of  $\lambda$  are also observed with the interpolation methods that show over-estimation (*i.e.*, B4-B6) – that is, over-estimation is more severe when  $\lambda$  is large. Similarly, the Lomb-Scargle method (C1a and C1c) performs worse (more serious underestimation) as  $\lambda$  increases. Finally, method C2 seems to perform the best when  $\lambda$  is large (1 or 10), but not well when  $\lambda$  is very small (0.01). This result highlights the sensitivity of the wavelet

method to the presence of large gaps in the time series. For such cases, a potentially more feasible approach is to break the whole time series into several segments (each without long gaps) and then apply the wavelet method (C2) to analyze each segment separately. If this can yield more accurate estimates, then further simulation experiments should be designed to systematically determine how long the gap needs to be to invoke such an approach.

Next, the method evaluation is extended to all the simulated LRD strengths, that is,  $\beta = 0, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75,$  and  $2$ . For ease of discussion, three quantitative criteria were proposed for evaluating performance, namely, bias (B), standard deviation (SD), and root-mean-squared error (RMSE), as defined below:

$$B_i = \bar{\beta}_i - \beta_{true} \quad (10.14)$$

$$SD_i = \sqrt{\frac{1}{29} \sum_{j=1}^{30} (\beta_{i,j} - \bar{\beta}_i)^2} \quad (10.15)$$

$$RMSE_i = \sqrt{B_i^2 + SD_i^2} \quad (10.16)$$

where  $\bar{\beta}_i$  is the mean of 30  $\beta$  values estimated by method  $i$ , and  $\beta_{true}$  is the prescribed  $\beta$  value for simulation of regular time series. In general, B and SD can be considered as the models' systematic error and random error, respectively, and RMSE serves as an integrated measure of both errors. For all evaluations, plots of bias and RMSE are provided in the main text. (Plots of SD are provided as Figures H5 and H10 for simulations with  $\mu = 1$  and  $\mu = 14$ , respectively.)

For simulations with  $\mu = 1$ , results of estimation bias and RMSE are summarized in Figure 10.6 and Figure 10.7, respectively. (More details are provided in Figures H1-H4

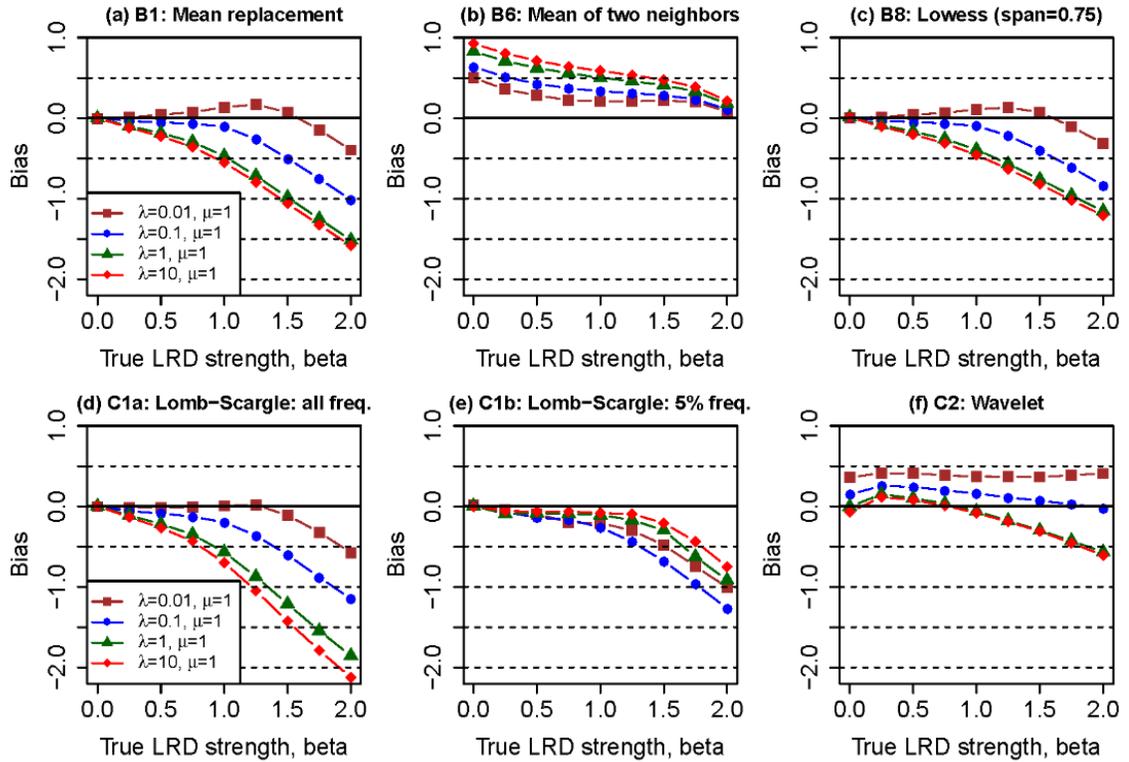


Figure 10.6. Comparison of bias in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 2.

in Appendix H.) For brevity, we focus on three direct methods (C1a, C1b and C2) and three representative interpolation methods. (Specifically, B1 represents B1-B3 and B7; B6 represents B4-B6, and B8 represents B8-B11.) Overall, these six methods show mixed performances. In terms of bias (Figure 10.6), B1 (global mean) and B8 (lowess with smoothing span of 0.75) tend to have negative bias, particularly for time series with (1) moderate-to-high  $\beta$  values and (1) high  $\lambda$  values (*i.e.*, less skewed gap intervals). B1 and B8 generally have minimal bias when (1)  $\beta$  is close to zero, since the simulated time series is close to white noise; and (2)  $\lambda$  is small (*e.g.*, 0.01), since interpolating a few large gaps cannot significantly affect the overall correlation structure. In addition, lowess interpolation with a larger smoothing window tends to yield more negatively biased

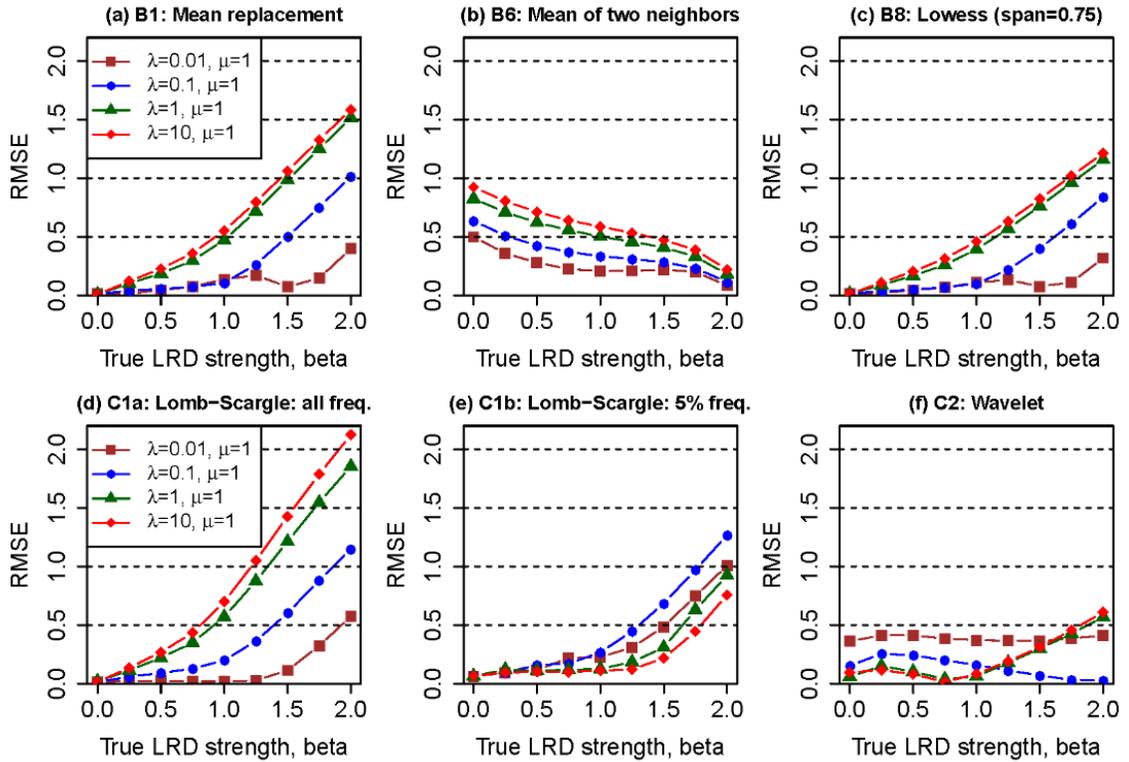


Figure 10.7. Comparison of root-mean-squared error (RMSE) in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 2.

estimates (data not shown). The other interpolation method, B6 (mean of two neighbors) tends to over-estimate  $\beta$ , particularly for time series with (1) low  $\beta$  values and (2) large  $\lambda$  values. At high  $\beta$  values (e.g., 2.0), the auto-correlation is already very strong such that taking the mean of two neighbors for gap-filling does not introduce much additional correlation, as opposed to the case of low  $\beta$  values. The Lomb-Scargle methods (C1a and C1b) generally have negative bias, particularly for time series with (1) moderate-to-high  $\beta$  values and (2) large  $\lambda$  values, which is similar to B1 and B8. In addition, C1b overall shows less severe bias than C1a. Finally, the wavelet method (C2) shows generally the smallest bias among all methods. However, its performance advantage is not as great when the time series has small  $\lambda$  values (i.e., very skewed gap intervals), as noted above,

which may be due to the fact that the aliasing filter was designed for regular time series.

In terms of SD (Figure H5), method C1b performs the worst among all methods (as noted above), method B6 and B8 perform poorly for large  $\beta$  values, and method C2 performs poorly for  $\beta = 0$ . In terms of RMSE (Figure 10.7), methods B1, B8, C1a, and C1b perform well for low  $\beta$  values and small  $\lambda$  values, whereas method B6 performs well for high  $\beta$  values and small  $\lambda$  values. In comparison, method C2 has the smallest RMSEs among all methods, and its RMSEs are similarly small for all evaluated  $\beta$  and  $\lambda$  values. In general, the wavelet method can be considered the most robust among all methods.

For simulations with  $\mu = 14$ , results of estimation bias and RMSE are summarized in Figure 10.8 and Figure 10.9, respectively. (More details are provided in Figures H6-H9

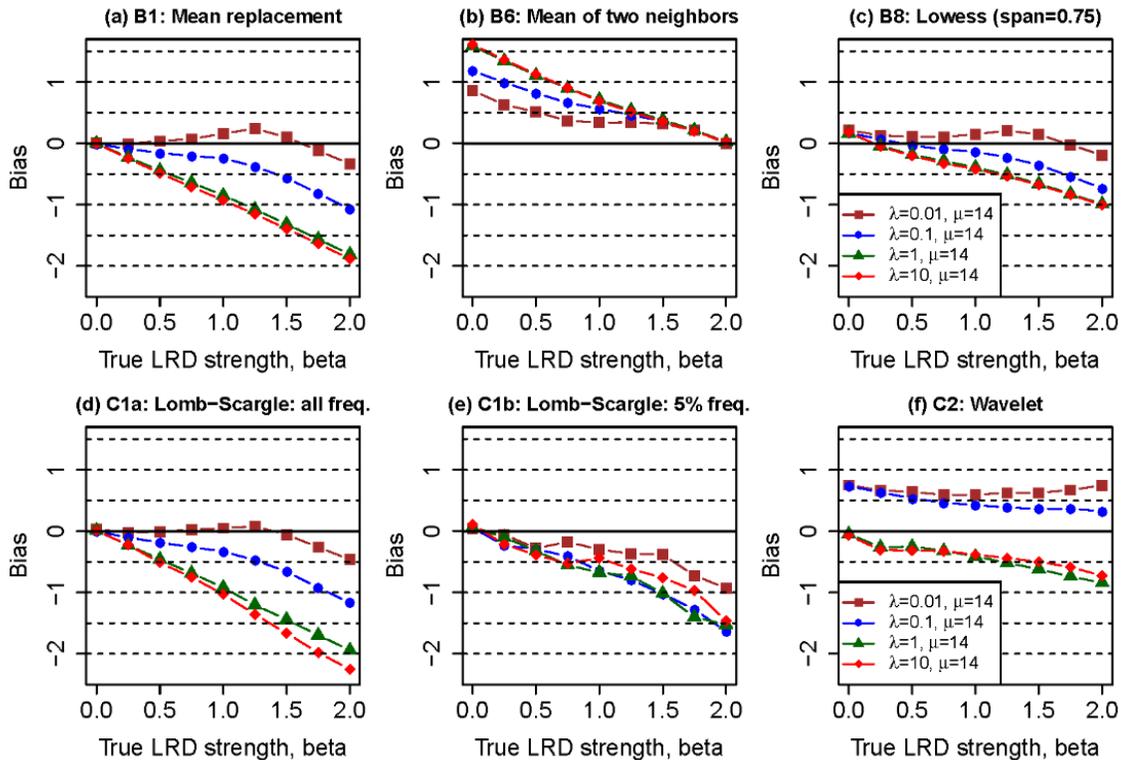


Figure 10.8. Comparison of bias in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 15.

in Appendix H.) Overall, these methods show mixed performances that are generally similar to the cases when  $\mu = 1$ , as discussed above. These results highlight the generality of these methods' estimation performances, which applies at least to the range of  $\mu = [1, 14]$ . In addition, all methods show generally larger RMSE for  $\mu = 14$  than  $\mu = 1$ , indicating their dependence on the mean gap interval (Figure 10.9). Perhaps the most notable difference is observed with method C2, which in this case shows positive bias for low  $\lambda$  values (0.01 and 0.1) and negative bias for high  $\lambda$  values (1 and 10) (Figure 10.8f). It nonetheless generally shows the smallest RMSEs among all the tested methods.

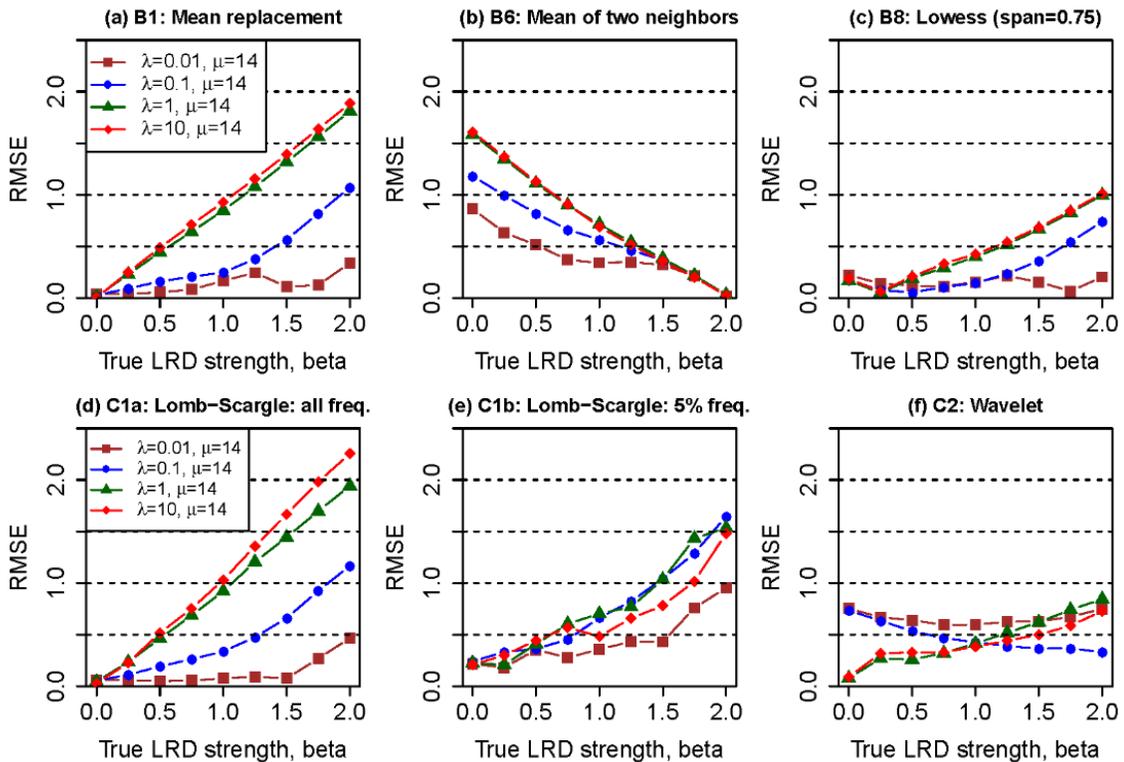


Figure 10.9. Comparison of root-mean-squared error (RMSE) in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 15.

### 10.3.3. Quantification of LRD in Real Water-Quality Data

In this section, the proposed estimation approaches were applied to quantify LRD in real water-quality data from the two regional monitoring programs presented in Section 10.2.2 (Table 10.1). As noted above in Section 10.1.3, such real-world data are typically much more complex than our simulated time series, because of (1) strong deviations from normal distributions and (2) effects of flow-dependence, seasonality, and temporal trend (Hirsch *et al.*, 1991; Helsel and Hirsch, 2002). In this regard, future research may simulate time series with these important characteristics and evaluate the performance of various LRD estimation approaches, perhaps following the framework described herein. Alternatively, one may quantify LRD in transformed time series after accounting for the above issues. In this work, we have taken the latter approach and made necessary transformation of the original concentration records using the recently-developed Weighted Regressions on Time, Discharge, and Season (WRTDS) method. This method is originally intended to estimate daily concentrations and loadings based on low-frequency concentration samples using time, season, and discharge as explanatory variables, *i.e.*,

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (10.17)$$

where  $C$  is concentration,  $Q$  is daily discharge,  $t$  is time in decimal years,  $\beta_i$  are fitted coefficients, and  $\varepsilon$  is the error term. The 2<sup>nd</sup> and 3<sup>rd</sup> terms on the right represent time and discharge effects, respectively, whereas the 4<sup>th</sup> and 5<sup>th</sup> terms collectively represent cyclical seasonal effects. For full details of this method, see Hirsch and De Cicco (2015). In brief, WRTDS develops a single model for each day of estimation, in which it pre-screens all available concentration samples and selects the most relevant samples to fit

Equation (10.17), with the “relevancy” being quantified and compared on three dimensions, *i.e.*, time, discharge, and season. The fitted coefficients are then used to estimate concentration for that particular day of estimation. In this work, WRTDS was applied to obtain the time series of estimated concentration for each constituent at each site. The difference between observed concentration ( $C_{obs}$ ) and estimated concentration ( $C_{est}$ ) was calculated in logarithmic space to obtain concentration residual, *i.e.*,

$$residuals = \ln(C_{obs}) - \ln(C_{est}) \quad (10.18)$$

Compared with the original concentration data, the residuals are more nearly normal and homoscedastic. For our data sets, histograms of concentration residuals (expressed in natural log concentration units) are shown in Figures H11-H14.

The estimated LRD values for the concentration residuals are summarized in Figure 10.10. Clearly, the estimated LRD varies considerably with the estimation method. In addition, the estimated LRD varies with site and constituent (*i.e.*, TP, TN, or NO<sub>x</sub>.) Our discussion below focuses on the wavelet method (C2), because it is established above that this method performs better than the other estimation methods under a wide range of gap conditions. While it is beyond our current scope to accurately quantify LRD in these water-quality data, our simulation results presented above (Section 10.3.2) can be used as references to qualitatively evaluate the reliability of C2 and/or other methods for these real-world data sets.

For TN and TP concentration data at the Chesapeake River Input Monitoring sites (Table 10.1),  $\mu$  varies between 9.5 and 24.4, whereas  $\lambda$  is  $\sim 1.0$ . Thus, the simulated gap scenario of NB( $\mu = 14, \lambda = 1$ ) can be used as a reasonable reference to assess methods' reliability (Figure 10.8). Based on method C2, the estimated LRD ranges between  $\beta =$

0.36 and  $\beta = 0.61$  for TN and between  $\beta = 0.30$  and  $\beta = 0.58$  for TP at these sites (Figure 10.10). For such ranges, the simulation results indicate that method C2 tends to moderately under-estimate LRD under this gap scenario (Figure 10.8), and hence LRD strengths at these Chesapeake sites are likely slightly higher than those presented above.

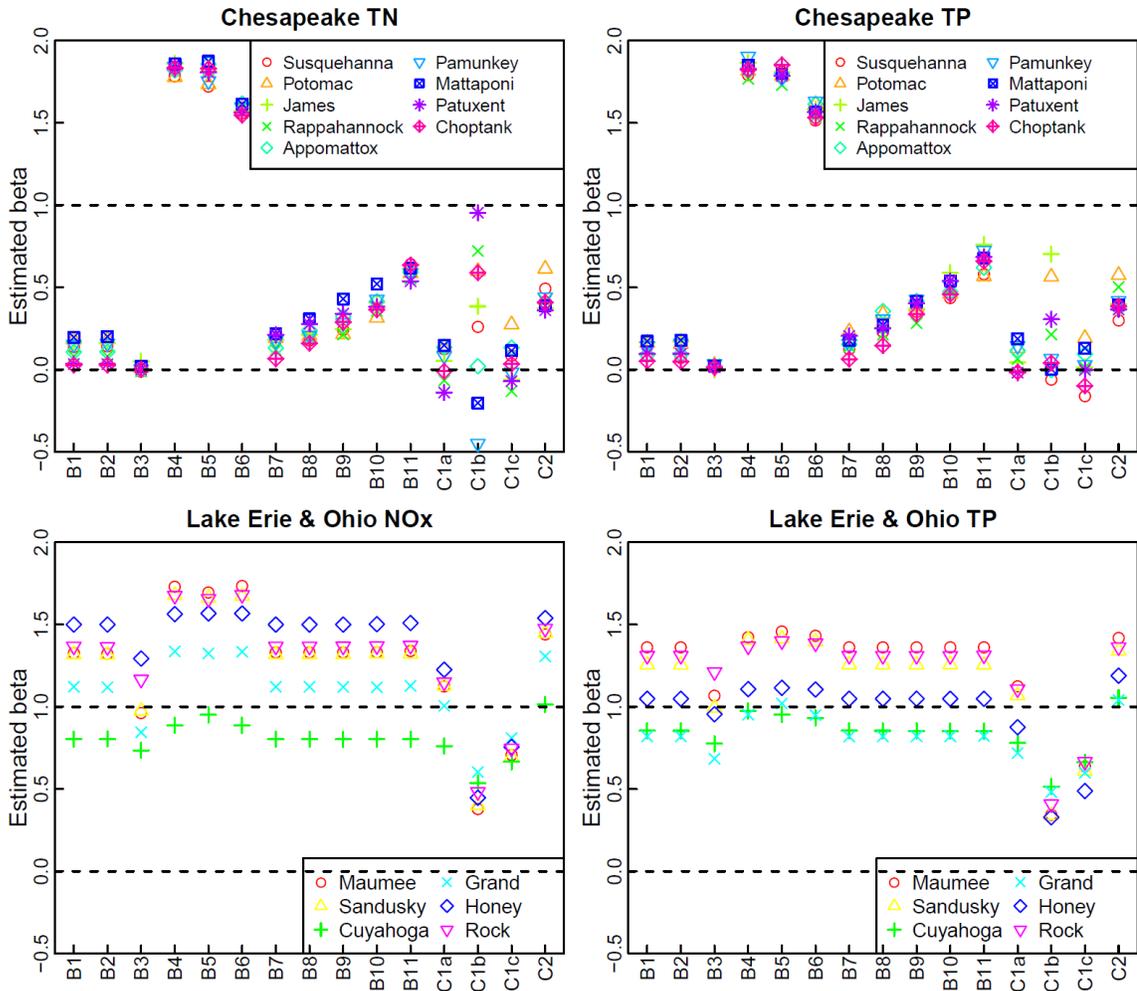


Figure 10.10. Quantification of LRD strengths in real water-quality data from the two regional monitoring networks, as estimated using the set of examined methods. All estimations were performed on concentration residuals (in natural log concentration units) after accounting for effects of time, discharge, and season. The two dashed lines in each panel indicate white noise ( $\beta = 0$ ) and flicker noise ( $\beta = 1$ ), respectively. See Table 10.1 for site and data details.

For NO<sub>x</sub> and TP concentration data at the Lake Erie and Ohio sites (Table 10.1),  $\mu$  varies between 0.06 and 0.22, whereas  $\lambda$  is  $\sim 0.01$ . Thus, the simulated gap scenario of NB( $\mu = 1, \lambda = 0.01$ ) can be used as a reasonable reference to assess methods' reliability (Figure 10.6). For such small  $\lambda$  (*i.e.*, a few gaps that are very dissimilar from others), C2 is not reliable for LRD estimation, as reflected by the generally positive bias in the simulation results. By contrast, methods B1 (interpolation with global mean) and B8 (LOWESS with span 0.75) both perform quite well under this gap scenario (Figure 10.6). These methods provide almost identical LRD estimates for each site-constituent combination, ranging from  $\beta = 1.0$  to  $\beta = 1.5$  for NO<sub>x</sub> and from  $\beta = 1.0$  to  $\beta = 1.4$  for TP (Figure 10.10).

Overall, the above analysis of real water-quality data has illustrated the wide variability of LRD estimates, with different choices of estimation methods yielding very different results. To our knowledge, these water-quality data have not heretofore been analyzed in this context. Our simulation experiments (Section 10.3.2) can be used as references to coarsely evaluate the reliability of each method under specific gap scenarios, thereby considerably narrowing the likely range of the estimated LRD strengths. Future research may consider applying Bayesian statistical analysis or other approaches to more accurately quantify the LRD strength and associated uncertainty for real water-quality data. Nonetheless, the results presented herein demonstrate that the analyzed water-quality time series can exhibit strong LRD, particularly at the Lake Erie and Ohio tributary sites. Thus, an important implication is that analysts should be cautious when applying standard statistical methods to analyze temporal trends in such water-quality data sets (Kirchner and Neal, 2013). In future work, our analyses presented

herein may be further extended to simulations of irregular time series that have prescribed LRD strengths and also superimposed temporal trends, which can then be used to evaluate and compare the validity of various statistical methods for identification of trend and the associated statistical significance.

#### **10.4. Summary and Prospects**

River water-quality time series have been observed to often exhibit long-range dependence, which presents challenges to the identification of deterministic trends. Because traditional LRD estimation methods are generally not applicable to irregular time series, we have examined two broad types of alternative estimation approaches and evaluated their performances against synthetic data with a wide range of prescribed LRD values and gap intervals representative of the sampling irregularity of real water-quality data.

This work has several important implications. First, the results remind us of the risks in using interpolation for gap-filling when examining auto-correlation, as the interpolation methods consistently under-estimate or over-estimate LRDs under a wide range of LRDs and gap distributions. Second, the long-established Lomb-Scargle spectral method also consistently under-estimates LRD. Its modified form, using the 5% lowest frequencies for spectral slope estimation, has very poor precision, although the overall bias is small. Third, the wavelet method, coupled with an aliasing filter, has the smallest bias and root-mean-squared error among all methods for a wide range of LRDs and gap distributions, except for small LRDs or very skewed gap distributions. Thus, the wavelet method is recommended for estimating LRD in irregular time series until improved methods are developed. Nonetheless, it should be noted that the wavelet method does not

perform as well when the time series has small  $\lambda$  values (*i.e.*, very skewed gap intervals) or small LRDs (*i.e.*, close to white noise). Thus, future research should aim to develop an aliasing filter that is more applicable to irregular time series. Finally, all methods' performances depend strongly on the sampling irregularity in terms of both the skewness and mean of gap-interval lengths, highlighting that the accuracy and precision of each method are data-specific.

Overall, these results provide new contributions in terms of better understanding and quantification of the proposed methods' performances for estimating LRD in irregularly sampled water-quality data. More generally, the findings and approaches may also be applicable to irregularly sampled data in other scientific fields. Last but not least, we note that accurate quantification of LRD in irregular water-quality time series remains an unresolved challenge for the hydrologic community and for other disciplines that must grapple with irregular sampling.

### **10.5. Supporting Information**

Supporting information to this chapter is provided in Appendix H.

### **10.6. Acknowledgements**

Q. Zhang was supported by the Maryland Sea Grant through awards NA10OAR4170072 and NA14OAR1470090 and by the Maryland Water Resources Research Center through a graduate fellowship. C. Harman's contribution to this work was supported by the National Science Foundation through grants CBET-1360415 and EAR-1344664. We thank Bill Ball (Johns Hopkins University) and Bob Hirsch (U.S. Geological Survey) for many useful discussions.

## 10.7. Literature Cited

- Beran, J., 2010. Long-range dependence. *Wiley Interdisciplinary Reviews: Computational Statistics* 2:26-35, DOI: 10.1002/wics.52.
- Beran, J., Y. Feng, S. Ghosh and R. Kulik, 2013. *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Berlin, Heidelberg, Springer Berlin Heidelberg, ISBN 978-3-642-35511-0
- Boutahar, M., V. Marimoutou and L. Nouria, 2007. Estimation Methods of the Long Memory Parameter: Monte Carlo Analysis and Application. *Journal of Applied Statistics* 34:261-301, DOI: 10.1080/02664760601004874.
- Box, G. E. P., G. M. Jenkins and G. C. Reinsel, 2008. *Time Series Analysis, Fourth Edition*. Hoboken, NJ, John Wiley & Sons, Inc., ISBN 9781118619193.
- Clarke, R. T., 2013. Calculating uncertainty in regional estimates of trend in streamflow with both serial and spatial correlations. *Water Resour. Res.* 49:7120-7125, DOI: 10.1002/wrcr.20465.
- Cleveland, W. S., 1981. LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *American Statistician* 35:54, DOI: 10.2307/2683591.
- Cohn, T. A. and H. F. Lins, 2005. Nature's style: Naturally trendy. *Geophys. Res. Lett.* 32:L23402, DOI: 10.1029/2005GL024476.
- Constantine, W. and D. Percival, 2014. *fractal: Fractal Time Series Modeling and Analysis*,
- Darken, P. F., C. E. Zipper, G. I. Holtzman and E. P. Smith, 2002. Serial correlation in water quality variables: Estimation and implications for trend analysis. *Water Resour. Res.* 38:1117, DOI: 10.1029/2001WR001065.

- Delignette-Muller, M. L. and C. Dutang, 2015. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* 64:1-34.
- Ehsanzadeh, E. and K. Adamowski, 2010. Trends in timing of low stream flows in Canada: impact of autocorrelation and long-term persistence. *Hydrol. Process.* 24:970-980, DOI: 10.1002/hyp.7533.
- Fatichi, S., S. M. Barbosa, E. Caporali and M. E. Silva, 2009. Deterministic versus stochastic trends: Detection and challenges. *Journal of Geophysical Research* 114:D18121, DOI: 10.1029/2009JD011960.
- Foster, G., 1996. Wavelets for period analysis of unevenly sampled time series. *The Astronomical Journal* 112:1709-1729, <http://articles.adsabs.harvard.edu/full/1996AJ....112.1709F>.
- Franzke, C., 2012a. Nonlinear Trends, Long-Range Dependence, and Climate Noise Properties of Surface Temperature. *J. Clim.* 25:4172-4183, DOI: 10.1175/JCLI-D-11-00293.1.
- Franzke, C., 2012b. On the statistical significance of surface air temperature trends in the Eurasian Arctic region. *Geophys. Res. Lett.* 39:L23705, DOI: 10.1029/2012GL054244.
- Godsey, S. E., W. Aas, T. A. Clair, H. A. de Wit, I. J. Fernandez, J. S. Kahl, I. A. Malcolm, C. Neal, M. Neal, S. J. Nelson, S. A. Norton, M. C. Palucis, B. L. Skjelkvåle, C. Soulsby, D. Tetzlaff and J. W. Kirchner, 2010. Generality of fractal 1/f scaling in catchment tracer time series, and its implications for catchment travel time distributions. *Hydrol. Process.* 24:1660-1671, DOI: 10.1002/hyp.7677.

- Graham, J., 2009. Missing Data Analysis: Making It Work in the Real World. *Annu. Rev. Psychol.* 60:549-576, DOI: 10.1146/annurev.psych.58.110405.085530.
- Helsel, D. R. and R. M. Hirsch, 2002. Statistical Methods in Water Resources. *U.S. Geological Survey Techniques of Water-Resources Investigations Book 4, Chapter A3*. U.S. Geological Survey, Reston, VA, p. 522.  
<http://pubs.usgs.gov/twri/twri4a3/>.
- Hirsch, R. M., R. B. Alexander and R. A. Smith, 1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resour. Res.* 27:803-813, DOI: 10.1029/91WR00259.
- Hirsch, R. M. and L. De Cicco, 2015. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). *U.S. Geological Survey Techniques and Methods Book 4, Chapter A10*, Reston, VA, p. 93. <http://dx.doi.org/10.3133/tm4A10>.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Hurst, H. E., 1951. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.* 116:770-808.
- Khaliq, M. N., T. B. M. J. Ouarda and P. Gachon, 2009. Identification of temporal trends in annual and seasonal low flows occurring in Canadian rivers: The effect of short- and long-term persistence. *Journal of Hydrology* 369:183-197, DOI: 10.1016/j.jhydrol.2009.02.045.

- Khaliq, M. N., T. B. M. J. Ouarda, P. Gachon and L. Sushama, 2008. Temporal evolution of low-flow regimes in Canadian rivers. *Water Resour. Res.* 44:W08436, DOI: 10.1029/2007WR006132.
- Kirchner, J., 2005. Aliasing in  $1/f^\alpha$  noise spectra: Origins, consequences, and remedies. *Physical Review E* 71:066110-066110, DOI: 10.1103/PhysRevE.71.066110.
- Kirchner, J. W., X. Feng and C. Neal, 2000. Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature* 403:524-527, DOI: 10.1038/35000537.
- Kirchner, J. W., X. Feng and C. Neal, 2001. Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations. *Journal of Hydrology* 254:82-101, DOI: 10.1016/S0022-1694(01)00487-5.
- Kirchner, J. W. and C. Neal, 2013. Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc. Natl. Acad. Sci. U. S. A.* 110:12213-12218, DOI: 10.1073/pnas.1304328110.
- Kirchner, J. W. and A. Weil, 1998. No fractals in fossil extinction statistics. *Nature* 395:337-338, DOI: 10.1038/26384.
- Langland, M. J., J. D. Blomquist, D. L. Moyer and K. E. Hyer, 2012. Nutrient and suspended-sediment trends, loads, and yields and development of an indicator of streamwater quality at nontidal sites in the Chesapeake Bay watershed, 1985-2010. U.S. Geological Survey Scientific Investigations Report 2012-5093, Reston, VA, p. 26. <http://pubs.usgs.gov/sir/2012/5093/pdf/sir2012-5093.pdf>.

- Lennartz, S. and A. Bunde, 2009. Trend evaluation in records with long-term memory: Application to global warming. *Geophys. Res. Lett.* 36:L16706, DOI: 10.1029/2009GL039516.
- Lomb, N. R., 1976. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science* 39:447-462, DOI: 10.1007/BF00648343.
- Montanari, A., R. Rosso and M. S. Taqqu, 1997. Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resour. Res.* 33:1035-1044, DOI: 10.1029/97WR00043.
- Montanari, A., R. Rosso and M. S. Taqqu, 2000. A seasonal fractional ARIMA Model applied to the Nile River monthly flows at Aswan. *Water Resour. Res.* 36:1249-1259, DOI: 10.1029/2000WR900012.
- Montanari, A., M. S. Taqqu and V. Teverovsky, 1999. Estimating long-range dependence in the presence of periodicity: An empirical study. *Mathematical and Computer Modelling* 29:217-228, DOI: 10.1016/S0895-7177(99)00104-1.
- National Center for Water Quality Research, 2015. User's Guide to the River Data Sets. <http://www.heidelberg.edu/academiclife/distinctive/newqr/data>. Accessed July 23, 2015.
- Noguchi, K., Y. R. Gel and C. R. Duguay, 2011. Bootstrap-based tests for trends in hydrological time series, with application to ice phenology data. *Journal of Hydrology* 410:150-161, DOI: 10.1016/j.jhydrol.2011.09.008.
- R Development Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3900051070. <http://www.r-project.org>.

- Rea, W., L. Oxley, M. Reale and J. Brown, 2009. Estimators for Long Range Dependence: An Empirical Study. *Electronic Journal of Statistics*, <http://arxiv.org/abs/0901.0762>.
- Sang, Y.-F., Z. Wang and C. Liu, 2014. Comparison of the MK test and EMD method for trend identification in hydrological time series. *Journal of Hydrology* 510:293-298, DOI: 10.1016/j.jhydrol.2013.12.039.
- Scargle, J. D., 1982. Studies in Astronomical Time-Series Analysis. II. Statistical Aspects of Spectral-Analysis of Unevenly Spaced Data. *Astrophysical Journal* 263:835-853, DOI: 10.1086/160554.
- Stroe-Kunold, E., T. Stadnytska, J. Werner and S. Braun, 2009. Estimating long-range dependence in time series: an evaluation of estimators implemented in R. *Behav. Res. Methods* 41:909-923, DOI: 10.3758/BRM.41.3.909.
- Taqqu, M. S., V. Teverovsky and W. Willinger, 1995. Estimators for long-range dependence: an empirical study. *Fractals*, <http://www.worldscientific.com/doi/abs/10.1142/S0218348X95000692>.
- Wang, Z., 2013. cts: An R Package for Continuous Time Autoregressive Models via Kalman Filter. *Journal of Statistical Software* 53:1-19.
- Witt, A. and B. D. Malamud, 2013. Quantification of Long-Range Persistence in Geophysical Time Series: Conventional and Benchmark-Based Improvement Techniques. *Surv. Geophys.* 34:541-651, DOI: 10.1007/s10712-012-9217-8.
- Yue, S., P. Pilon, B. Phinney and G. Cavadias, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Process.* 16:1807-1829, DOI: 10.1002/hyp.1095.

- Zeileis, A. and G. Grothendieck, 2005. zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software* 14:1-27.
- Zetterqvist, L., 1991. Statistical Estimation and Interpretation of Trends in Water Quality Time Series. *Water Resour. Res.* 27:1637-1648, DOI: 10.1029/91wr00478.
- Zhang, Q., D. C. Brady, W. Boynton and W. P. Ball, 2015. Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake Watershed: An Assessment of Progress by River and Season. *J. Am. Water Resour. Assoc.* 51:1534-1555, DOI: 10.1111/1752-1688.12327.

## **Chapter 11. Summary, Application, and Future Work**

### **11.1. Summary and Implications**

This research has provided new understanding of nutrient and sediment export from the watershed through WRTDS-based retrospective analysis and synthesis of decadal-scale records, which have direct bearing toward better management of the Chesapeake watershed. In addition, this work has made important methodological advancements with respect to WRTDS estimation performance, interpretation of riverine concentration-discharge relationships, and quantification of long-range dependence in irregularly sampled data. Last but not least, results from this work effectively illustrate how science-based management can benefit from maintaining open-access, high-quality monitoring data at multiple locations in watersheds. Specific contributions are briefly summarized below for each chapter.

In Chapter 2, long-term seasonal trends of flow-normalized nutrient and sediment loads were analyzed for locations above and below Conowingo Reservoir. Loads of N, P, and SS have generally declined in the SRB above the reservoir, which can most likely be attributed to a suite of management control actions on point, agricultural, and stormwater sources. By contrast, loads of SS, PP, and PN at Conowingo Dam have been generally rising since the mid-1990s. The reservoir's capacity to trap these materials has been diminishing, and it has neared its sediment storage capacity.

In Chapter 3, sediment and nutrient data at sites above and below the reservoir system were further analyzed for the period of 1986-2013 in a mass-balance context and as a function of river discharge. Results consistently show declined net deposition rates of SS

and TP in recent years under a wide range of flow conditions that include flows well below the literature-reported scour threshold. Future progress in Chesapeake Bay restoration will depend on accurate predictions of how inputs of SS, TP, and TN to the reservoirs will be modulated by sediment accumulation and remobilization behind Conowingo Dam.

In Chapter 4, all seven sites of long-term monitoring in the Susquehanna River basin were analyzed for a tempo-spatial synthesis of nutrient and sediment export. Riverine loadings have generally declined at sites upstream of Conowingo Reservoir. These declines, particularly those of TN, seem to have followed source input reductions in the concurrent period of 1987-2011. For both TN and TP, however, the generally  $<1.0$   $\Delta_{\text{FN-Yield}}/\Delta_{\text{Input}}$  ratios suggest the possibility of legacy contributions and reinforce the importance of considering lag time for river quality improvement. The concentration-discharge patterns generally show chemostasis effects for dissolved species and mobilization effects for particulate species. Finally, long-term nutrient and sediment yields were found to be positively correlated with the area fraction of non-forested land.

In Chapter 5, analysis was extended to nine major tributaries from the 1970s to 2012. For the non-tidal Chesapeake Bay watershed (*i.e.*, sum of the nine tributaries), particulate nutrients and SS loads have risen since the mid-1990s. The majority of these rises were from Susquehanna River and relate to diminished net trapping at Conowingo Reservoir. Rising SS trends were also observed, however, in other rivers. Moreover, the summed rise in particulate P load from other rivers is of similar magnitude as from Susquehanna. Dissolved nutrient loads have dropped in the upland rivers, but risen in two small Coastal-Plain rivers affected by lagged groundwater input. The seven upland rivers

yielded surprisingly steady fractions of N, suggesting consistent input and modulation across the watershed. TN:TP ratios have declined in most rivers, suggesting the potential for changes in nutrient limitation in downstream estuaries. These results may have important implications for the coastal regions with respect to algal growth, hypoxia extent, and biogeochemical processes. Finally, seasonal comparisons revealed that load magnitudes have been generally highest in Jan-Mar and lowest in Jul-Sep, but the temporal trends have followed similar decadal-scale patterns in all seasons, with some notable exceptions. One example is Susquehanna River, where the sediment trend in summer (July-September) was remarkably different from the other seasons, owing to storm events and associated effects on reservoir deposition and remobilization. Another example is Choptank River, where particulate N load has increased in winter (January-March) but decreased in all other seasons. This remains unexplained but would be consistent with a hypothesis of intensified and/or earlier fertilizer applications in winter.

In Chapter 6, new uncertainty and sensitivity analyses were conducted to quantify the effects of historical extreme-flow concentration samples on the WRTDS-modeled results with respect to sediment and nutrients processing behind Conowingo Dam. These new modeling experiments reveal high uncertainties in estimated “true-condition” WRTDS loadings during the most extreme discharges and in estimated “flow-normalized” loadings during specific times and times of year that are proximate to those discharges. These uncertainties, however, do not have major impacts on our previously published conclusions – *i.e.*, those regarding the temporal trends of decreasing input, increasing output, and overall decreasing net deposition through the reservoir system. To the

contrary, these new modeling experiments have provided additional statistical evidence to support those prior general conclusions.

In Chapter 7, the long-standing topic of riverine concentration-discharge ( $C-Q$ ) relationships was revisited. Sediment data from Susquehanna River at Conowingo were used to illustrate several common issues, including non-linear  $\ln(C)-\ln(Q)$  relations; variations in  $C-Q$  relations over time and season; and different variations at different discharges. In this context, WRTDS is proposed to offer improved interpretation of  $C-Q$  relationships. Specifically, the  $\beta_2$  coefficients can be organized in various ways to provide new insights to the interpretation of river water-quality data. As demonstrated using the Conowingo example, the approach can provide more nuanced insights about long-term declines in reservoir trapping performance and perform uncertainty analyses on the results.

In Chapter 8, the proposed  $\beta_2$  approach was applied to interpret the nature and change of nutrient and sediment  $C-Q$  relationships for nine major Chesapeake tributaries. The analyses revealed that constituent export has been dominated by mobilization patterns for SS and TP and chemostasis patterns for TN under many discharge conditions. The general lack of dilution patterns may suggest that none of these constituents has been supply-limited in these watersheds. SS and TP coefficients have generally followed a clear positive monotonic pattern with respect to discharge, exhibiting threshold behaviors. In many cases, coefficients show clear temporal non-stationarity under selected discharges, possibly reflecting changes in dominant watershed sources due to management actions.

In Chapter 9, improved WRTDS models for riverine concentration and flux estimation were developed by accounting for antecedent discharge conditions. The modified models show improvement over the original WRTDS model under all three sampling strategies considered. Major improvements were achieved for  $\text{NO}_x$  by the long-term flow-anomaly (LTFA) model and for chloride by the average discounted flow (ADF) model and the short-term flow-anomaly (STFA) model. Moderate improvements were achieved for SS, TP, and TKN by the ADF model. By contrast, no such achievement was achieved for SRP by any proposed model. Among the three alternative sampling strategies, performance of all models (including the original) was found to be the best using the sampling strategy that involved 12 regular samples and 8 storm samples per year (*i.e.*, sampling strategy “C”) and worst for the strategy that had the fewest samples and no storm sampling (*i.e.*, strategy “A”, or 12 random samples per year). Such sampling strategy effects were especially pronounced for SS, TP, and SRP, confirming the value of routinely collecting storm-flow samples.

In Chapter 10, a series of approaches for estimating long-range dependence (LRD) in irregular time series were evaluated. All methods’ performances depend strongly on the sampling irregularity (as quantified by both the skewness and mean of gap-interval lengths). The interpolation methods consistently under-estimate or over-estimate LRDs under a wide range of conditions of LRDs and gap distributions. The well-established Lomb-Scargle method was found to consistently under-estimates LRD. Its modified form, using the 5% lowest frequencies for spectral slope estimation, has very poor precision, although the overall bias is small. In comparison, the recently developed wavelet method (coupled with an aliasing filter) can better suppress spectral leakage and

aliasing artifacts. Among all methods considered, the wavelet method has the smallest bias and root-mean-squared error in most scenarios examined, although it does not perform as well when the series has very skewed gap intervals or small LRDs. Thus, the wavelet method is recommended for LRD estimation until improved methods are developed.

## **11.2. Application and Relevancy**

Overall, the body of research described in this dissertation has demonstrated the utility of various statistical modeling approaches toward large-scale analysis and synthesis of long-term river water-quality data at multiple watershed locations. The methods used and developed here are generally applicable and can be similarly conducted for major watersheds elsewhere, provided that water-quality constituent samples are adequate in both quantity and quality. Specifically, such water-quality data record is recommended to (1) cover a period of at least 20 years; (2) have an average sampling frequency of at least 10 days per year; (3) be accompanied with a complete record of daily discharge over the period of interest; and preferably, (4) contain a fair number of routinely-collected samples during high discharges. In general, these requirements can be met by monitoring data collected at many long-term sites in the United States and elsewhere. In this section, the application and relevancy of the completed studies are summarized.

The input-output analysis for the Lower Susquehanna River Reservoir System (Chapters 2-3) may be the first detailed mass-balance analysis of long-term sediment and nutrient accumulations for a major reservoir system that is close to the end of its period of effective sediment trapping. The Chesapeake Bay Program partnership should be able

to use these efforts in its on-going management efforts – specifically, the Bay Program will need a realistic representation of the current and future trapping efficiency of the reservoir system, preferably as a function of discharge and storage capacity, in order to improve its modeling of sediment and nutrient loadings from Susquehanna River to Chesapeake Bay. Our results may be the best existing information to describe how the reservoir system has historically transformed inputs into outputs and how this modulation has evolved over the last few decades. Specially, our results can help constrain and inform the development and application of improved predictive models of reservoir performance, and particularly the incorporation of such models in the ongoing upgrade of the Chesapeake Bay Partnership’s Watershed Model (Linker *et al.*, 2013; Shenk and Linker, 2013). More broadly, our approach and findings may be generally applicable to other reservoir systems that may be similarly approaching a state of dynamic equilibrium with respect to sediment storage.

The multi-site analysis for the Susquehanna River basin (Chapter 4) illustrates how relatively simple and traditional mass-balance and concentration-discharge approaches can be used to understand important patterns and controls of constituent export (*i.e.*, source inputs, reservoirs, streamflow, and land use). This examination of various types of data for the Susquehanna River basin may foster comparisons with rivers in other geographical regions (within the Chesapeake watershed and beyond) for better understanding patterns and controls of constituent export.

The trend analysis of nine major non-tidal Chesapeake tributaries (Chapter 5) demonstrates how integrated comparison of historical data can provide new insights on the impacts of management practice and highlights the importance of maintaining long-

term monitoring at multiple watershed locations. Such insights are an important basis for future work for better understanding the impacts of numerous potential factors affecting water-quality trends. Our seasonal focus has provided some new insights on water-quality trends that would otherwise be obscured from annual trends, which are described with examples in Section 11.1 and in more details in Section 5.4 (“Comparison of Loads among Seasons”).

The analysis of uncertainties and sensitivities of the WRTDS method (Chapter 6) serves as an important case study that demonstrates the value of rigorous statistical approaches for analysis of sparse water-quality data while also emphasizing the uncertainties associated with the modeled results. The results confirm that WRTDS is a reasonably robust tool for inferring general water-quality conditions and trends from long-term data, but that uncertainties can sometimes be high and need to be recognized.

The analysis of  $C$ - $Q$  relationships using WRTDS regression coefficients (Chapters 7-8) demonstrates its capability to make robust and informative interpretation of  $C$ - $Q$  relationships in low-frequency monitoring data. This approach allows simultaneous examination of various discharge conditions within an uncertainty framework to provide sound conclusions. With a focus on the sensitivity of concentration to discharge, the WRTDS coefficients provide an alternative approach for interpretation of water-quality data and for generation of sensible hypotheses on the dominant watershed processes. The proposed approach can be implemented by running additional codes within the *EGRET* package. These codes are publicly available (Zhang and Ball, 2016), making the approach easily applicable to water-quality data sets at other monitoring sites.

The analysis with the modified WRTDS models (Chapter 9) has provided a comprehensive set of statistical evidence for incorporating antecedent discharge into the WRTDS model for concentration and flux estimation, thereby combining the advantages of recent developments in regard to both the consideration of antecedent conditions and WRTDS modeling. The quantified effects of the proposed flow variables can inform the modification of WRTDS and/or other similar regression-based models. In addition, the evaluation of model estimation performance under three different sampling strategies has provided evidence on the value of routinely collecting storm-flow samples during river water-quality monitoring. Such insights are immediately relevant to many monitoring programs with respect to water-quality data collection and analysis. I would recommend that additional studies be conducted using data from different types of watershed to place these results into a more generalizable context that can inform further improvement of estimation methods.

Finally, the analysis of LRD estimation (Chapter 10) illustrates the apparent limitations in using interpolation for gap-filling when examining auto-correlation. The results show that the wavelet method has the smallest bias and root-mean-squared error among all methods for a wide range of conditions of LRDs and gap distributions, except for small LRDs or very skewed gap distributions. These results provide new understanding and quantification of the methods' performances for estimating LRD in irregular water-quality data, which has immediate relevancy to the hydrological community as well as other scientific fields.

### 11.3. Recommendations for Future Work

The various studies described in this dissertation have potential for immediate application in several existing contexts, especially as related to management of the Chesapeake Bay watershed, and these completed studies also help to motivate future areas of research. Some recommendations for future work are described below.

The analysis of the evolution of the Lower Susquehanna River Reservoir System can help constrain and inform the development and application of improved predictive models of reservoir performance, and particularly the incorporation of such models in the ongoing upgrade of the Chesapeake Bay Partnership's Watershed Model (Linker *et al.*, 2013; Shenk and Linker, 2013). Future progress in Chesapeake Bay restoration will depend on accurate predictions of how inputs of SS, TP, and TN to the reservoirs will be modulated by reservoir processes. Management actions will need to be adjusted to reflect the future role that sediment accumulation and remobilization behind Conowingo Dam will have on the delivery of SS, TP, and TN to the Bay. Continued monitoring and modeling of this system will be crucial to estimating inputs to and outputs from the reservoirs as well as the uncertainties of these estimates, which form the basis to advance our understanding of this complex and ever evolving system (Hirsch, 2012).

The analysis of seasonal loading trends using WRTDS has advanced our understanding in nutrient and sediment export across the Chesapeake watershed, which complements and updates findings from prior studies (Hirsch *et al.*, 2010; Hirsch, 2012; Langland *et al.*, 2012; Moyer *et al.*, 2012). Our seasonal focus has provided some new insights on water-quality trends that would otherwise be obscured from annual trends. In future work, additional data compilation and analysis are needed to more thoroughly

understand why temporal trends have been different among seasons in some tributaries. In addition, our understandings of temporal trends are a critical basis for future work for better understanding the impacts of numerous potential driving factors, including spatial and temporal changes in land use, fertilizer and manure application rates, atmospheric deposition, and point source loadings. In this regard, a collaboration of federal, state, private, and academic partners from throughout the Chesapeake watershed has begun work within the Chesapeake Bay Partnership to seek explanations for the various loading changes and trends (Keisman *et al.*, 2015). Moreover, continued analysis will be needed to refine our understandings of the riverine loading trends as new monitoring data and modeling methods become available. In future work, investigators may consider adopting an integrated and holistic approach -- as demonstrated here in our analyses (Chapter 5) -- to understanding complex export behaviors across watersheds.

The analysis of  $C$ - $Q$  relationships using WRTDS coefficients demonstrates the capability of this approach for generating sensible hypotheses on the dominant watershed processes. The application of this approach toward 15 long-term sites in the Chesapeake Bay watershed has revealed diverse trends in WRTDS coefficients (*i.e.*, sensitivity of constituent concentration to discharge) under different discharge conditions, which possibly reflect changes in dominant watershed sources due to management actions. In this regard, further research is needed to systematically explore factors that have driven the observed functional changes under different discharge conditions for different watershed systems. Moreover, while this study was intended to seek common patterns among the 15 watersheds, further investigation is recommended to investigate factors that have driven the unique behavior of each individual watershed, which may include but not

limited to, flow-generation processes, flood-plain structure, groundwater and surface-water interactions, river flashiness, watershed input, density of reservoirs and dams, and land uses.

The analysis of riverine concentration/flux estimation using modified WRTDS models has provided concrete statistical evidence for incorporating antecedent discharge into the WRTDS model for concentration and flux estimation. Accurate estimation of riverine constituent concentrations and fluxes has been an active research area for decades and will remain a critical challenge to the hydrological community (Cohn *et al.*, 1989; Cohn *et al.*, 1992; Hirsch *et al.*, 2010; Hirsch, 2014). In recognition of the need for additional future work, several limitations of this study need to be acknowledged. First, this study has been limited to only six constituents for nine watersheds within a specific geographical location. Model performance may vary for constituents not covered here. In addition, more studies need to be carried in different types of watershed in other geographical regions to place these results into a more generalizable context to inform further improvement of estimation methods. Second, all modified models contain only one additional flow variable for the sake of model parsimony. It is possible that a mixture of these terms can better incorporate export dynamics at different temporal scales and thus provide additional improvement. Third, this work examined three sampling strategies that are representative of common monitoring practice, but not strategies with different sampling frequencies (*e.g.*, weekly *vs.* monthly), record lengths (*e.g.*, 10 *vs.* 30 years), or proportions of routine versus storm sampling. Further understanding of these aspects can help inform the allocation of valuable (and often limited) resources toward river quality monitoring. Fourth, in addition to the time, discharge, and season windows

in the current WRTDS model, an additional window may be introduced for the new variable on antecedent discharge, such that higher weights are assigned to concentration samples that have similar antecedent discharge conditions to the estimation day. Fifth, antecedent discharge variables may also be considered toward use with the flow-normalization algorithm in WRTDS for water-quality trend analysis. Finally, this work has focused entirely on traditional monitoring data that derive from infrequently measured water-quality samples. With increasing use of new technology that allows very high-frequency or even continuous data collection from in-situ probes, a practical challenge facing the hydrological community is how to combine the use of both traditional low-frequency data and in-situ high-frequency data toward riverine water quality analysis and load estimation (Pellerin *et al.*, 2014; Halliday *et al.*, 2015). Moving forward, there is need for additional research to advance understandings on these aspects and to incorporate all available understanding into the development and upgrade of software that can be broadly applied.

Finally, the analysis of LRD estimation methods provides new understanding and quantification of the proposed methods' performances for estimating LRD in irregular data. Accurate quantification of LRD in irregular water-quality time series remains an unresolved challenge for the hydrologic community and for other disciplines that must grapple with irregular sampling (Kirchner and Neal, 2013; Witt and Malamud, 2013). Several areas of research deserve attention and can be potentially fruitful. First, for time series with very long gaps, an alternative and perhaps more feasible approach is to break the whole time series into several segments (each without long gaps) and then apply the wavelet method to analyze each segment separately. If this can yield more accurate

estimate, then further simulation experiments should be designed to systematically determine how long the gap needs to be to invoke such an approach. Second, the simulation experiments described herein can be used as references to coarsely evaluate the reliability of estimation methods and thus quantify the likely bias and error in their LRD estimates. Future research may consider applying Bayesian statistical analysis or other approaches to more accurately quantify the LRD strength and associated uncertainty for real water-quality data. Third, given that real water-quality data often exhibit non-Gaussian distributions and strong effects of flow-dependence and seasonality (Hirsch *et al.*, 1991; Helsel and Hirsch, 2002), future research should simulate time series with these important characteristics and then evaluate the performance of various LRD estimation approaches, perhaps following the framework described herein. Last but not least, it is well established that the presence of LRD, if not accounted for, can increase Type I errors when inferring the statistical significance of temporal trends (Cohn and Lins, 2005). In this regard, the analyses presented herein may be further extended to simulations of irregular time series that have prescribed LRD strengths and also superimposed temporal trends, which can then be used to evaluate and compare the validity of various statistical methods for identification of temporal trend and the associated statistical significance.

Overall, this dissertation research has provided new understanding of nutrient and sediment export from the watershed through retrospective analysis and synthesis of decadal-scale records. The findings are useful in the context of adaptive watershed management, tributary water-quality monitoring, and continued development and calibration of the Chesapeake Bay Watershed Model. After decades of management

efforts, the goal of Chesapeake Bay restoration has not yet been fulfilled. In this regard, continued watershed-wide data collection and integration, methods development and improvement, and coupled watershed-estuary modeling are much needed to provide additional science-based understanding of nutrient and sediment fluxes and uncertainties at the watershed scale as well as their connections to temporally and spatially changing watershed environments (*e.g.*, more extreme rainfall events due to climate change, land use change, legacy pools of nutrient and sediment, and intensifying agricultural production and associated fertilizer use) and to downstream estuary dynamics (*e.g.*, dissolved oxygen condition). The Chesapeake Bay Program partnership, which is a precious and long-standing model that engages and unites federal and state agencies, local governments, non-profit organizations, and academic institutions, needs to continue working on these grand challenges in the coming decades. More broadly, in a world with seemingly ubiquitous nutrient enrichment and water quality degradation, lessons that will be learned on the 64,000 mi<sup>2</sup> Chesapeake Bay watershed and its various tributaries can become valuable resources that may guide and facilitate the protection and restoration of water resources (*i.e.*, rivers, lakes, and estuaries) in other geographical locations.

#### **11.4. Literature Cited**

- Clarke, R. T., 2013. Calculating uncertainty in regional estimates of trend in streamflow with both serial and spatial correlations. *Water Resour. Res.* 49:7120-7125, DOI: 10.1002/wrcr.20465.
- Cohn, T. A., D. L. Caulder, E. J. Gilroy, L. D. Zynjuk and R. M. Summers, 1992. The validity of a simple statistical model for estimating fluvial constituent loads: An Empirical study involving nutrient loads entering Chesapeake Bay. *Water Resour.*

- Res.* 28:2353-2353, DOI: 10.1029/92wr01008.
- Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch and D. K. Wells, 1989. Estimating constituent loads. *Water Resour. Res.* 25:937-942, DOI: 10.1029/WR025i005p00937.
- Cohn, T. A. and H. F. Lins, 2005. Nature's style: Naturally trendy. *Geophys. Res. Lett.* 32:L23402, DOI: 10.1029/2005GL024476.
- Darken, P. F., C. E. Zipper, G. I. Holtzman and E. P. Smith, 2002. Serial correlation in water quality variables: Estimation and implications for trend analysis. *Water Resour. Res.* 38:1117, DOI: 10.1029/2001WR001065.
- Halliday, S. J., R. A. Skeffington, A. J. Wade, M. J. Bowes, E. Gozzard, J. R. Newman, M. Loewenthal, E. J. Palmer-Felgate and H. P. Jarvie, 2015. High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. *Hydrol. Process.* 10.1002/hyp.10453, DOI: 10.1002/hyp.10453.
- Helsel, D. R. and R. M. Hirsch, 2002. Statistical Methods in Water Resources. *U.S. Geological Survey Techniques of Water-Resources Investigations Book 4, Chapter A3*. U.S. Geological Survey, Reston, VA, p. 522.  
<http://pubs.usgs.gov/twri/twri4a3/>.
- Hirsch, R. M., 2012. Flux of Nitrogen, Phosphorus, and Suspended Sediment from the Susquehanna River Basin to the Chesapeake Bay during Tropical Storm Lee, September 2011, as an indicator of the effects of reservoir sedimentation on water quality. U.S. Geological Survey Scientific Investigations Report 2012-5185, Reston, VA, p. 17. <http://pubs.usgs.gov/sir/2012/5185/>.

- Hirsch, R. M., 2014. Large Biases in Regression-Based Constituent Flux Estimates: Causes and Diagnostic Tools. *J. Am. Water Resour. Assoc.* 50:1401-1424, DOI: 10.1111/jawr.12195.
- Hirsch, R. M., R. B. Alexander and R. A. Smith, 1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resour. Res.* 27:803-813, DOI: 10.1029/91WR00259.
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Keisman, J., J. Blomquist, S. Phillips, G. Shenk and E. Yagow, 2015. Estimating Land Management Effects on Water Quality Status and Trends. *Scientific and Technical Advisory Committee Publication Number 14-009*. Edgewater, Maryland, p. 28.
- Kirchner, J. W. and C. Neal, 2013. Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc. Natl. Acad. Sci. U. S. A.* 110:12213-12218, DOI: 10.1073/pnas.1304328110.
- Langland, M. J., J. D. Blomquist, D. L. Moyer and K. E. Hyer, 2012. Nutrient and suspended-sediment trends, loads, and yields and development of an indicator of streamwater quality at nontidal sites in the Chesapeake Bay watershed, 1985-2010. U.S. Geological Survey Scientific Investigations Report 2012-5093, Reston, VA, p. 26. <http://pubs.usgs.gov/sir/2012/5093/pdf/sir2012-5093.pdf>.
- Linker, L. C., R. A. Batiuk, G. W. Shenk and C. F. Cerco, 2013. Development of the Chesapeake Bay Watershed Total Maximum Daily Load Allocation. *J. Am. Water*

- Resour. Assoc.* 49:986-1006, DOI: 10.1111/jawr.12105.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Pellerin, B. A., B. A. Bergamaschi, R. J. Gilliom, C. G. Crawford, J. Saraceno, C. P. Frederick, B. D. Downing and J. C. Murphy, 2014. High frequency measurement of nitrate concentration in the Lower Mississippi River, USA. *Environ. Sci. Technol.* 48:12612-12619, DOI: 10.1021/es504029c.
- Shenk, G. W. and L. C. Linker, 2013. Development and Application of the 2010 Chesapeake Bay Watershed Total Maximum Daily Load Model. *J. Am. Water Resour. Assoc.* 49:1042-1056, DOI: 10.1111/jawr.12109.
- Witt, A. and B. D. Malamud, 2013. Quantification of Long-Range Persistence in Geophysical Time Series: Conventional and Benchmark-Based Improvement Techniques. *Surv. Geophys.* 34:541-651, DOI: 10.1007/s10712-012-9217-8.
- Yue, S., P. Pilon, B. Phinney and G. Cavadias, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Process.* 16:1807-1829, DOI: 10.1002/hyp.1095.
- Zhang, Q. and W. P. Ball, 2016. Data associated with Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-Quality Data. Johns Hopkins University Data Archive, Baltimore, MD.  
<http://dx.doi.org/10.7281/T18G8HM0>, DOI: 10.7281/T18G8HM0.

**Appendix A. Supporting Information to Chapter 2**

**“Long-term Seasonal Trends of Nitrogen, Phosphorus, and Suspended Sediment  
Load from the Non-tidal Susquehanna River Basin to Chesapeake Bay”**

## Appendix A1. Overview of the estimation process in the WRTDS method

This section briefly summarizes the basic structure of the estimation process in the “weighted regressions on time, discharge, and season (WRTDS)” method, for both *true-condition* and *flow-normalized* estimates. The WRTDS method has also been fully described elsewhere, *e.g.*, Hirsch *et al.* (2010), and Sprague *et al.* (2011).

As described by Hirsch *et al.* (2010), the “core” equation of WRTDS is,

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (\text{A1})$$

where  $C$  is concentration of analyte,  $Q$  is daily streamflow discharge,  $t$  is time in decimal years,  $\beta_i$  are fitted coefficients, and  $\varepsilon$  is the error term. For convenience, we can classify any day in the dataset as either a “Sampled Day” or an “Unsampled Day,” depending on whether it has measured concentration for the analyte of interest. Note that measured discharge is available every day. To estimate the “*true-condition*” concentration ( $C_{tc}$ ) on a particular day (called “Estimation Day”) with known parameters of time ( $t_0$ ) and discharge ( $Q_0$ ), the first step in WRTDS is to select at least 100 Sampled Days (*i.e.*, days with both measured discharge and concentration) that are sufficiently close in time, discharge, and season in relative to the Estimation Day so as to ultimately have a non-zero weight (Figure A1) (Hirsch *et al.*, 2010). In general, for a Sampled Day with time parameter ( $t_i$ ), measured discharge ( $Q_i$ ), and measured concentration ( $C_i$ ), weight is assigned by first calculating the time distance, discharge distance, and seasonal distance between that Sampled Day and the Estimation Day (Hirsch *et al.*, 2010). According to a previously developed tri-cube weight function (Tukey, 1977) and pre-defined half-window widths, the three distances are converted to time weight ( $w_t$ ), discharge weight ( $w_d$ ), and seasonal weight ( $w_s$ ), respectively. The weight assigned to that Sampled Day ( $w_i$ ) is then computed as product of the three individual weights (Hirsch *et al.*, 2010),

$$w_i = w_t \cdot w_d \cdot w_s \quad (\text{A2})$$

Mathematically, the method assigns relatively little weight to any Sampled Day that is far away from the Estimation Day in one or more of the three dimensions (*i.e.*, time, discharge, or season). In this study, the half-window widths were defined as 10 years and 0.5 years for the time and seasonal dimensions, respectively. For the discharge dimension, the window was selected such that, for a given discharge  $Q$  (as reference),

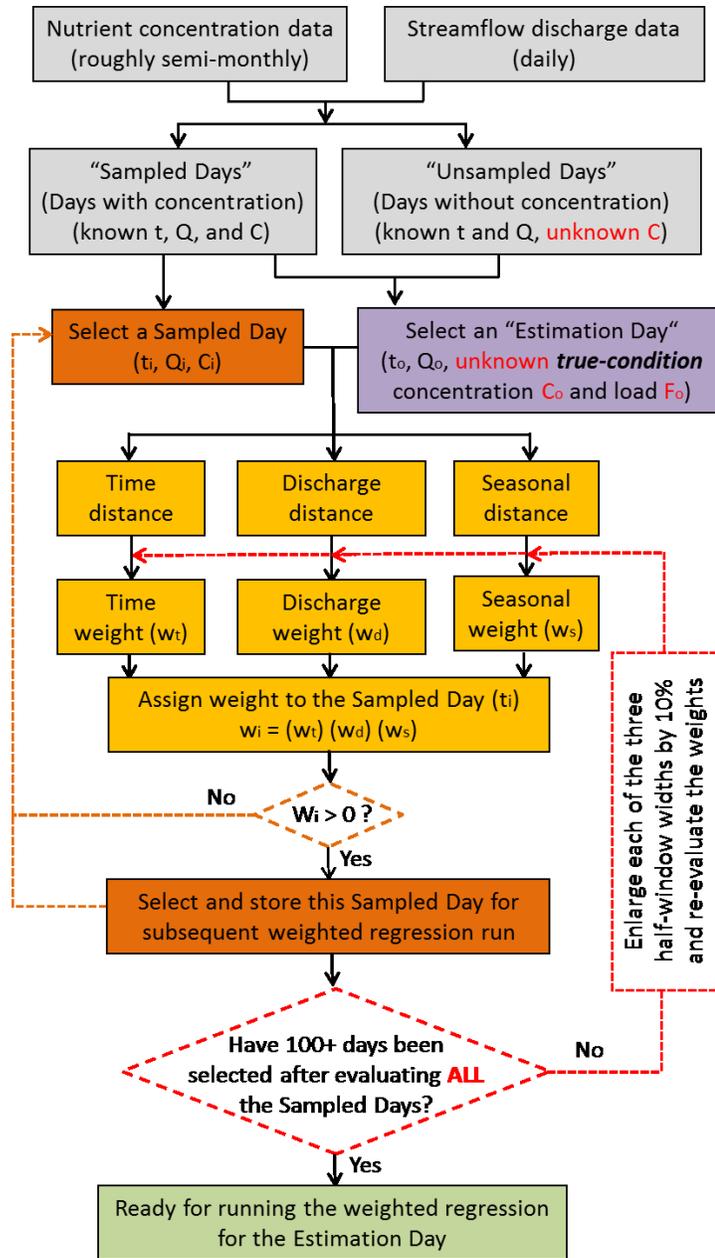


Figure A1. Flowchart illustrating the first step of the WRTDS method – selection of data to be used in the weighted regression. The objective is to find at least 100 “Sampled Days” that are sufficiently “proximate” to the “Estimation Day” for which the *true-condition* concentration ( $C_{tc}$ ) is to be estimated. For a Sampled Day with parameters of time ( $t_i$ ), measured discharge ( $Q_i$ ), and measured concentration ( $C_i$ ), the proximity to the “Estimation Day” (with known parameters  $t_o$  and  $Q_o$ ) is calculated for each of the three dimensions -- *i.e.*, time, discharge, and season.

non-zero weights would be assigned only to discharges falling between  $Q_{\min} = Q/\exp(2)$  and  $Q_{\max} = Q \exp(2)$ , or in other words,  $\ln(Q_{\max}/Q) = \ln(Q/Q_{\min}) = 2$ . Hirsch *et al.* (2010) have suggested the proper range for each half-window width, and considered the above values appropriate for Chesapeake Bay tributaries based on testing.

The above-described weight assignment is performed for all Sampled Days in the record. If less than 100 Sampled Days are assigned non-zero weight, the algorithm would automatically enlarge each of the three half-window widths by 10% to allow more Sampled Days to have non-zero weight (Hirsch *et al.*, 2010). This process is repeated, if necessary, until 100 or more Sampled Days are assigned non-zero weight. These selected Sampled Days are then compiled to run a weighted regression using Eq. (A.1) to determine the fitted coefficients ( $\beta_i$ ) (Figure A2a) (Hirsch *et al.*, 2010). The *true-condition* concentration on the Estimation Day ( $C_{tc}$ ) is then obtained by substituting  $t_0$  and  $Q_0$  into Eq. (A.1), and the *true-condition* load ( $F_{tc}$ ) is calculated as follows,

$$F_{tc} = C_{tc} \cdot Q_0 \quad (\text{A3})$$

Both  $C_{tc}$  and  $F_{tc}$  are referred as the *true-condition* estimates because the streamflow discharge (*i.e.*,  $Q_0$ ) used in the estimation process represents the true hydrological condition on the Estimation Day (Hirsch *et al.*, 2010).

The above process is separately implemented for all days in the record to obtain a complete time series of daily *true-condition* estimates of concentration and load (Figure A2a). Note that when a Sampled Day is considered as the Estimation Day, the above process is implemented by assuming that the concentration is unknown. Thus, the estimated concentration on any Sampled Day would not be necessarily the same as the measured concentration. In WRTDS, a unique set of fitted coefficients in Eq. (A1) is obtained separately for each Estimation Day (Hirsch *et al.*, 2010). This key feature allows the dependencies of concentration on time, discharge, and season to change flexibly.

Another feature of the WRTDS method is a flow-normalization algorithm, which estimates the target daily concentration and load using the full set of discharges occurring on that day of all years in the entire record, while holding the time and seasonal components constant (Figure A2b) (Hirsch *et al.*, 2010). To better illustrate this algorithm, we consider the task of estimating the *flow-normalized* daily concentration

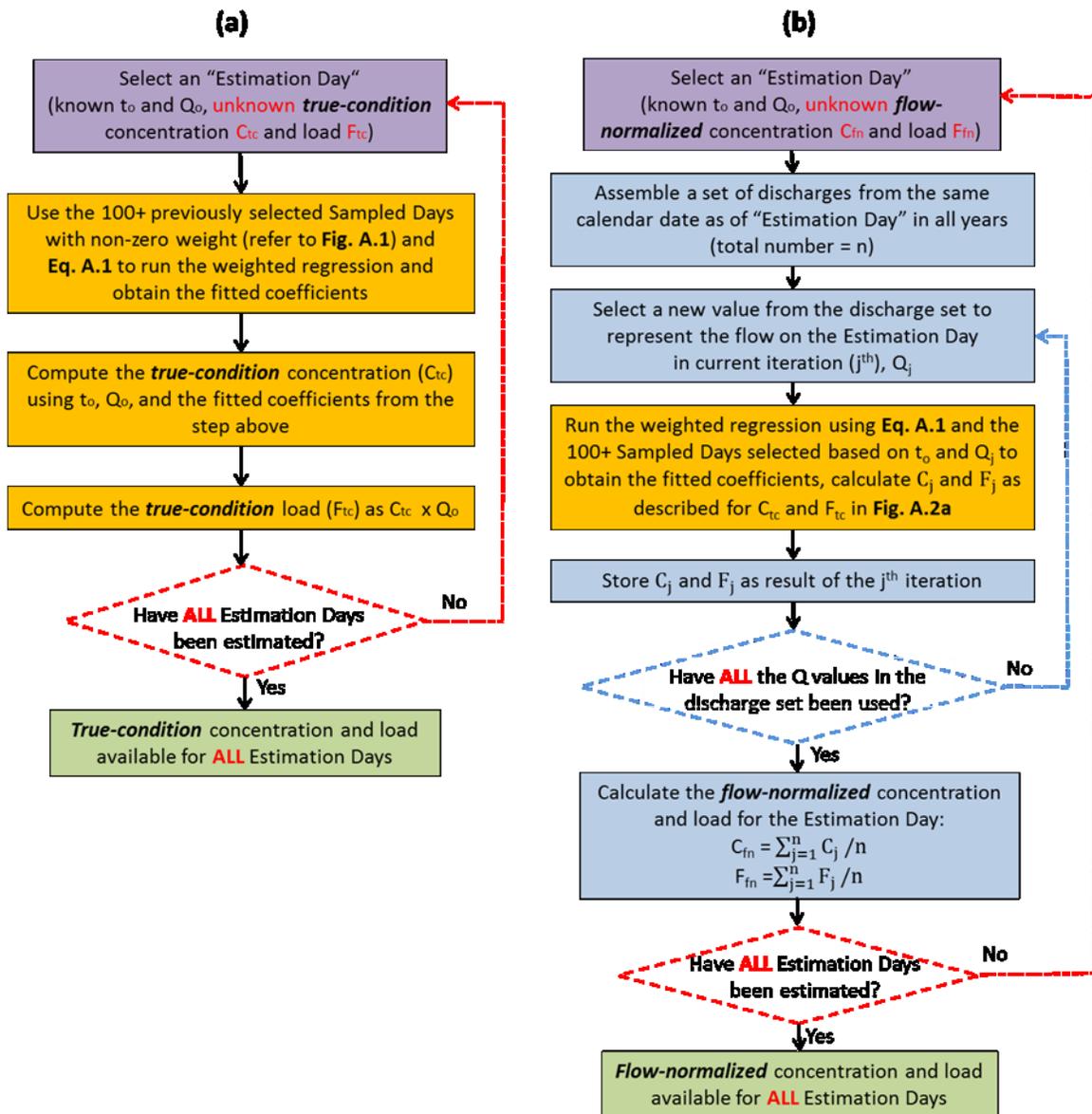


Figure A2. Flow diagrams illustrating the WRTDS method for calculating (a) the “*true-condition*” estimates of concentration and load for all “Estimation Days”, and (b) the “*flow-normalized*” estimates of concentration and load for all “Estimation Day.”

and load on 1 January 1978 (the “Estimation Day” in this example) for total nitrogen (TN) in the Susquehanna River at Conowingo. The available data include continuous daily discharge record and intermittent TN concentration record at Conowingo from 1 January 1978 to 31 December 2011 (34 years). The main idea of flow-normalization is to link the estimation on 1 January 1978 to the full range of hydrological conditions for the

calendar date of the Estimation Day, *i.e.*, all the historical discharges occurring on January 1 (Figure A3a). Let  $j$  be an index of a certain year in the entire record (*e.g.*,  $j = 1$  for 1978,  $j = 2$  for 1979, ...,  $j = 34$  for 2011), thus, for the date of January 1,  $Q_1 = Q_{1/1/1978}$ ,  $Q_2 = Q_{1/1/1979}$ , ...,  $Q_{34} = Q_{1/1/2011}$ . In the first iteration,  $Q_1$  is assigned as the discharge occurring on the Estimation Day, and the estimated concentration ( $C_1$ ) is obtained using the weighted regression procedure described above. The estimated load ( $F_1$ ) is then calculated as follows:

$$F_1 = Q_1 \cdot C_1 \quad (\text{A4})$$

Because  $Q_1$  is the true flow condition on 1 January 1978, the estimates ( $C_1, F_1$ ) are indeed the *true-condition* estimates for the Estimation Day. In the subsequent iterations,  $Q_j$  ( $2 \leq j \leq 34$ ) is set to be the discharge, and the estimates ( $C_j, F_j$ ) are obtained using the same procedure as in the first iteration. Correspondingly, a total of 34 pairs of concentration and load estimates are obtained. Finally, by assigning equal probability to the 34 discharges occurring on January 1, the *flow-normalized* estimates of concentration ( $C_{fn}$ ) and load ( $F_{fn}$ ) for the Estimation Day are calculated as the average of the estimates obtained in the 34 iterations:

$$C_{fn} = \frac{1}{34} \sum_{j=1}^{34} C_j \quad (\text{A5})$$

$$F_{fn} = \frac{1}{34} \sum_{j=1}^{34} F_j \quad (\text{A6})$$

Unlike the *true-condition* estimates (Eq. [A.3]), there is no apparent relation between the *flow-normalized* concentration and the *flow-normalized* load. Similarly, the above process is separately implemented for all the other days in the record to obtain a complete time series of daily *flow-normalized* estimates of concentration and load.

The *true-condition* and the *flow-normalized* TN estimates at Conowingo for 1 January in each year of the record (1978-2011) are presented in Figure A3 (b and c). By responding to the full hydrological cycle, the *flow-normalized* estimates can largely remove the sometimes dramatic influence of random variations in streamflow, and can thus render inter-annual trends easier to detect and understand than they would be with the *true-condition* estimates. For this example, the *true-condition* daily TN loads were lower than the total maximum daily load (TMDL) of 101,000 kg N day<sup>-1</sup> set for the

Susquehanna River (US Environmental Protection Agency, 2010) in some years (Figure A3c), which may be attributed to random low-flow conditions, or effective management actions, or a combination of both. In contrast, the *flow-normalized* loads were always higher than the TMDL, suggesting that the nitrogen reduction goal has not yet been met.

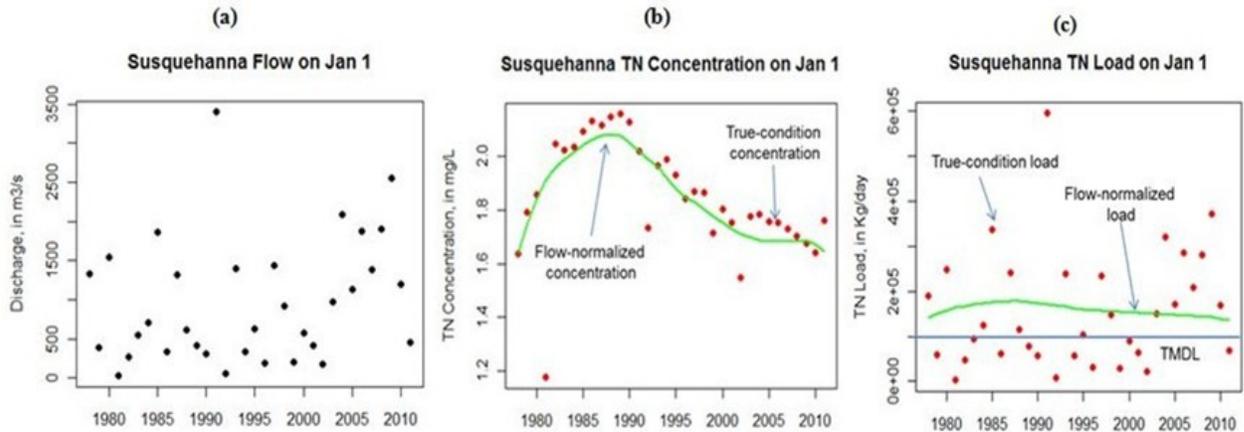


Figure A3. Three plots illustrating application of WRTDS to estimate concentration and load for total nitrogen (TN) in the Susquehanna River at Conowingo on January 1 in each year of the record (1978-2011). Part (a) shows all historical discharges occurring on the calendar date of the Estimation Day (*i.e.*, January 1). Using the flow record, WRTDS produces both the *true-condition* and the *flow-normalized* estimates of (b) daily TN concentration and (c) daily TN load on January 1 in each year (refer to Figure A2). For comparison, the TMDL of 101,000 kg N day<sup>-1</sup> set for the Susquehanna River (US Environmental Protection Agency, 2010) is inserted in (c).

## Appendix A2. Linear regression models relating monthly TN to NO<sub>x</sub> loads at the Conowingo Station

This section presents the linear regression models relating monthly TN load to nitrate plus nitrite (NO<sub>x</sub>) load at Conowingo for each calendar month (Table A1). The load estimates were obtained using the WRTDS method based on available monitoring data from 1981 to 2010 (30 years). The slope coefficient and the model-fit parameter (R<sup>2</sup>) were statistically significant in all twelve models.

Table A1. Regression models relating monthly TN to NO<sub>x</sub> loads at the Conowingo Station. <sup>a</sup>

Month	TN = $\beta_1$ (NO <sub>x</sub> ) + $\beta_2$		
	Slope, $\beta_1$ (p-value)	Intercept, $\beta_2$ <sup>b</sup> (p-value)	Model fit, R <sup>2</sup> (p-value)
January	1.3206 (< 2 x10 <sup>-16</sup> )	0	0.9965 (< 2.2 x10 <sup>-16</sup> )
February	1.3691 (< 2 x10 <sup>-16</sup> )	0	0.9848 (< 2.2 x10 <sup>-16</sup> )
March	1.3822 (< 2 x10 <sup>-16</sup> )	0	0.9958 (< 2.2 x10 <sup>-16</sup> )
April	1.4588 (< 2 x10 <sup>-16</sup> )	0	0.9961 (< 2.2 x10 <sup>-16</sup> )
May	1.4909 (< 2 x10 <sup>-16</sup> )	0	0.9982 (< 2.2 x10 <sup>-16</sup> )
June	1.5814 (< 2 x10 <sup>-16</sup> )	0	0.9915 (< 2.2 x10 <sup>-16</sup> )
July	1.5655 (< 2 x10 <sup>-16</sup> )	0	0.9986 (< 2.2 x10 <sup>-16</sup> )
August	1.4884 (< 2 x10 <sup>-16</sup> )	975.3 (0.075)	0.9986 (< 2.2 x10 <sup>-16</sup> )
September	1.8776 (< 2 x10 <sup>-16</sup> )	0	0.9647 (< 2.2 x10 <sup>-16</sup> )
October	1.3879 (< 2 x10 <sup>-16</sup> )	0	0.9987 (< 2.2 x10 <sup>-16</sup> )
November	1.3417 (< 2 x10 <sup>-16</sup> )	0	0.9981 (< 2.2 x10 <sup>-16</sup> )
December	1.3407 (< 2 x10 <sup>-16</sup> )	0	0.9956 (< 2.2 x10 <sup>-16</sup> )

<sup>a</sup> developed using monthly WRTDS-estimates of TN and NO<sub>x</sub> loads from 1981 to 2010.

<sup>b</sup> forced to be zero, if negative.

### **Appendix A3. Overview of some historical changes in Susquehanna watershed management and practices**

To provide context for our discussion of loading changes in the main manuscript, this section presents an overview of some relevant historical changes in management and practices associated with point, agricultural, and stormwater sources in the Susquehanna River basin (SRB), with more focus on Pennsylvania (PA) and New York (NY) than Maryland (MD) due to their dominance in watershed area.

The most important historical management controls at point sources in the SRB are probably the P-ban in detergents and the introduction of increasingly effective nutrient removal technologies (NRTs) at wastewater treatment plants (WWTPs). Among the three jurisdictions in the SRB, NY led the implementation of state-wide P-ban in detergents (0.5% weight limit), which became effective in June 1973 for household products and in January 1976 for commercial products (Litke, 1999). MD and PA introduced the state-wide P-ban in detergents in December 1985 and March 1990, respectively (Litke, 1999). NRT upgrades in WWTP include the general adoption of biological nutrient removal (BNR) practices to achieve higher removal efficiencies for both N and P. As an early attempt, PA adopted a point source nutrient control program for the lower Susquehanna River in 1970 that required 80% P-removal for all “new and modified dischargers” to the river and its tributaries below the Juniata River, and later extended it to a state-wide nutrient regulation in 1983 (Chesapeake Executive Council, 1988). By 1998, PA has implemented BNR in 11 significant point source facilities in the SRB, which greatly offset nutrient load rise due to population expansion (Chesapeake Bay Program, 1998). In 2004, PA allocated \$250 million to further support NRT installation (Pennsylvania Department of Environmental Protection, 2004). Further upstream, NY was expected to complete BNR facility construction at the Binghamton-Johnson City WWTP, the largest point source of NY in the SRB, by 2001 (Blankenship, 1999). Later NY installed an enhanced level of treatment (so-called “enhanced nutrient removal” [ENR]) in this WWTP in July 2007 (New York State Department of Environmental Conservation, 2007). Overall, these NRT upgrades, together with the P-ban in detergents, have likely been responsible for a significant part of observed reductions in point-source N and point-source P in the SRB. In fact, the overall reduction of point-source P has been

substantial, for example, Sprague *et al.* (2000) reported a 41% decline of total phosphorus (TP) load from point sources in the SRB from 1985 to 1998.

For agricultural sources, many management strategies have focused on control of fertilizer and manure on the agricultural land. As early as 1977, PA first adopted regulations on storage and usage of animal manure (Pennsylvania Department of Environmental Protection, 2004). In September 1987, the Chesapeake Bay Program provided \$1.2 million to fund best management practices (BMPs) in PA (Chesapeake Executive Council, 1988). Also in 1987, the PA Environmental Quality Board reviewed waste management regulations and published them in April 1988 (Chesapeake Executive Council, 1988). PA's Nutrient Management Act of 1993, which included regulations on concentrated animal operations (CAOs), became fully effective in 1997 (Pennsylvania Department of Environmental Protection, 2004). The regulations on concentrated animal feeding operations (CAFOs) were further updated in 2002, and farms raising non-production animal began to be considered for CAOs in 2004 (Pennsylvania Department of Environmental Protection, 2004). Further upstream, NY initiated the Agricultural Environmental Management Program in 2000 with some specific CAFO policies (New York State Department of Environmental Conservation, 2007). In 2004, the NY State Soil and Water Conservation Committee urged all CAFOs to develop comprehensive nutrient management plans (New York State Department of Environmental Conservation, 2007). In addition, another related regulation, the Unified Strategy for Animal Feeding Operations, was jointly issued by the U.S. Department of Agriculture (USDA) and the U.S. Environmental Protection Agency (USEPA) in 1999 (US Department of Agriculture and US Environmental Protection Agency, 1999). This regulation listed some P-based nutrient management strategies including the P-index approach (Weld *et al.*, 2002) that had been previously developed by USDA in cooperation with the scientific community in 1993 (Lemunyon and Gilbert, 1993; Weld *et al.*, 2002). The P-index approach evaluates P-loss risk by considering both source and transport factors (Sharpley *et al.*, 2003). It has since been adopted in some areas in the SRB, including the southeast PA, which has high animal densities and was reported to be most influenced by this approach (Kogelmann *et al.*, 2004).

For stormwater management, the USEPA has initiated the National Pollutant

Discharge Elimination System Phase I regulations in 1990 for Municipalities with Separate Storm Sewer Systems (MS4s) serving populations of 0.1 million or more (US Environmental Protection Agency, 2000). In 1999, the program was expanded to smaller MS4s with the publication of Phase II Final Rule in 1999 (US Environmental Protection Agency, 2000). In PA, the Comprehensive Stormwater Management Policy became effective in September 2002 (Pennsylvania Department of Environmental Protection, 2004). In NY, under the State Pollutant Discharge Elimination System, permit controls were initiated for stormwater from industrial and construction activities in 1998 and 2002, respectively, and for MS4 management in 2002 (New York State Department of Environmental Conservation, 2007).

Among the reported trends, an interesting observation is the sharp reduction of DOP load above the reservoirs since around 2002 (Figure 2.7d). This trend is likely related to three concurrent changes in the Pennsylvania watershed, namely, the initiation of regulations on concentrated animal operations (CAOs) in 1997, the updating of regulations on concentrated animal feeding operations (CAFOs) in 2002, and the introduction of P-index to nutrient management in the early 2000s. Overall, it is likely that the source-based management strategies and associated controls on transport and transformation processes, were responsible for the downward nutrient and sediment trends in the SRB at locations above the reservoirs.

## Literature Cited

- Blankenship, K., 1999. NY sewage treatment plant to install BNR, boost Bay efforts. Chesapeake Bay Journal (March 1999).
- Chesapeake Bay Program, 1998. Chesapeake Bay watershed model application and calculation of nutrient and sediment loadings, Appendix F: Point source loadings. Annapolis, MD, p. 693.  
[http://www.chesapeakebay.net/content/publications/cbp\\_12313.pdf](http://www.chesapeakebay.net/content/publications/cbp_12313.pdf).
- Chesapeake Executive Council, 1988. Baywide nutrient reduction strategy: an agreement commitment report. Annapolis, MD.  
[http://www.chesapeakebay.net/content/publications/cbp\\_12114.pdf](http://www.chesapeakebay.net/content/publications/cbp_12114.pdf).
- Hirsch, R. M., D. L. Moyer and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46:857-880, DOI: 10.1111/j.1752-1688.2010.00482.x.
- Kogelmann, W. J., H. S. Lin, R. B. Bryant, D. B. Beegle, A. M. Wolf and G. W. Petersen, 2004. A statewide assessment of the impacts of phosphorus-index implementation in Pennsylvania. *J. Soil Water Conserv.* 59:9-18,  
<http://www.jswconline.org/content/59/1/9.short>.
- Lemunyon, L. J. and G. R. Gilbert, 1993. The concept and need for a phosphorus assessment tool. *J. Prod. Agric.* 6:483-486,  
<http://www.mendeley.com/research/concept-need-phosphorus-assessment-tool/>.
- Litke, D. W., 1999. Review of phosphorus control measures in the United States and their effects on water quality. *Water-Resources Investigations Report 99-4007*. US Geological Survey, Denver, CO, p. 43. <http://pubs.usgs.gov/wri/wri994007/>.
- New York State Department of Environmental Conservation, 2007. New York State Tributary Strategy for Chesapeake Bay Restoration.  
[http://www.dec.ny.gov/docs/water\\_pdf/cbaystratfinal.pdf](http://www.dec.ny.gov/docs/water_pdf/cbaystratfinal.pdf).
- Pennsylvania Department of Environmental Protection, 2004. Pennsylvania's Chesapeake Bay Tributary Strategy. <http://www.elibrary.dep.state.pa.us/dsweb/Get/Version-45267/3900-BK-DEP1656.pdf>.
- Sharpley, A. N., T. Daniel, T. Sims, J. Lemunyon, R. Stevens and R. Parry, 2003.

- Agricultural phosphorus and eutrophication, 2<sup>nd</sup> ed. *Agricultural Research Service, ARS-149*. US Department of Agriculture, p. 44.  
<http://ddr.nal.usda.gov/dspace/bitstream/10113/26693/1/CAT30907360.pdf>.
- Sprague, L. A., R. M. Hirsch and B. T. Aulenbach, 2011. Nitrate in the Mississippi River and its tributaries, 1980 to 2008: are we making progress? *Environ. Sci. Technol.* 45:7209-7216, DOI: 10.1021/es201221s.
- Sprague, L. A., M. J. Langland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. *Water-Resources Investigations Report 00-4218*. US Geological Survey, Richmond, VA, p. 109.  
[http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).
- Tukey, J. W., 1977. *Exploratory data analysis*. Reading, MA, Addison-Wesley, ISBN 0201076160
- US Department of Agriculture and US Environmental Protection Agency, 1999. Unified National Strategy for Animal Feeding Operations.  
<http://www.epa.gov/npdes/pubs/finafost.pdf>.
- US Environmental Protection Agency, 2000. Fact Sheet 1.0 - Stormwater Phase II Final Rule: An Overview. *Stormwater Phase II Final Rule Fact Sheet Series: EPA 833-F-00-001*.
- US Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment.  
<http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html>.
- Weld, J. L., R. L. Parsons, D. B. Beegle, A. N. Sharpley, W. J. Gburek and W. R. Clouser, 2002. Evaluation of phosphorus-based nutrient management strategies in Pennsylvania. *J. Soil Water Conserv.* 57:448-454,  
<http://www.jsowconline.org/content/57/6/448.short>.

*Page intentionally left blank*

**Appendix B. Supporting Information to Chapter 3**

**“Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to  
Chesapeake Bay: Effects of Reservoir Sedimentation”**

## Appendix B1. Supporting information to Section 3.2.1

This appendix contains Table B1 that is relevant to Section 3.2.1.

Table B1. Details of the river monitoring sites.<sup>23</sup>

USGS ID	Station name	Drainage area (km <sup>2</sup> )	Annual flow <sup>a</sup>							
			Min	Median	Mean	Max	Min	Median	Mean	Max
			(m <sup>3</sup> /s)				(m/year)			
01576000	Susquehanna River at Marietta, PA	67,314	690	1,021	1,122	2,081	0.32	0.48	0.53	0.98
01576754	Conestoga River at Conestoga, PA	1,217	10	18	20	33	0.27	0.48	0.51	0.85
01578310	Susquehanna River near Conowingo, MD	70,189	667	1,006	1,152	2,041	0.30	0.45	0.52	0.92

<sup>a</sup> calculated based on annual average flow data between 1987 and 2013 (U.S. Geological Survey, 2014).

<sup>23</sup> Relative to our published paper (Zhang *et al.*, 2016), values in the last three columns of this table have been corrected.

## **Appendix B2. Supporting information to Section 3.2.2**

This appendix contains figures that are relevant to Section 3.2.2.

### **B2-I. Normal probability plots of WRTDS model residuals**

Following procedures described by Qian,<sup>2</sup> we have constructed normal probability plots of the model residuals (in natural log units) at each of the three study sites: Susquehanna River at Conowingo, MD (Figure B1), Susquehanna River at Marietta, PA (Figure B2), and Conestoga River at Conestoga, PA (Figure B3). This analysis shows that the model residuals in natural log units are approximately normal, with only limited exceptions at the extremes of the distributions.

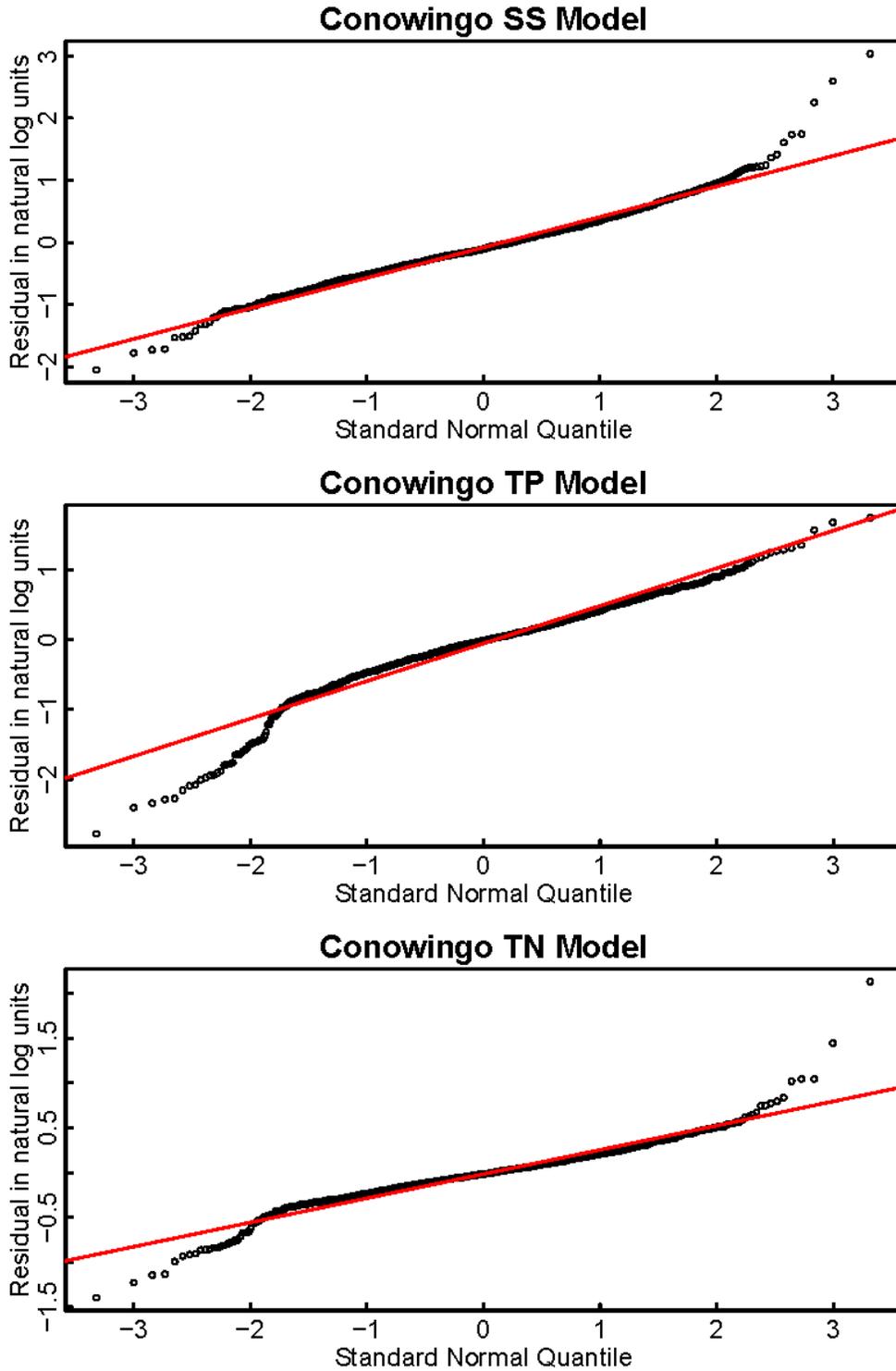


Figure B1. Normal probability plots for residuals (in natural log units) from the WRTDS models for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the Susquehanna River at Conowingo, MD.

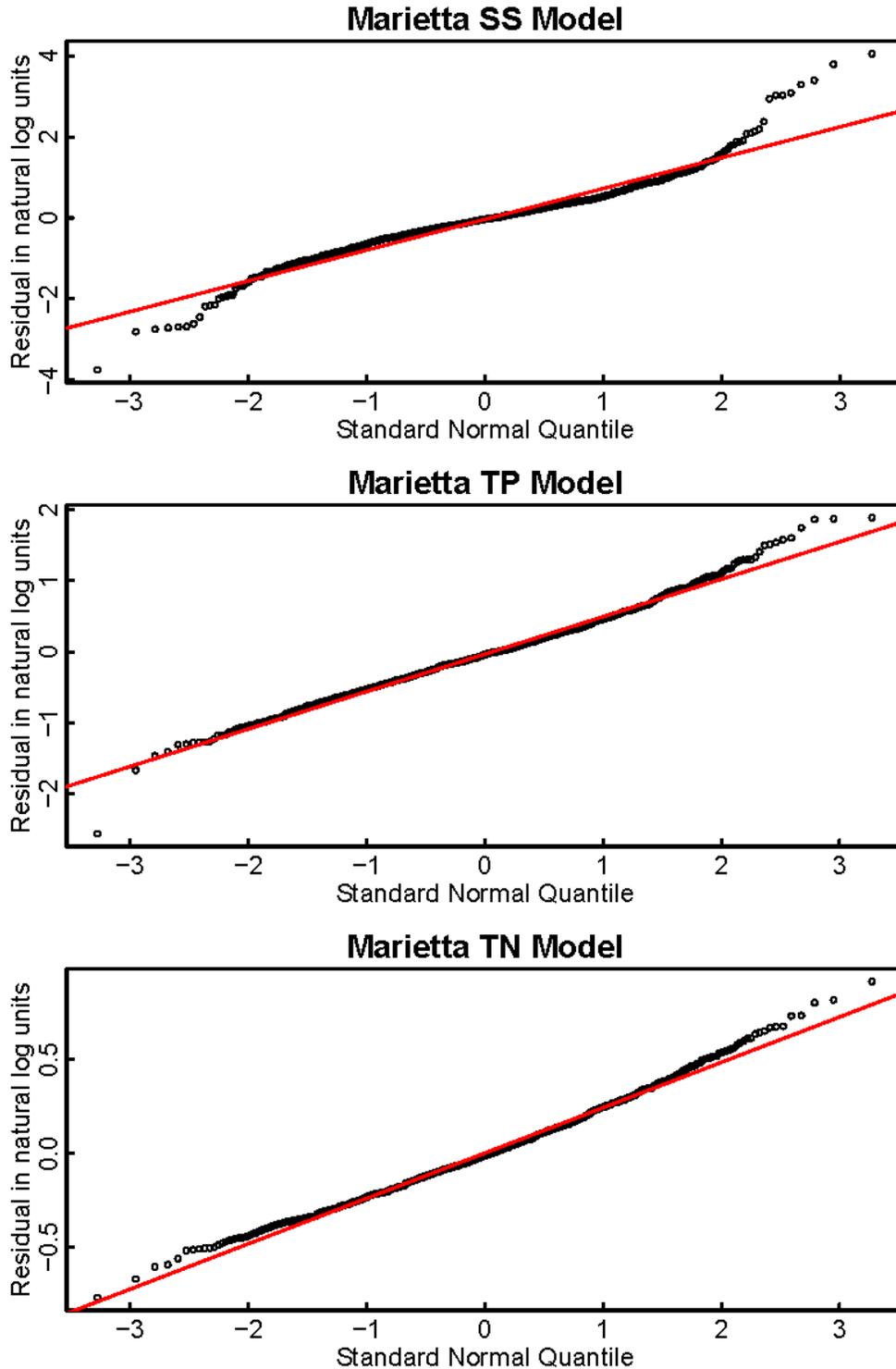


Figure B2. Normal probability plots for residuals (in natural log units) from the WRTDS models for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the Susquehanna River at Marietta, PA.

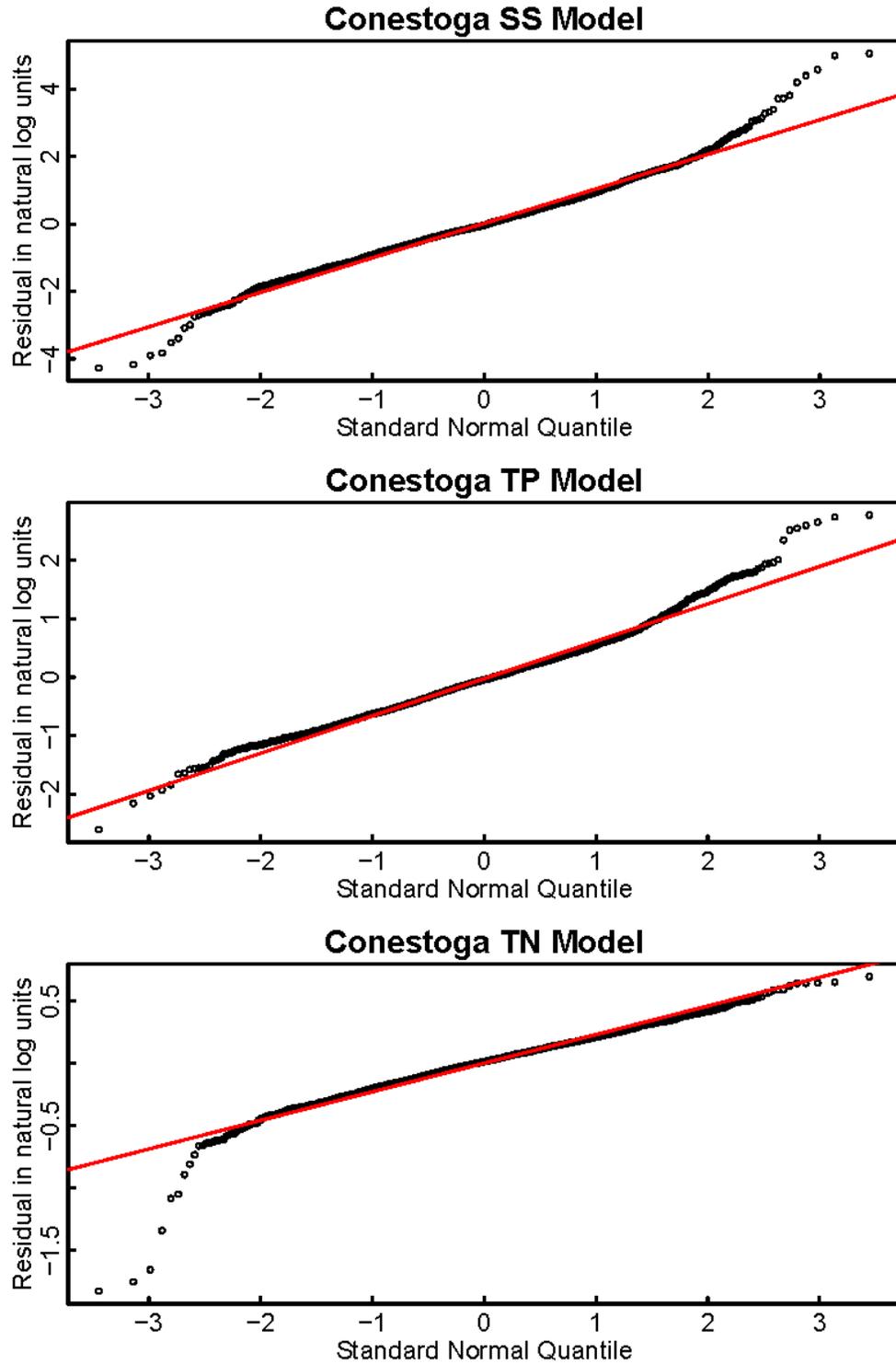


Figure B3. Normal probability plots for residuals (in natural log units) from the WRTDS models for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the Conestoga River at Conestoga, PA.

## **B2-II. WRTDS regression surface plots and model residual plots**

Figure B4 shows examples of WRTDS-derived concentration regression surface for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) in the Susquehanna River at Conowingo, MD.

Figures B5-B7 are model residual plots for SS for the three study sites, which are selected as representatives of all WRTDS analyses conducted in this study. Each figure contains four graphs that plot model residuals (observed log concentration minus estimated log concentration) against (a) estimated concentrations, (b) time, (c) discharges, and (d) months, respectively. Note that on some plots (*e.g.*, Figure B7a) the vertical stripes correspond to left-censored concentration samples. As a general rule of WRTDS, such concentrations are treated as intervals and plotted with vertical lines. As a result, their corresponding estimates and residuals are also indicated with vertical lines graphically. (Readers are referred to the “Censored Data” section in Moyer *et al.*<sup>3</sup> for details.) Another issue evident in some figures (*e.g.*, Figure B6a) is a clear pattern of diagonal line-up of some data points. This phenomenon relates to rounding issues of USGS sample recording and to the fact that the details of “round-up” varied over time.

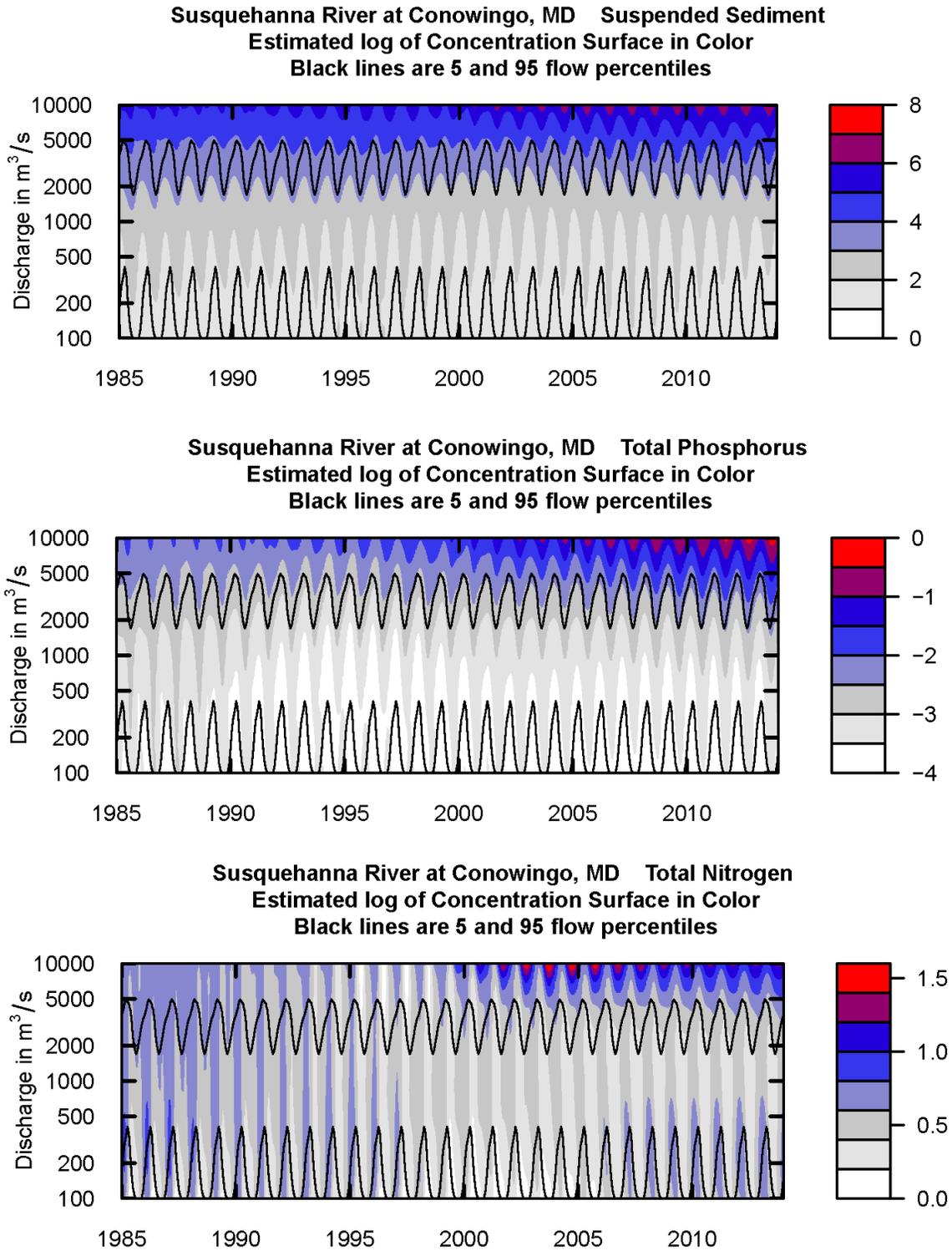


Figure B4. Contour plots showing estimated log of concentration surface from the standard WRTDS models of suspended sediment (top), total phosphorus (middle), and total nitrogen (bottom) in the Susquehanna River at Conowingo, MD.

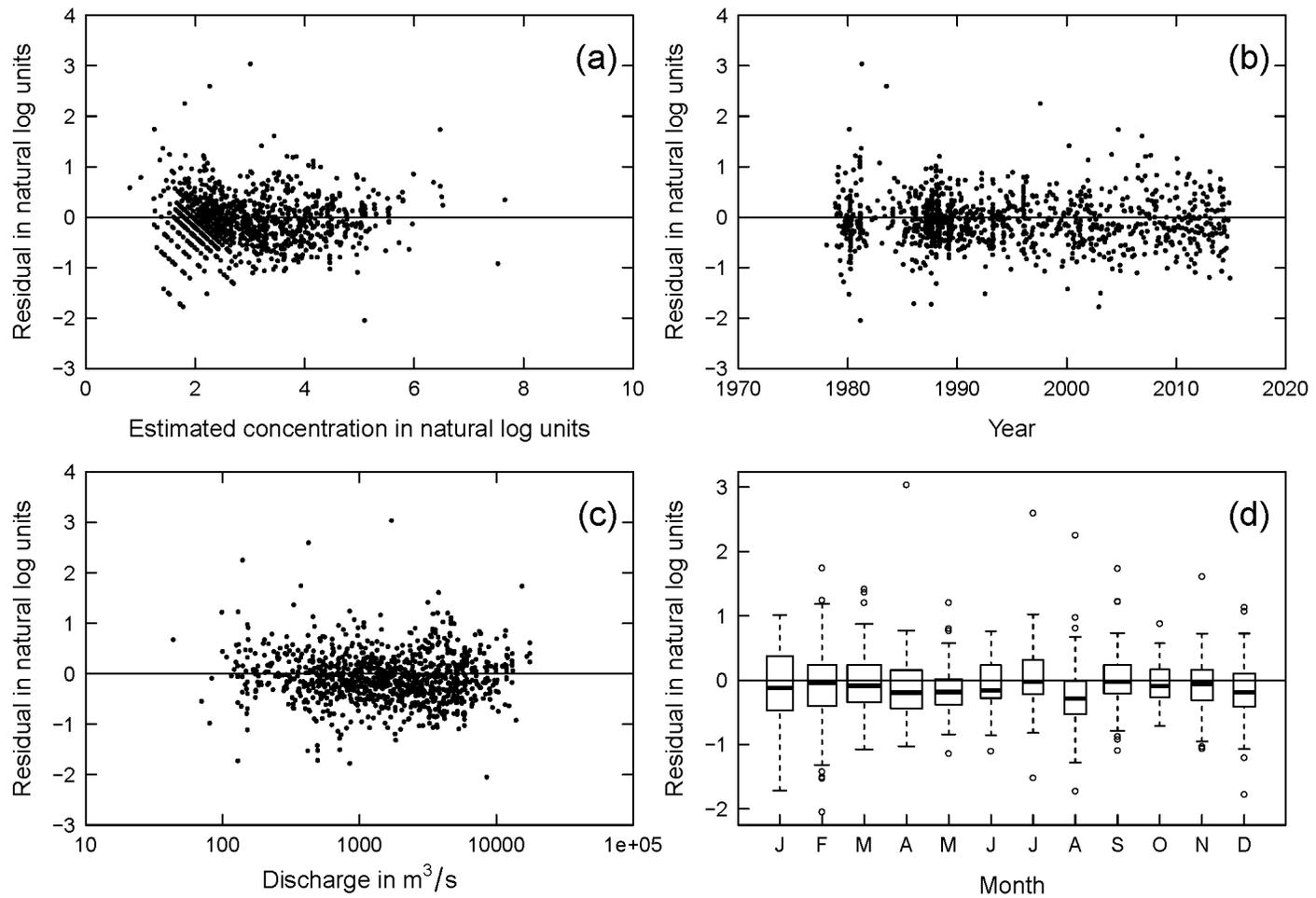


Figure B5. Model residual plot for suspended sediment in the *Susquehanna River at Conowingo, MD*, showing residual (in natural log units) as a function of (a) estimated concentration, (b) year, (c) discharge, and (d) month.

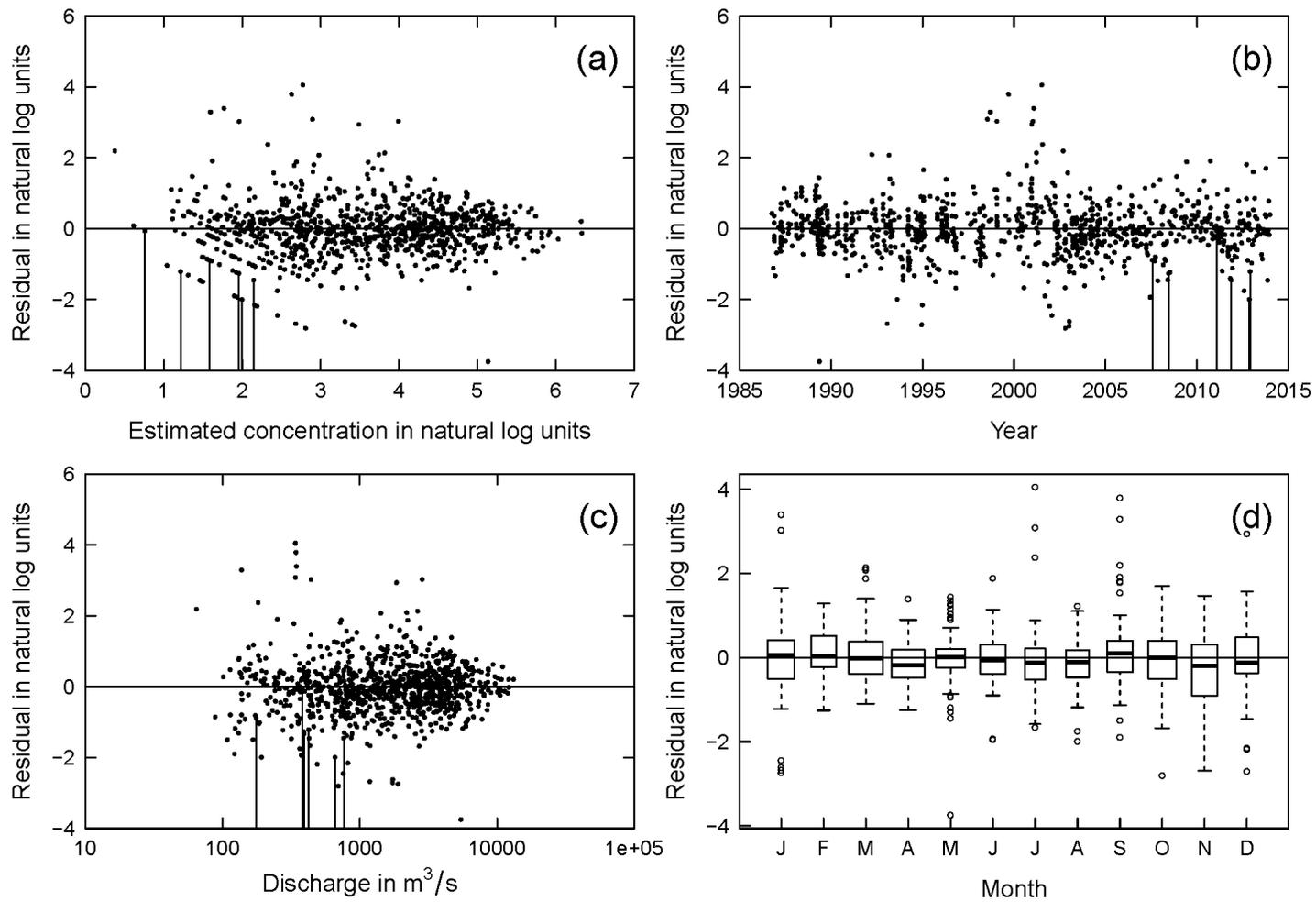


Figure B6. Model residual plot for suspended sediment in the *Susquehanna River at Marietta, PA*, showing residual (in natural log units) as a function of (a) estimated concentration, (b) year, (c) discharge, and (d) month.

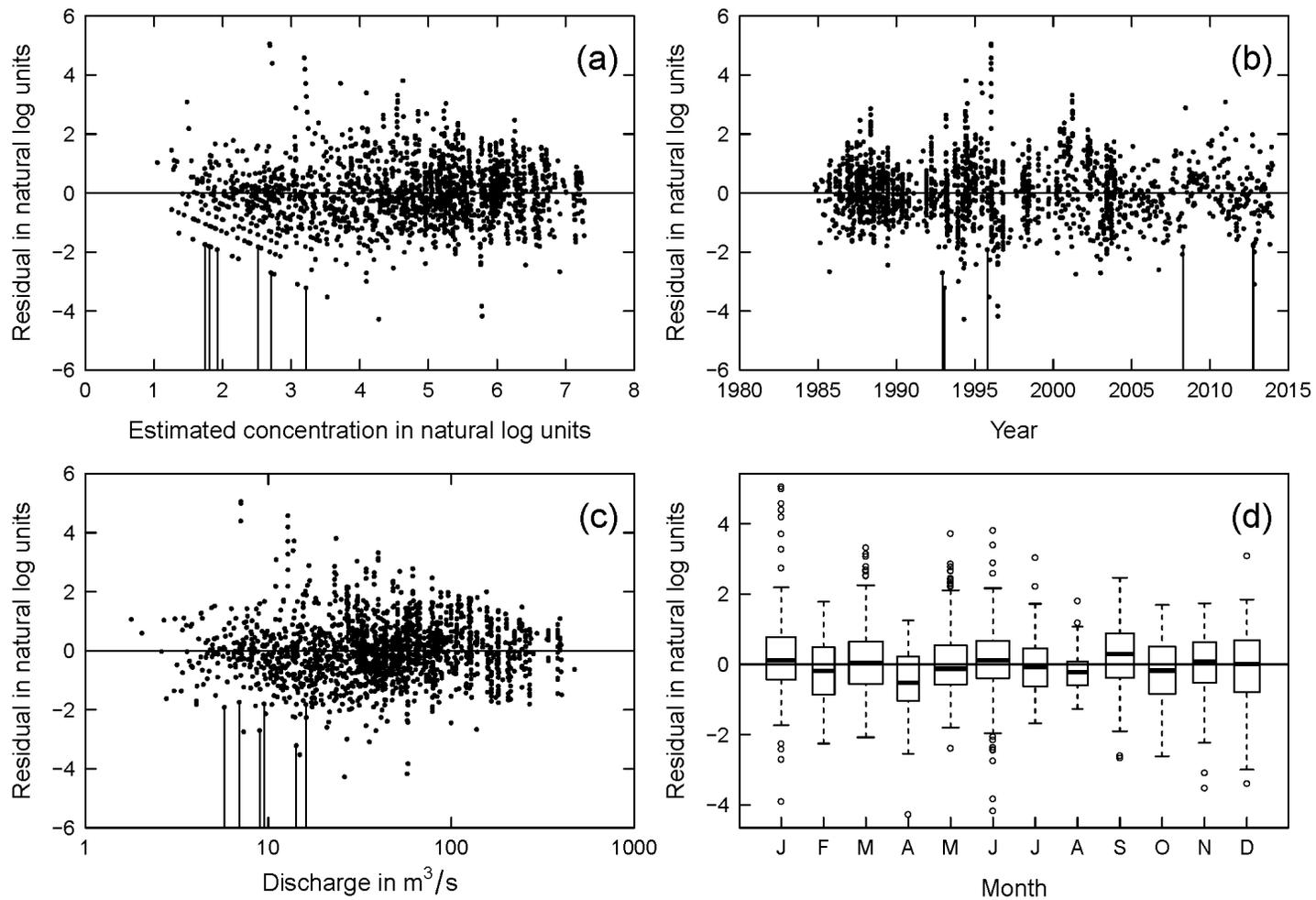


Figure B7. Model residual plot for suspended sediment in the *Conestoga River at Conestoga, PA*, showing residual (in natural log units) as a function of (a) estimated concentration, (b) year, (c) discharge, and (d) month.

### **B2-III. Uncertainty analysis of WRTDS loading estimates**

To provide estimates of the uncertainty of the WRTDS-estimated loadings, we have employed the “block-bootstrap” method of Hirsch *et al.* (2015) which involves resampling (with replacement) of the raw concentration data to obtain multiple realizations of representative data sets and associated WRTDS estimates of daily and annual loadings. This approach serves to both (1) “approximately [maintain] the short term serial correlation structure that exists in the data set but without having to attempt to model that serial correlation structure” and (2) “[prevent] the bootstrapping procedure from oversampling ... the denser sampled periods of the record.”<sup>4</sup> In brief, one single date is randomly sampled from the period of record and then all samples within a fixed block length after the selected date are chosen to constitute a first set of samples within the bootstrap replicate set. This procedure is repeated and each newly selected set of samples is added to the bootstrap replicate set until the replicate set has the same number of samples as the original data set. The selection of an appropriate block length for representative analysis has been investigated thoroughly by Hirsch *et al.* (2015) The authors reported that for water-quality data similar in duration and sampling frequency to those studied here, a block length of 100 days is adequate and that larger values give similar results. Following Hirsch *et al.* (2015) we have used a block length of 200 days. This block-bootstrap analysis was performed for each constituent at Conowingo 100 times to obtain 100 sets of daily reservoir output loading estimates. Similarly, the procedure was performed on Marietta and Conestoga concentration data to obtain 100 sets of daily inputs. Mass-balance analysis of these input and output loadings resulted in 100 sets of estimates for daily net deposition, which can then be averaged to obtain annual net deposition. Uncertainty results for such annualized average loadings are shown in Figure B8 as an indication of the level of uncertainty in the WRTDS estimates. These same replicate data sets are also used for uncertainty analysis on the output/input ratios that we calculated from 35-day running averages of daily reservoir output and input loadings – see Figure 3.3 in Section 3.3.2.2.

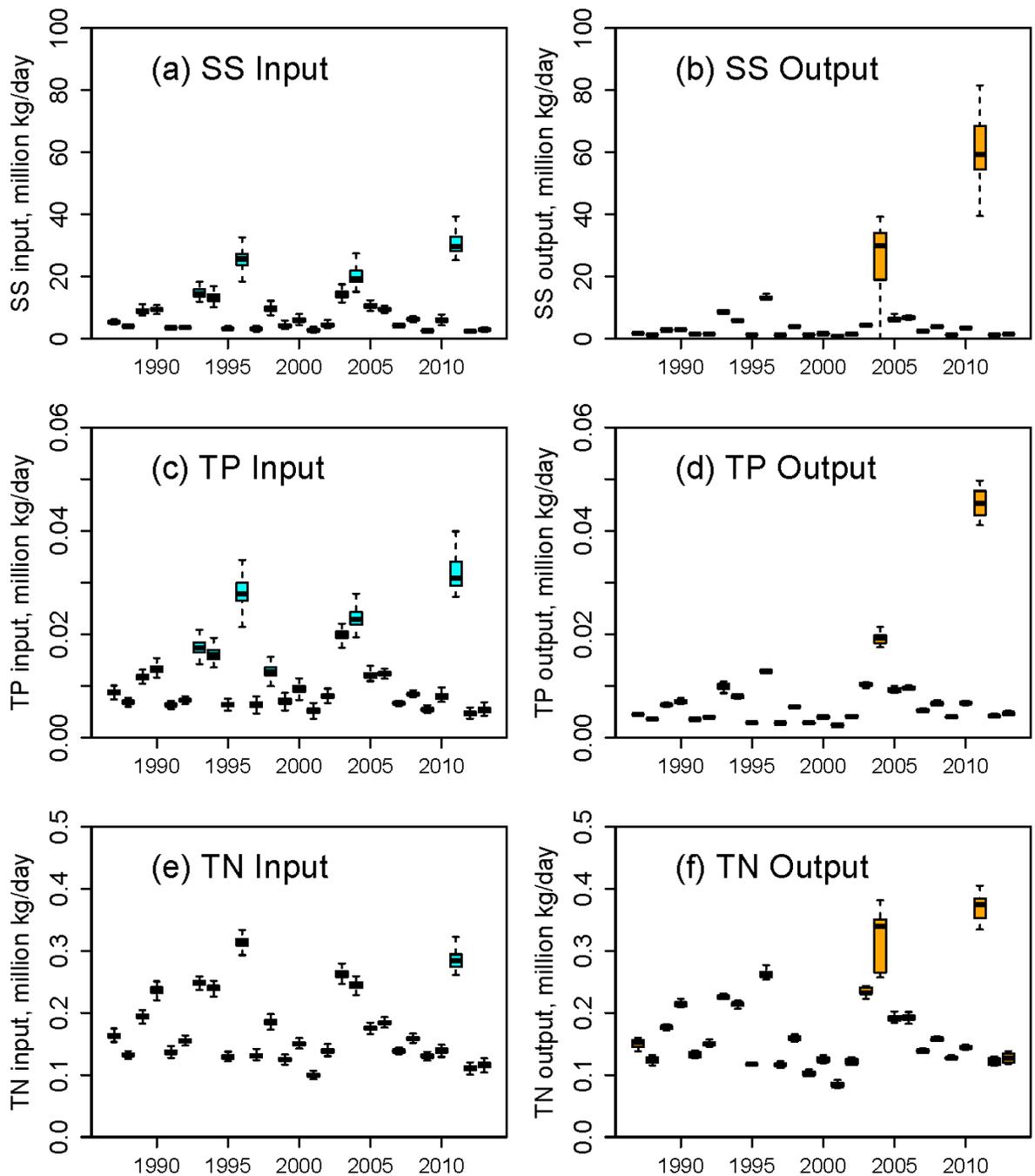
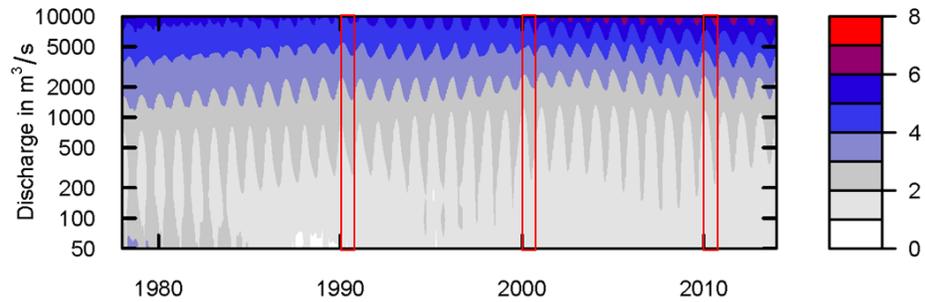


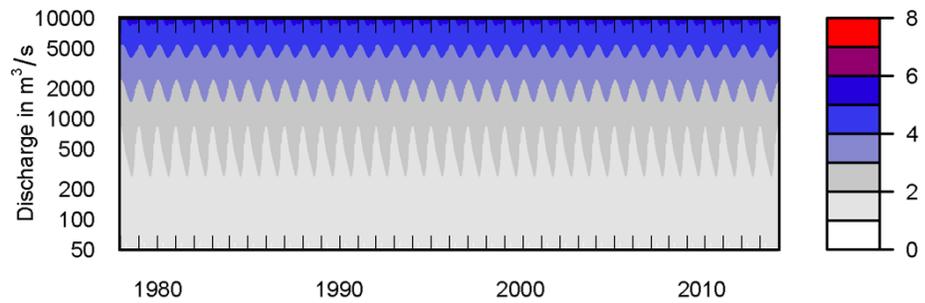
Figure B8. Uncertainty in WRTDS-estimated annual loadings in reservoir inputs and outputs of suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN). Each annual boxplot indicates the 95% confidence interval of loading, as based on 100 replicate runs conducted using a method of sampling with replacement – see text for details.

## B2-IV. Development of stationary WRTDS concentration surfaces

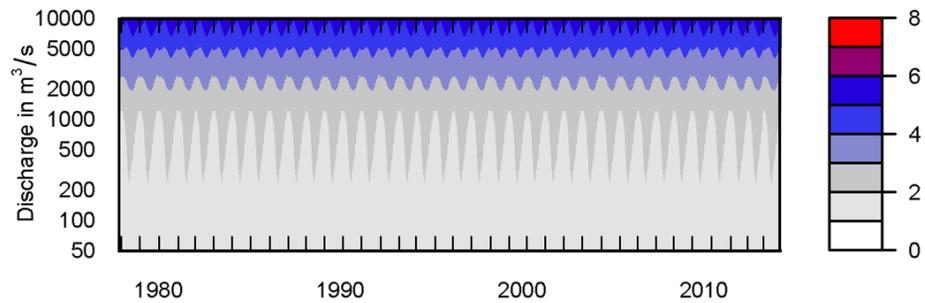
(a) The Standard WRTDS Model for Conowingo



(b) The 1990-surface-based Stationary Model



(c) The 2000-surface-based Stationary Model



(d) The 2010-surface-based Stationary Model

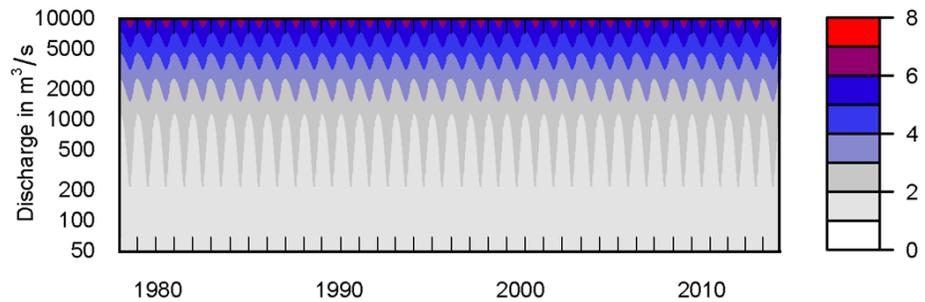


Figure B9 (caption appears on the next page)

Figure B9. Development of three “stationary” concentration surfaces for suspended sediment in the Susquehanna River at Conowingo, MD, with estimated log concentration shown by color. Plot (a) shows the concentration surface from the standard WRTDS model for the period of 1978-2014 and the selection of three one-year-wide (*i.e.*, 1990, 2000, and 2010) surfaces (indicated by red-bordered intervals). Each of the three 1-year surfaces was separately repeated to fill in the entire time-span, thus producing three stationary surfaces for the entire record, as illustrated in plots (b)-(d). These period-of-record surfaces were then applied to the actual history of daily discharge to estimate three scenarios of daily loadings.

### Appendix B3. Supporting information to Section 3.3.1

This appendix contains Figure B10 that is relevant to Section 3.3.1.

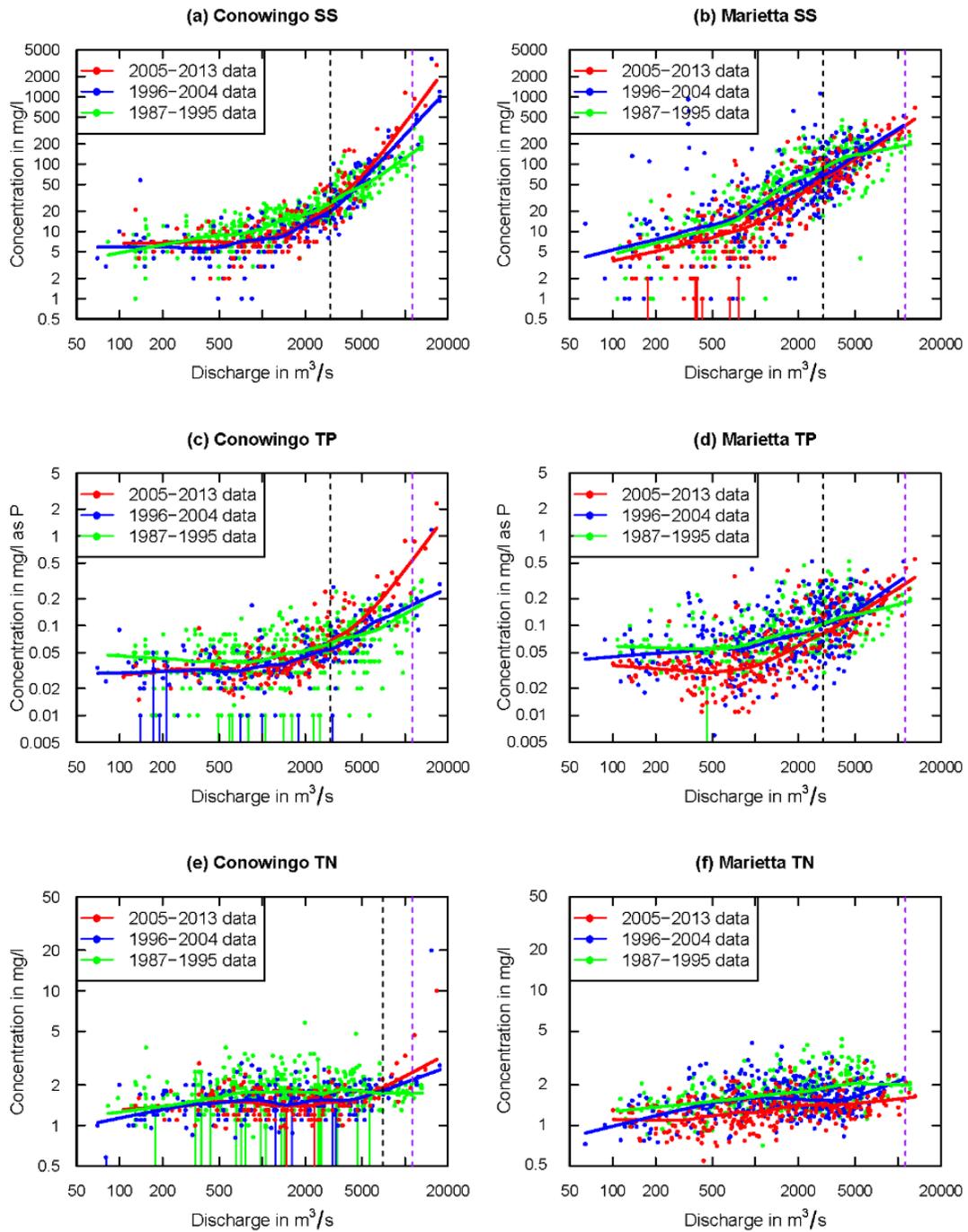


Figure B10 (caption appears on the next page)

Figure B10. Observed concentration-discharge relationships for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN) at Conowingo and Marietta for three separate periods between 1987 and 2013. Solid lines are fitted LOWESS (locally weighted scatterplot smoothing) curves. Vertical black dashed lines correspond to 3,000 m<sup>3</sup>/s in (a)-(d) and 7,000 m<sup>3</sup>/s in (e). Vertical purple dashed lines in (a)-(f) indicate the literature scour threshold of 11,300 m<sup>3</sup>/s (400,000 ft<sup>3</sup>/s). Data points with vertical solid lines indicate left-censored concentration samples. (Detection limit varied with samples.)

## Appendix B4. Supporting information to Section 3.3.2

This appendix contains tables and figures that are relevant to Section 3.3.2.

### B4-I. Reconstruction of cumulative sediment deposition behind Conowingo Dam

Cumulative sediment deposition behind the Conowingo Dam was reconstructed between 1987 and 2013 by applying mass-balance analysis on WRTDS-estimated loadings. Estimates of annual sediment net deposition between 1987 and 2013 were interpreted on a storage-volume basis by assuming that the 2008 bathymetry-based estimate of sediment capacity<sup>5</sup> is correct. Our estimates of annual net deposition can then be used to estimate capacity in years before and after 2008 (Figure B11). For each year, the 100 estimates of annual net deposition (see Appendix B2-III) were ranked to determine the 2.5<sup>th</sup> and 97.5<sup>th</sup> quantiles, which formed a 95% confidence interval (CI). This CI was then pinned to the 2008 bathymetry data<sup>5</sup> to obtain a 95% CI on the cumulative deposition in 1987-2013, as shown with blue and purple points in Figure B11.

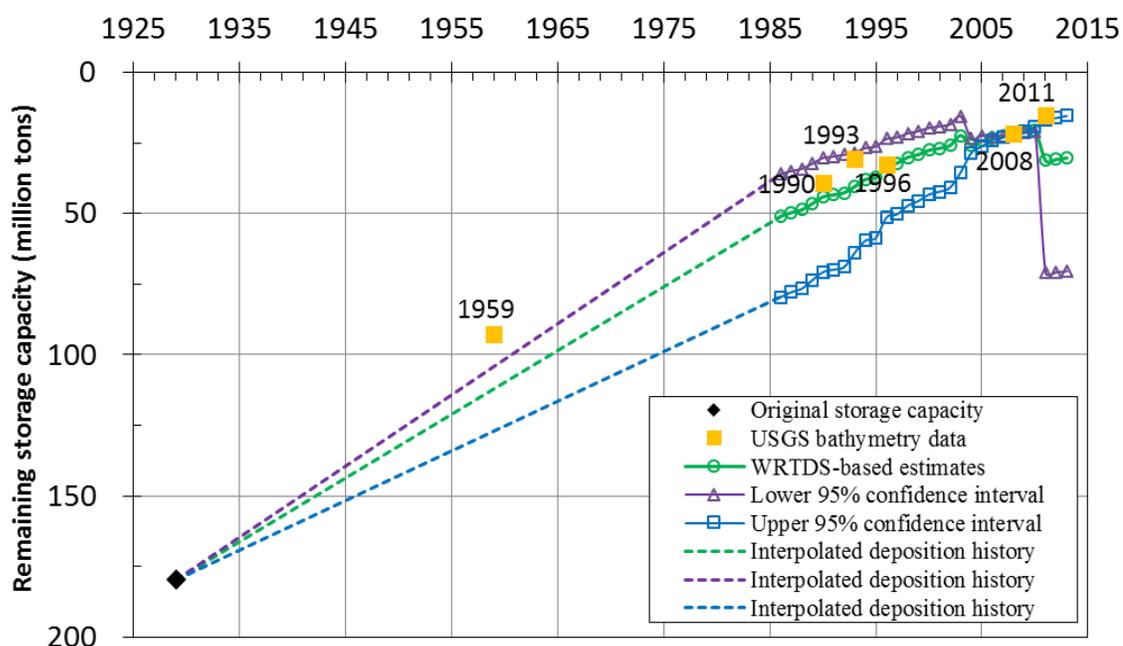


Figure B11. Reconstruction of the history of cumulative sediment deposition in Conowingo Reservoir between 1929 and 2013. See text above for method details.

Table B2. Average rate of sediment net deposition rate in Conowingo Reservoir.

	Average net deposition <sup>a</sup> (million tons/year)	95% confidence interval <sup>b</sup> (million tons/year)
1929-1987 <sup>c</sup>	2.24	1.75 ~ 2.49
1987-2000 <sup>d</sup>	1.70	1.18 ~ 2.66
2000-2010 <sup>d</sup>	0.72	-0.13 ~ 2.38

<sup>a</sup> based on standard WRTDS estimates of annual net deposition for each year in 1987-2010.

Results correspond to the green data in Figure B11.

<sup>b</sup> based on 100 set of annual net deposition estimates for each year in 1987-2010; see text for details. Results correspond to the purple and blue data in Figure B11.

<sup>c</sup> based on 1929 original storage capacity (Langland, 2015) and estimate of 1987 remaining capacity as calculated from 2008 bathymetry and WRTDS estimates of annual net deposition for each year in 1987-2008.

<sup>d</sup> based on 1987, 2000, and 2010 remaining capacities as calculated from 2008 bathymetry and WRTDS estimates of annual net deposition for each year in 1987-2010.

## B4-II. Estimation of travel time and evaluation of its effects on output/input ratios

The travel time of river flow across the reservoir system may complicate the analysis of output/input ratio. To deal with this issue, we have quantified the travel time distribution across the reservoir system. Based on compiled information from literature (Hainly *et al.*, 1995; Langland and Hainly, 1997; Langland, 2009; Langland, 2015) the three reservoirs (*i.e.*, lake Clarke, Lake Aldred, and Conowingo Reservoir) had, as of 2008, remaining (water plus sediment) storage capacities of 65,000, 50,000, and 162,000 acre-feet, respectively, which sum to 277,000 acre-feet or  $3.4 \times 10^8 \text{ m}^3$ . This estimated capacity is considered applicable to our study period (*i.e.*, 1987-2013), because the Conowingo Reservoir's storage has not decreased significantly in this period (~10%) and the other two reservoirs have been filled for decades. Based on this storage estimate and the daily flow distribution at Marietta (system inlet), the travel time distribution across the reservoir system can be calculated, with results summarized in Table B3.

Table B3. Travel time distribution across the reservoir system.

Flow condition	Marietta flow ( $\text{m}^3/\text{s}$ )	Travel time (days)
Minimum	63	62
1 <sup>th</sup> percentile	112	35
5 <sup>th</sup> percentile	161	25
25 <sup>th</sup> percentile	387	10
Median	774	5.1
Mean	1,149	3.5
75 <sup>th</sup> percentile	1,418	2.8
Maximum	18,040	0.2

Therefore, it takes 35 days or less for water to travel through the reservoir system for 99% of the days in the study period of 1987-2013. In light of this result and also the concern that human intervention in dam releases may vary by weekday, it is appropriate to use 35 days (five weeks) as the temporal scale for the output/input ratio analysis. Specifically, we have calculated the 35-day moving average for daily input, the 35-day moving average for daily output, and their ratios. These results are presented in Figure 3.3 in the main text.

Sensitivity of trend in output/input ratio to the selection of averaging period is shown in Figures B12-B14 for suspended sediment (SS), total phosphorus (TP), and total

nitrogen (TN), respectively. These figures show that similar trends are observed irrespective of averaging period within the range of daily to 35 days. (Note that this moving-average approach allows a fair comparison among the three temporal scales by consistently plotting between 330 and 365 days in each year.) In general, all three figures indicate that, as the averaging period gets longer, fewer outliers are present because the issue of travel time has been effectively taken into account.

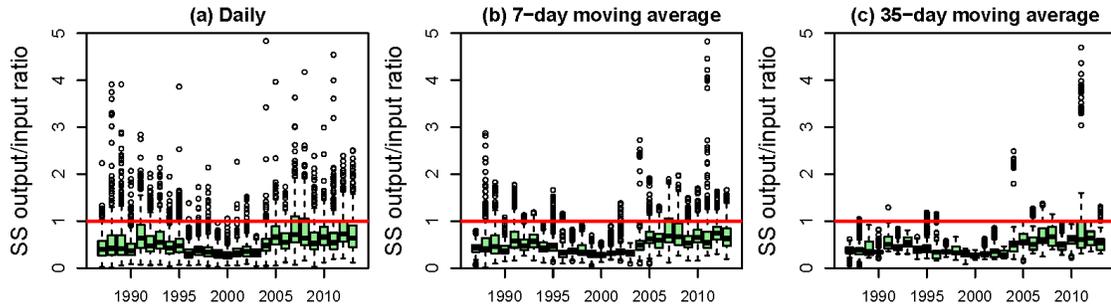


Figure B12. Boxplots of suspended sediment (SS) output/input ratio based on (a) daily loads, (b) 7-day moving averages, and (c) 35-day moving averages.

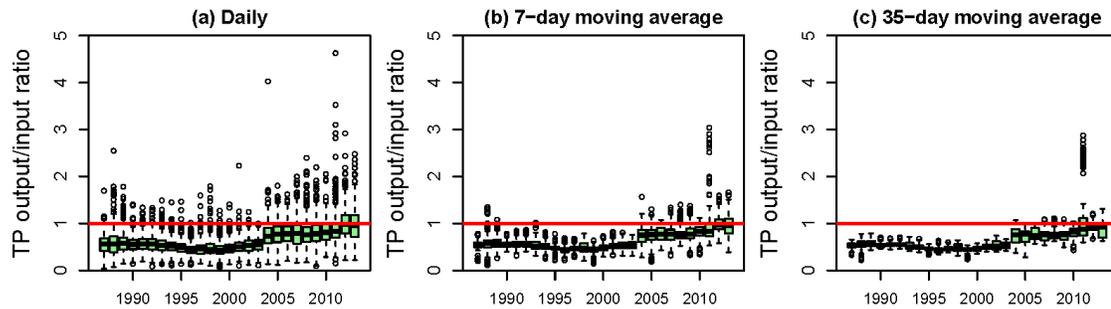


Figure B13. Boxplots of total phosphorus (TP) output/input ratio based on (a) daily loads, (b) 7-day moving averages, and (c) 35-day moving averages.

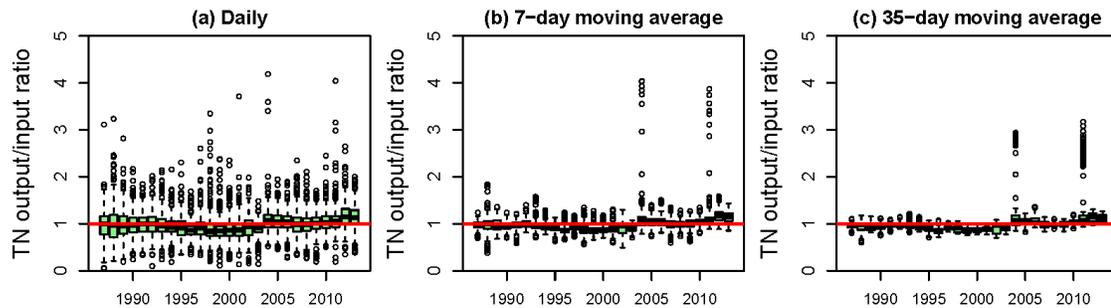


Figure B14. Boxplots of total nitrogen (TN) output/input ratio based on (a) daily loads, (b) 7-day moving averages, and (c) 35-day moving averages.

### B4-III. Uncertainty in the numbers of excursions of output/input ratios in each year

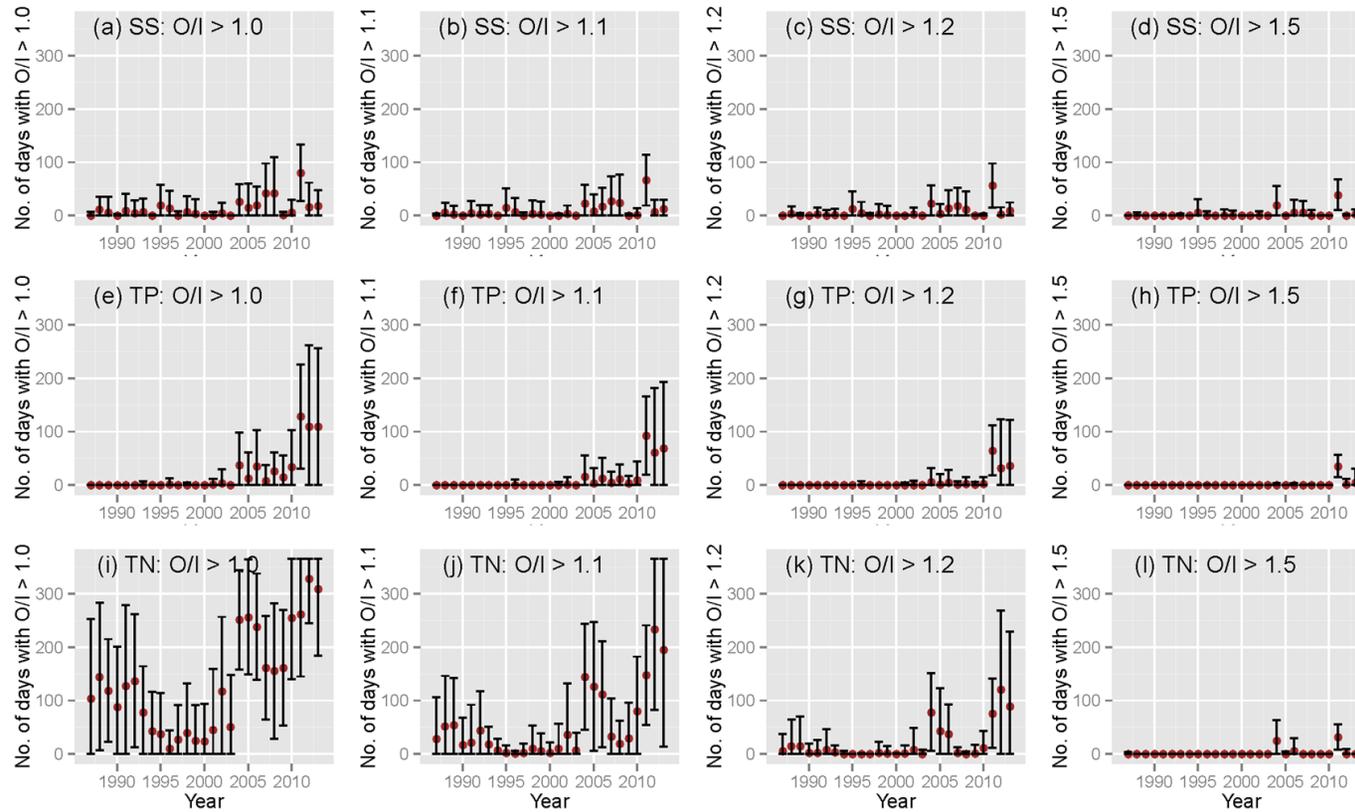


Figure B15. Uncertainty in the number of excursions of output/input ratios (O/I) for suspended sediment (SS), total phosphorus (TP), and total nitrogen (TN). The excursions indicate conditions of net scour as defined by O/I greater than four different cut-off thresholds, *i.e.*, 1, 1.1, 1.2, and 1.5. Each linechart shows the mean +/- 2 standard deviation (95% confidence interval) of the counts, which were obtained with 100 replicate runs on resampled (with replacement) concentration data (see Appendix B2-III for details).

#### B4-IV. Sensitivity analysis of results to highflow samples at Marietta and Conowingo

Figures B16-B18 below summarize the available samples in the Susquehanna River at the Marietta and Conowingo sites. Sampling at Conowingo and Marietta began in 1978 and 1986, respectively. For visual clarity, the y-axis in these figures only shows those days when Conowingo flow exceeded 5,000 m<sup>3</sup>/s (~176,600 ft<sup>3</sup>/s). For the period when both sites were monitored, data of suspended sediment (SS) show that (Figure B16):

- (1) neither site was sampled when flow exceeded ~ 20,000 m<sup>3</sup>/s in 2011 (1 occurrence: 2011/09/09),
- (2) only Conowingo was sampled when flows were between 15,000 m<sup>3</sup>/s and 20,000 m<sup>3</sup>/s (3 occurrences: 1996/01/21, 2004/09/20, and 2011/09/08),
- (3) both sites were sampled similarly when flows were between 10,000 m<sup>3</sup>/s and 15,000 m<sup>3</sup>/s, and
- (4) both sites were sampled with comparative missing rates when flows were between 5,000 m<sup>3</sup>/s and 10,000 m<sup>3</sup>/s.

In addition, data of total phosphorus (TP) and total nitrogen (TN) present very similar results to those of suspended sediment, as shown in Figures B17 and B18, respectively. Based on these observations, the major distinction between Marietta and Conowingo with respect to highflow sampling appear to lie in the region of 15,000-20,000 m<sup>3</sup>/s, for which three dates were sampled at Conowingo but not Marietta.

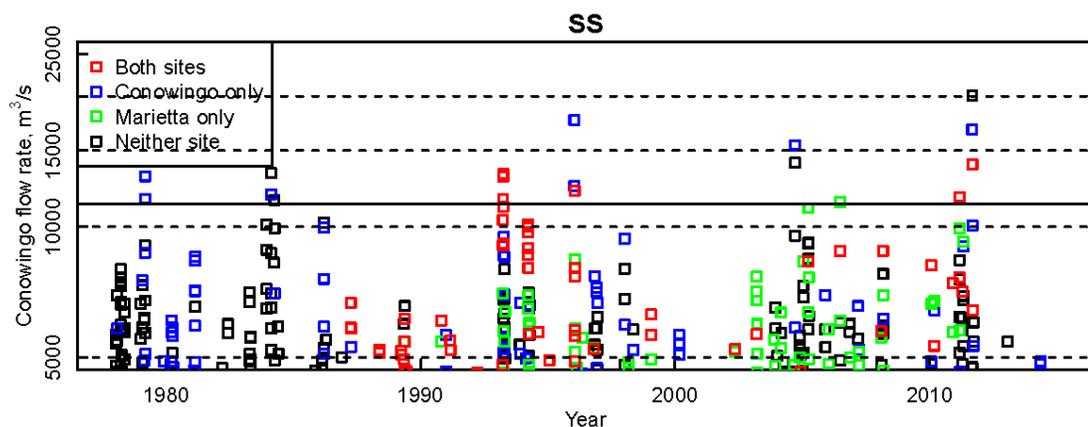


Figure B16. Monitoring data of suspended sediment (SS) in the Susquehanna River at Marietta and Conowingo. Solid horizontal line indicates the literature-reported scour threshold of 11,300 m<sup>3</sup>/s (400,000 ft<sup>3</sup>/s). Note that y-axis is on log-scale.

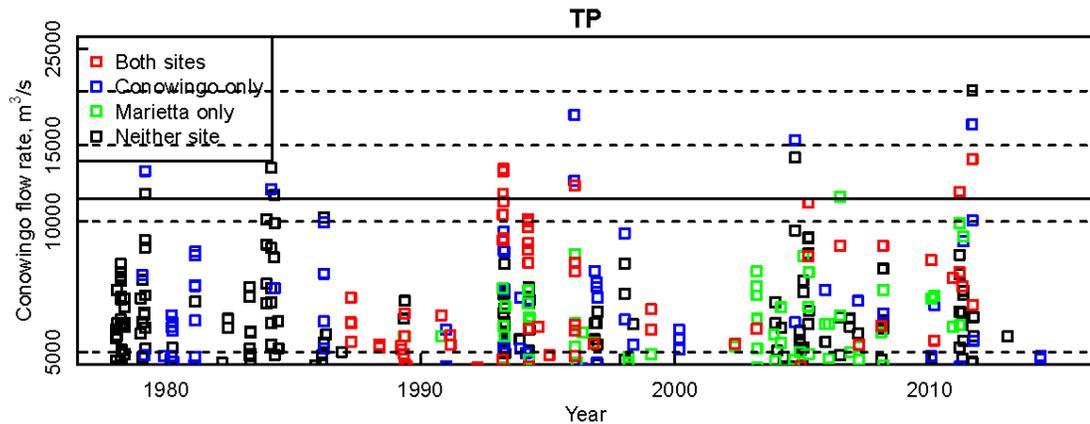


Figure B17. Monitoring data of total phosphorus (TP) in the Susquehanna River at Marietta and Conowingo. Solid horizontal line indicates the literature-reported scour threshold of 11,300 m<sup>3</sup>/s (400,000 ft<sup>3</sup>/s). Note that y-axis is on log-scale.

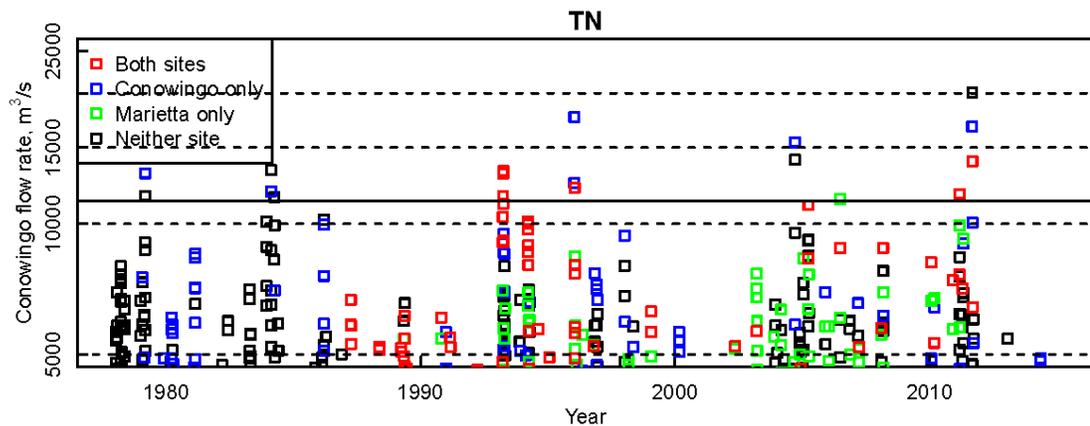


Figure B18. Monitoring data of total nitrogen (TN) in the Susquehanna River at Marietta and Conowingo. Solid horizontal line indicates the literature-reported scour threshold of 11,300 m<sup>3</sup>/s (400,000 ft<sup>3</sup>/s). Note that y-axis is on log-scale.

In this regard, a sensitivity analysis was conducted to investigate whether this difference in highflow sampling has induced any major effects on the conclusions with respect to changing performance of Conowingo Reservoir. Specifically, we have re-run our analysis with the standard WRTDS model on Marietta and Conowingo by excluding samples collected during flows above 15,000 m<sup>3</sup>/s. This threshold was chosen because

these are flows that have the most distinction in number of samples between the two sites. Note that this value is well above the literature-reported scour threshold of 11,300 m<sup>3</sup>/s (400,000 ft<sup>3</sup>/s) and that the resulting censored data set will therefore generate WRTDS surfaces that are more uncertain for estimating concentrations and loads during > 15,000 m<sup>3</sup>/s flows at both sites. This is relevant because such flows still exist in the estimation data even though the related samples have not been used in the model generation.

This sensitivity analysis was limited to SS and TP, because the two species showed diminished trapping by the reservoir. As in the main text (Section 3.3.2), we quantified output/input ratios for SS and TP for each complete year (*i.e.*, 1987-2013), as shown in Figure B19. Clearly, the SS output/input ratios have been rising since the early 2000s and the TP output/input ratios have shown a similar trend. These results are consistent with those obtained based on WRTDS runs on all samples. (See Figure 3.3 in the main text.)

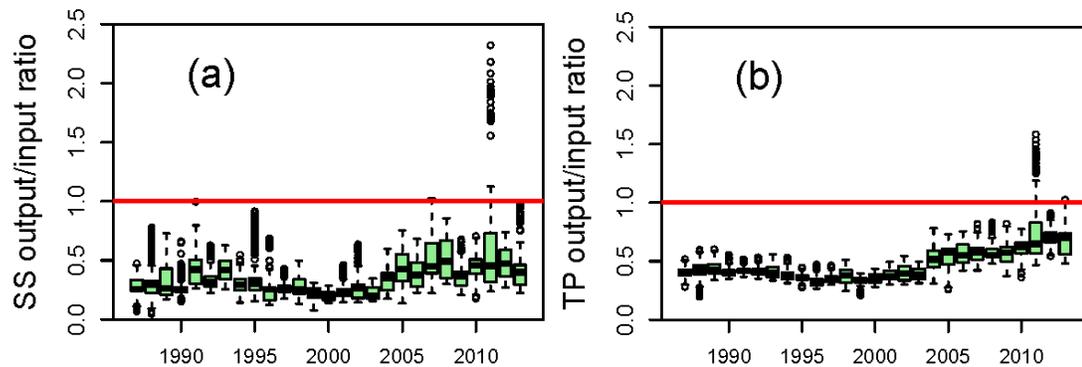


Figure B19. Boxplots of output/input ratios based on 35-day moving averages of output and input for (a) suspended sediment (SS) and (b) total phosphorus (TP), as obtained from standard WRTDS models runs that exclude samples collected during highflow days (*i.e.*, flow > 15,000 m<sup>3</sup>/s) from the analysis.

### B4-V. WRTDS estimates of input, output, and output/input ratio under different flow conditions

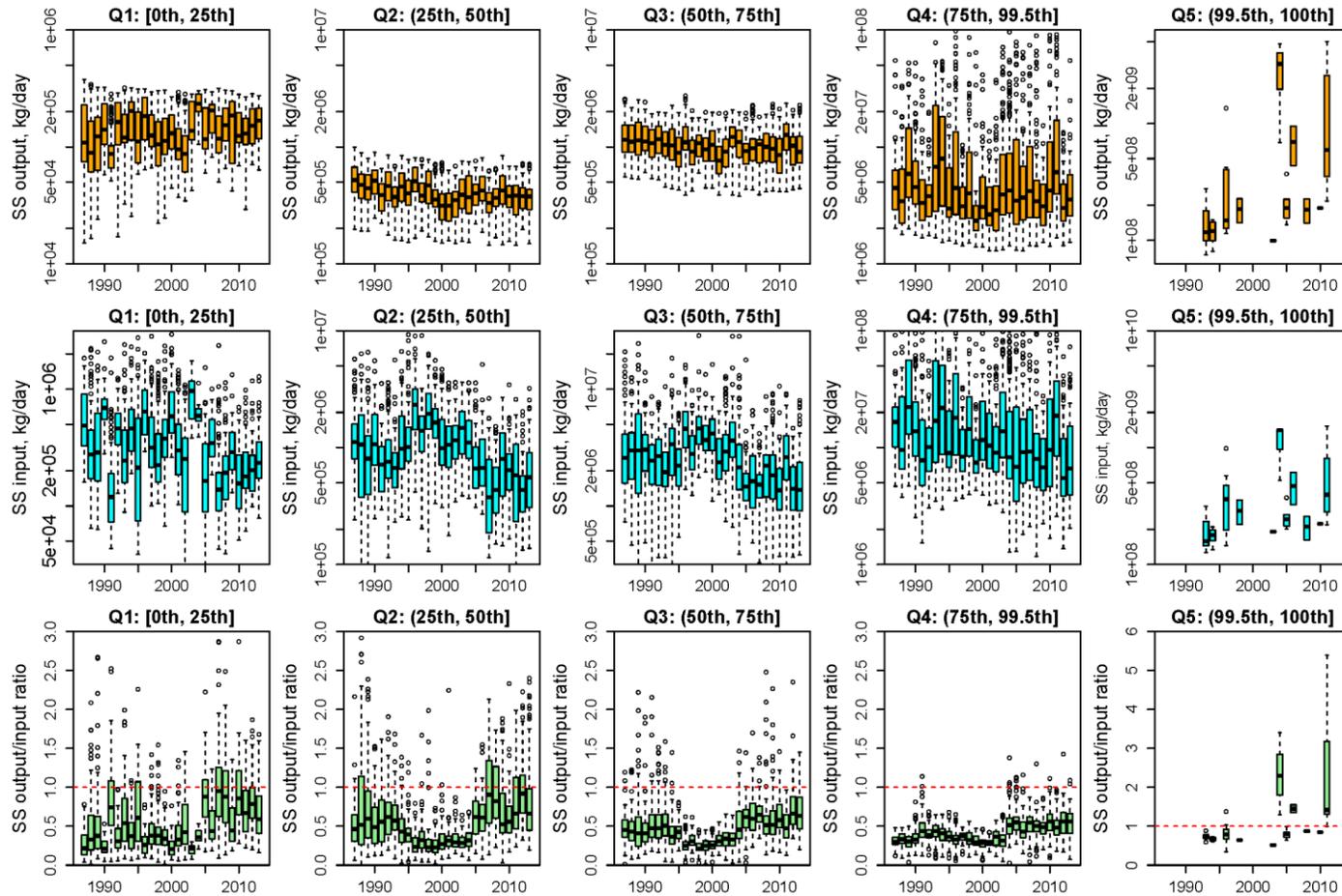


Figure B20. Annual boxplots of WRTDS daily estimates of suspended sediment (SS) for reservoir output (top), reservoir input (middle), and output/input ratio (bottom) under different flow conditions. Note that y-axis is on log-scale for top and middle panels.

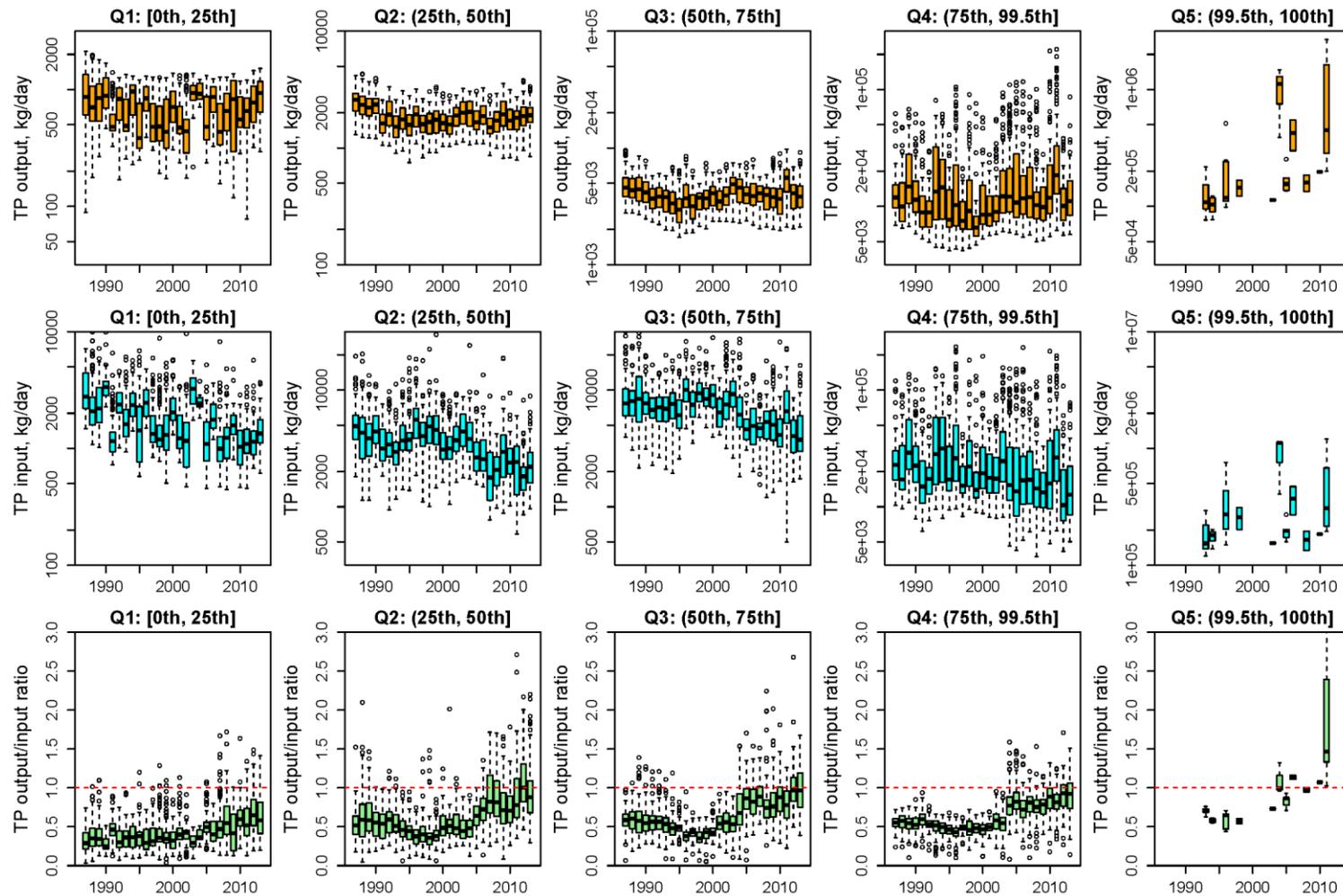


Figure B21. Annual boxplots of WRTDS daily estimates of total phosphorus (TP) for reservoir output (top), reservoir input (middle), and output/input ratio (bottom) under different flow conditions. Note that y-axis is on log-scale for top and middle panels.

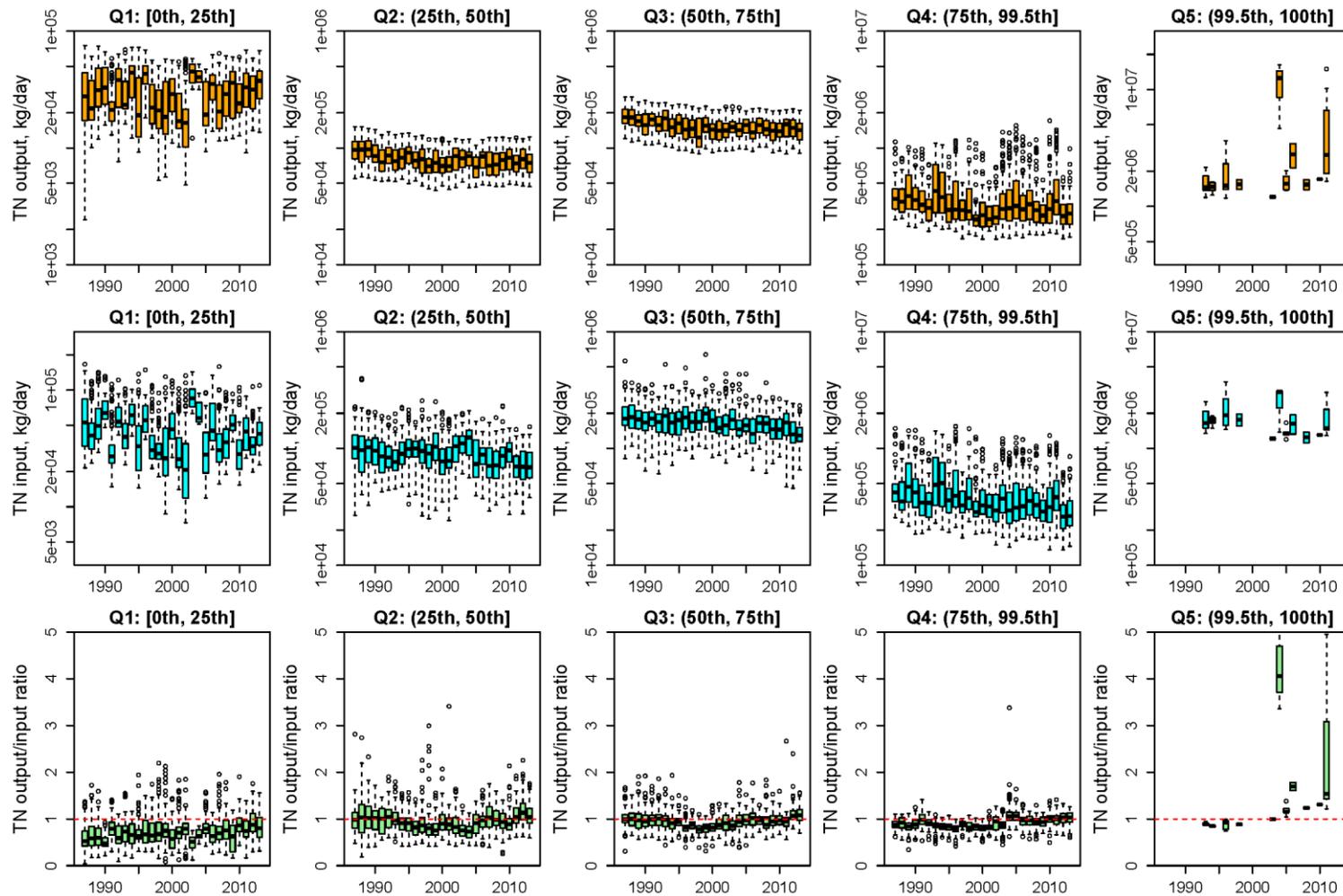


Figure B22. Annual boxplots of WRTDS daily estimates of total nitrogen (TN) for reservoir output (top), reservoir input (middle), and output/input ratio (bottom) under different flow conditions. Note that y-axis is on log-scale for top and middle panels.

## Appendix B5. Supporting information to Section 3.3.3

This appendix contains figures that are relevant to Section 3.3.3. Specifically:

- Figures B23-B31 (Appendix B5-I) are frequency plots of ranked loadings for reservoir output, output, and net deposition based on the three stationary WRTDS models, which are discussed in Section 3.3.3.1. These plots rank the estimated loadings and the count the number of days per year that these loadings exceed certain values under each model. For visual clarity, these plots also contain three enlarged portions, that is, 0-15, 100-115, and 200-215 days/year. Note that y-axis is on log-scale.
- Figures B32-B40 (Appendix B5-II) show loading-discharge relationships for reservoir output, output, and net deposition based on the three stationary WRTDS models, which are discussed in Section 3.3.3.2. These loading-discharge relationships were modeled with LOWESS (locally weighted scatterplot smoothing) curves using the function “*lowess*” in *R* (with a smoother span of 0.5). For visual clarity, these plots also contain three enlarged portions, that is, 500-2000, 2000-5000, and 5000-11300 m<sup>3</sup>/s. Note that both axes are on log-scale. Readers should appreciate that these statistically modeled loading-discharge relationships are highly uncertain at the highest discharges (*i.e.*, those associated with infrequent weather-related events) due to the scarcity of concentration monitoring data at one or more locations during these discharges and the related fact that available measurements may not accurately represent the proper flow-weighted distribution of concentration. To reflect these issues, we also add the observed loading data (*i.e.*, observed daily concentration x observed daily discharge) on these figures. Note that, however, the LOWESS curves were not obtained by directly fitting to these observed loading data.
- Figures B41-B43 (Appendix B5-III) show modeled cumulative reservoir output, cumulative reservoir input, and cumulative reservoir storage over the course of three selected wet, average, and dry calendar years (*i.e.*, 2003, 2007, and 2001) based on the three stationary WRTDS models, which are discussed in Section 3.3.3.3.

- All the above figures are shown with uncertainties. For figures associated with input or output, dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. For figures associated with net increase in storage, dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from the resulting 10x10 input x output matrix.

**B5-I. Frequency plots of ranked loadings for reservoir output, input, and net deposition based on stationary WRTDS models**

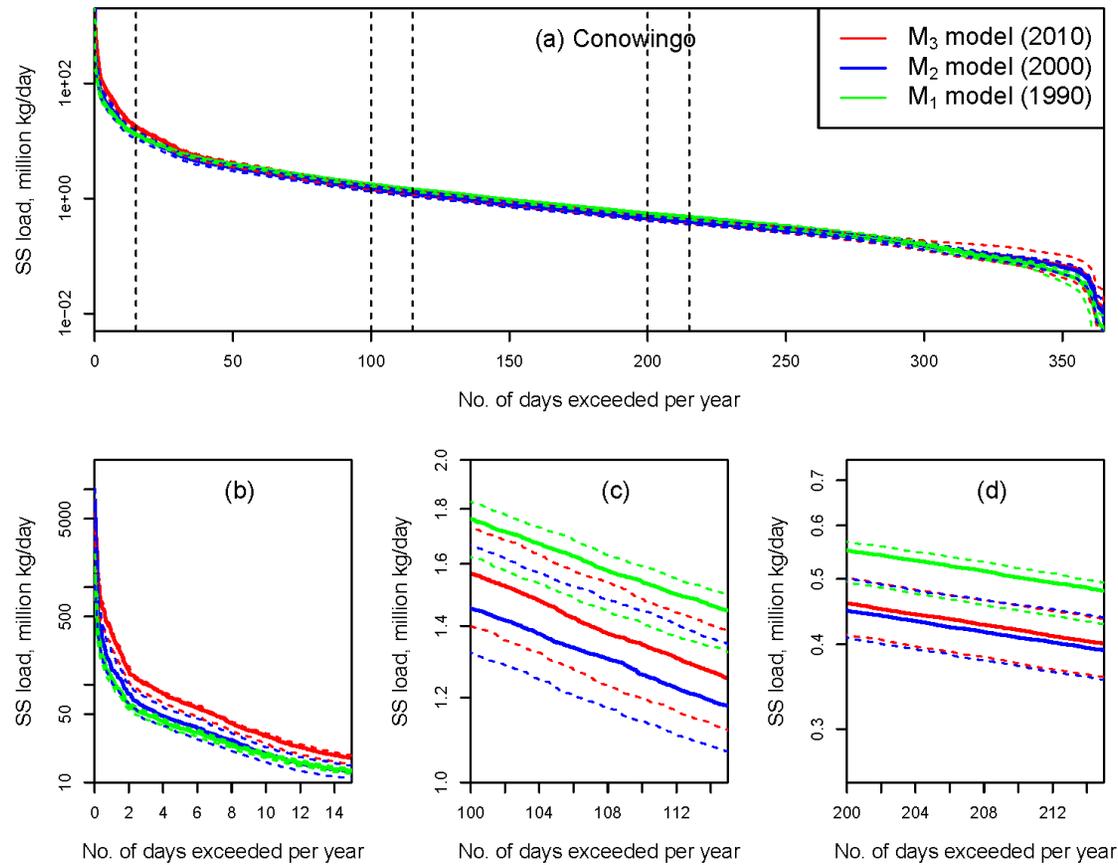


Figure B23. Frequency plots of ranked loadings for reservoir output loading of suspended sediment (SS) based on the three stationary WRTDS models. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Note that y-axis is on log-scale.

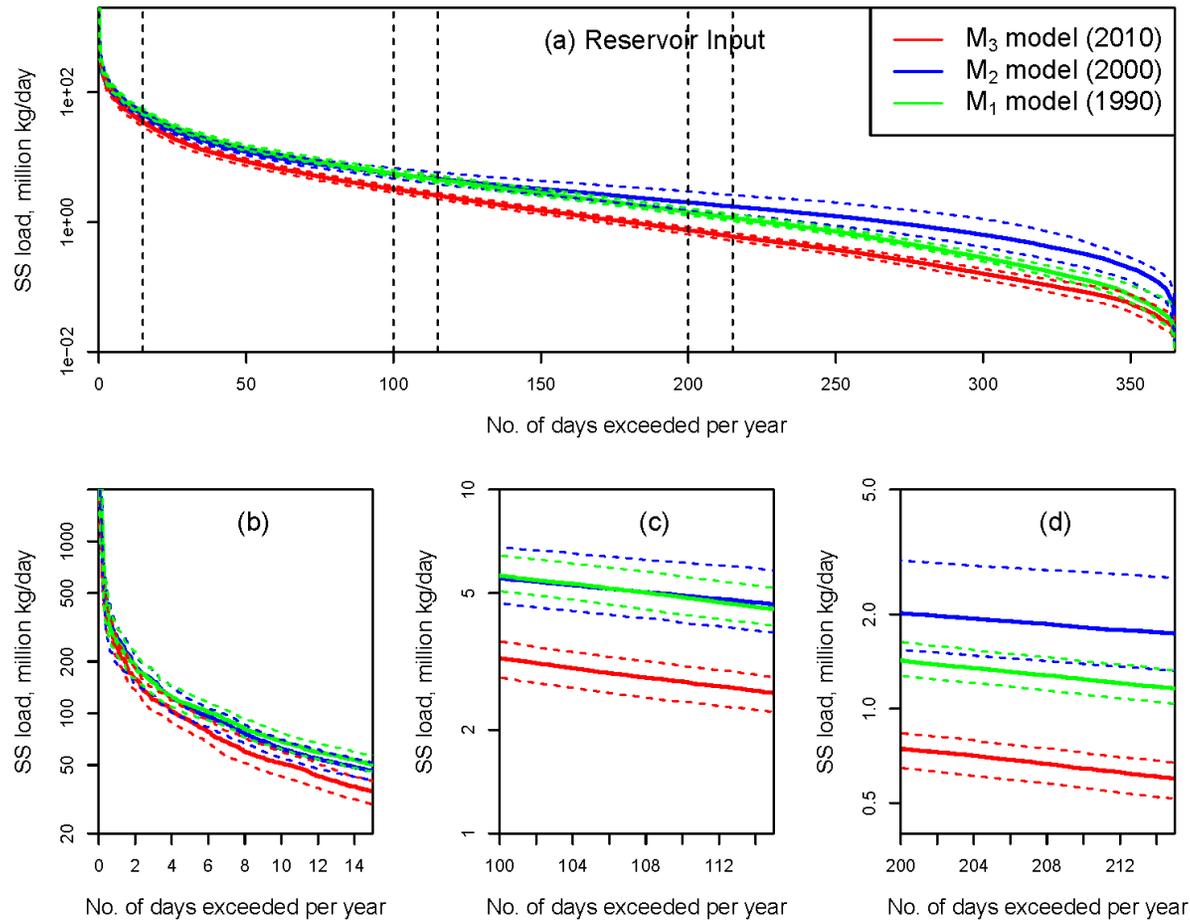


Figure B24. Frequency plots of ranked loadings for reservoir input loading of suspended sediment (SS) based on the three stationary WRTDS models. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Note that y-axis is on log-scale.

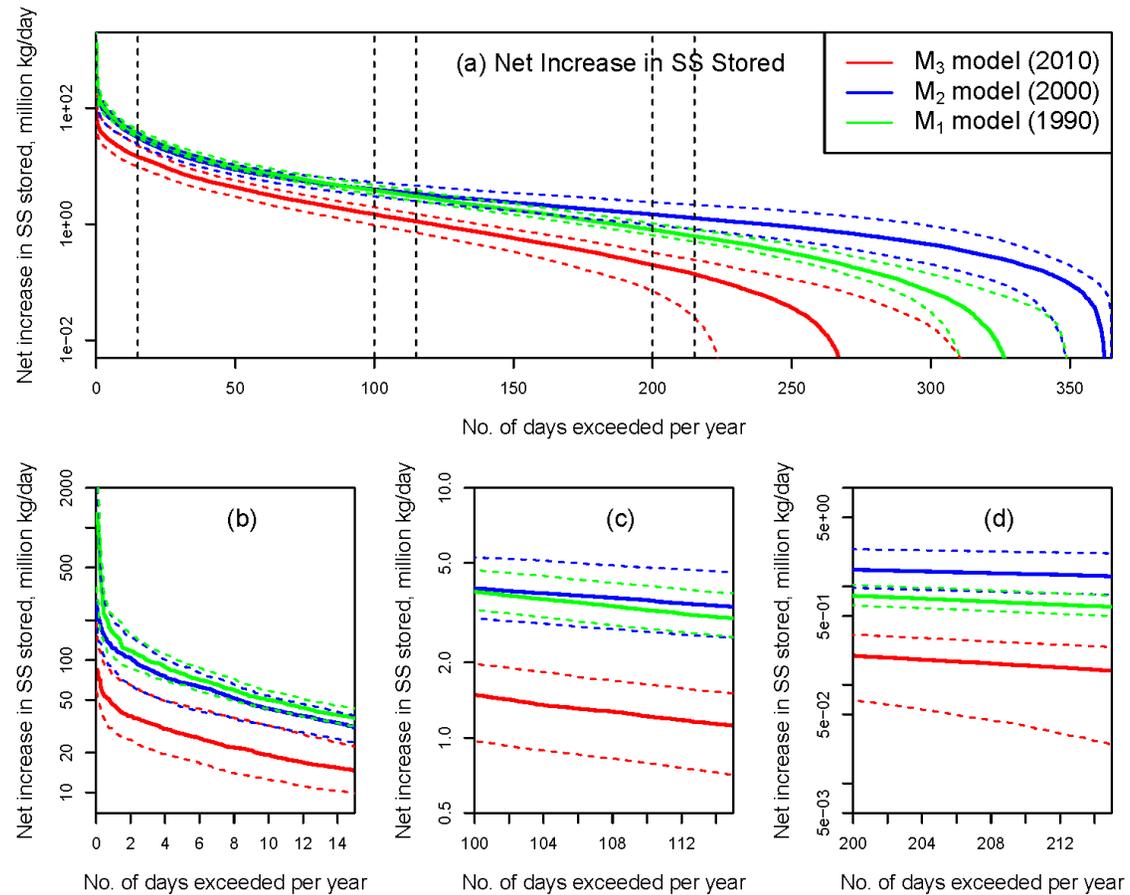


Figure B25. Frequency plots of ranked loadings for net increase in reservoir storage of suspended sediment (SS) based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. Note that y-axis is on log-scale.

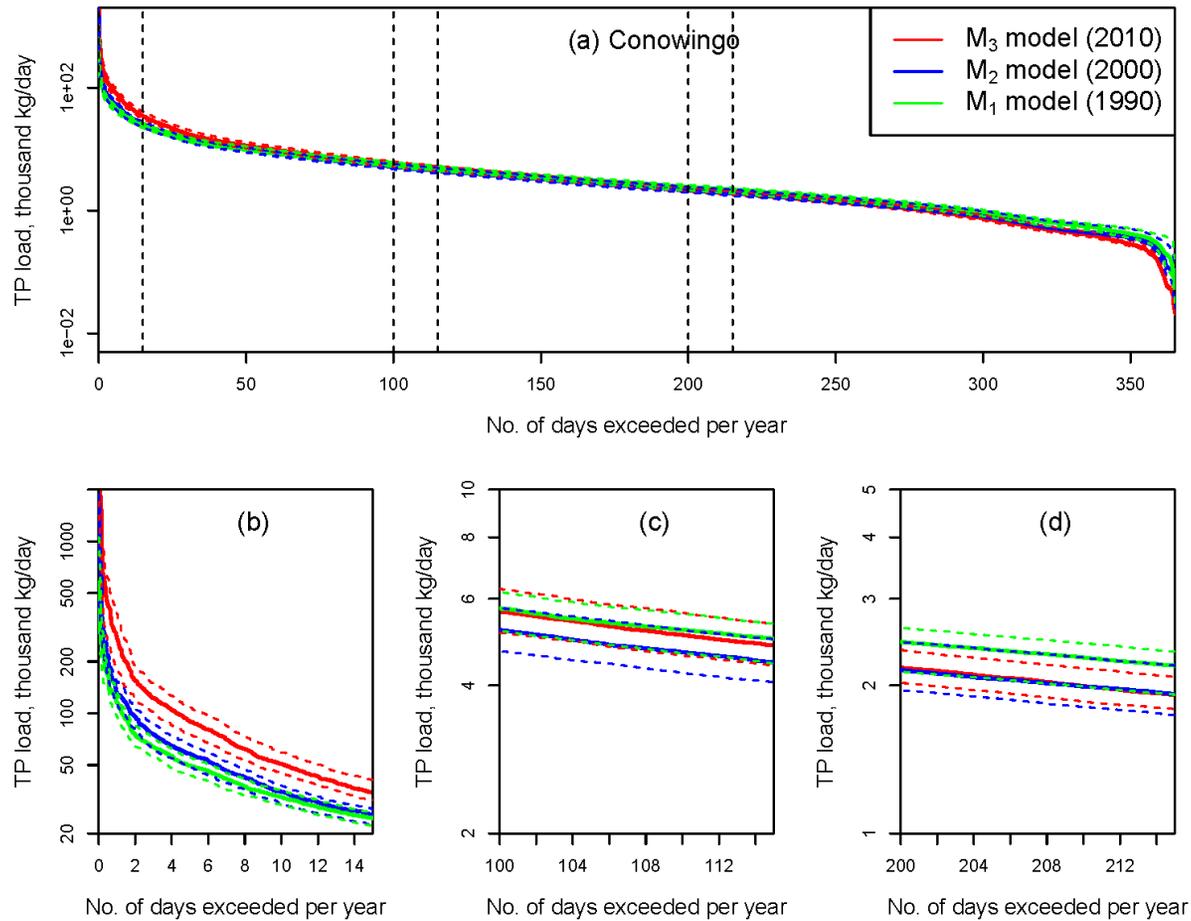


Figure B26. Frequency plots of ranked loadings for reservoir output loading of total phosphorus (TP) based on the three stationary WRTDS models. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Note that y-axis is on log-scale.

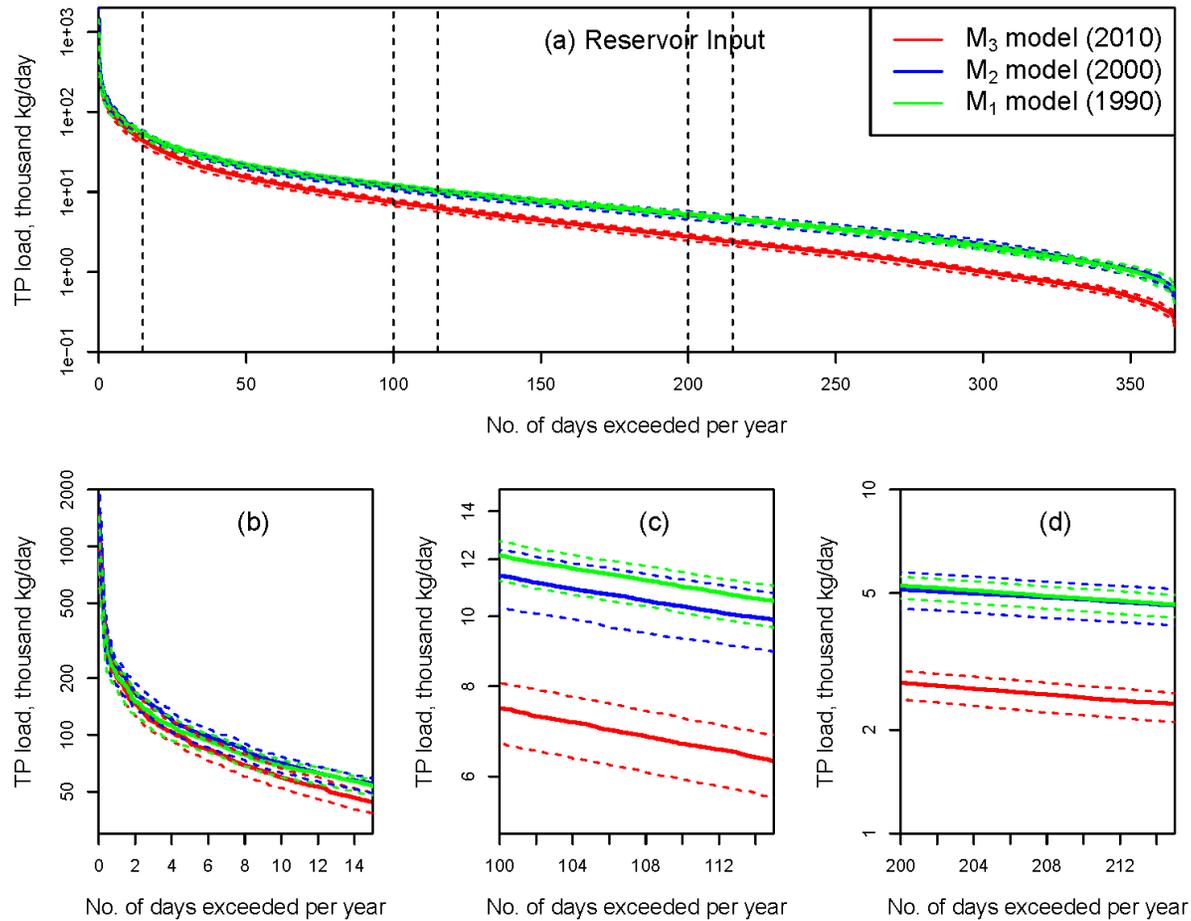


Figure B27. Frequency plots of ranked loadings for reservoir input loading of total phosphorus (TP) based on the three stationary WRTDS models. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Note that y-axis is on log-scale.

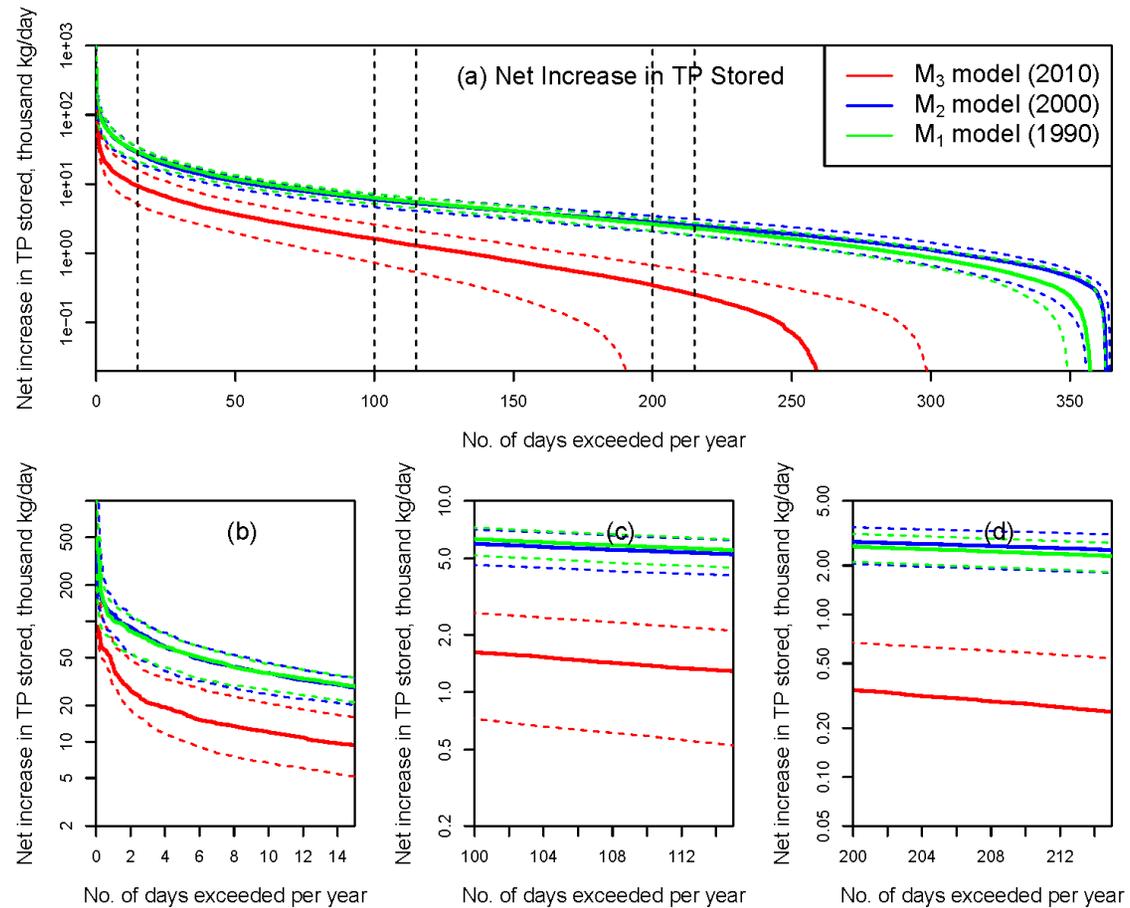


Figure B28. Frequency plots of ranked loadings for net increase in reservoir storage of total phosphorus (TP) based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. Note that y-axis is on log-scale.

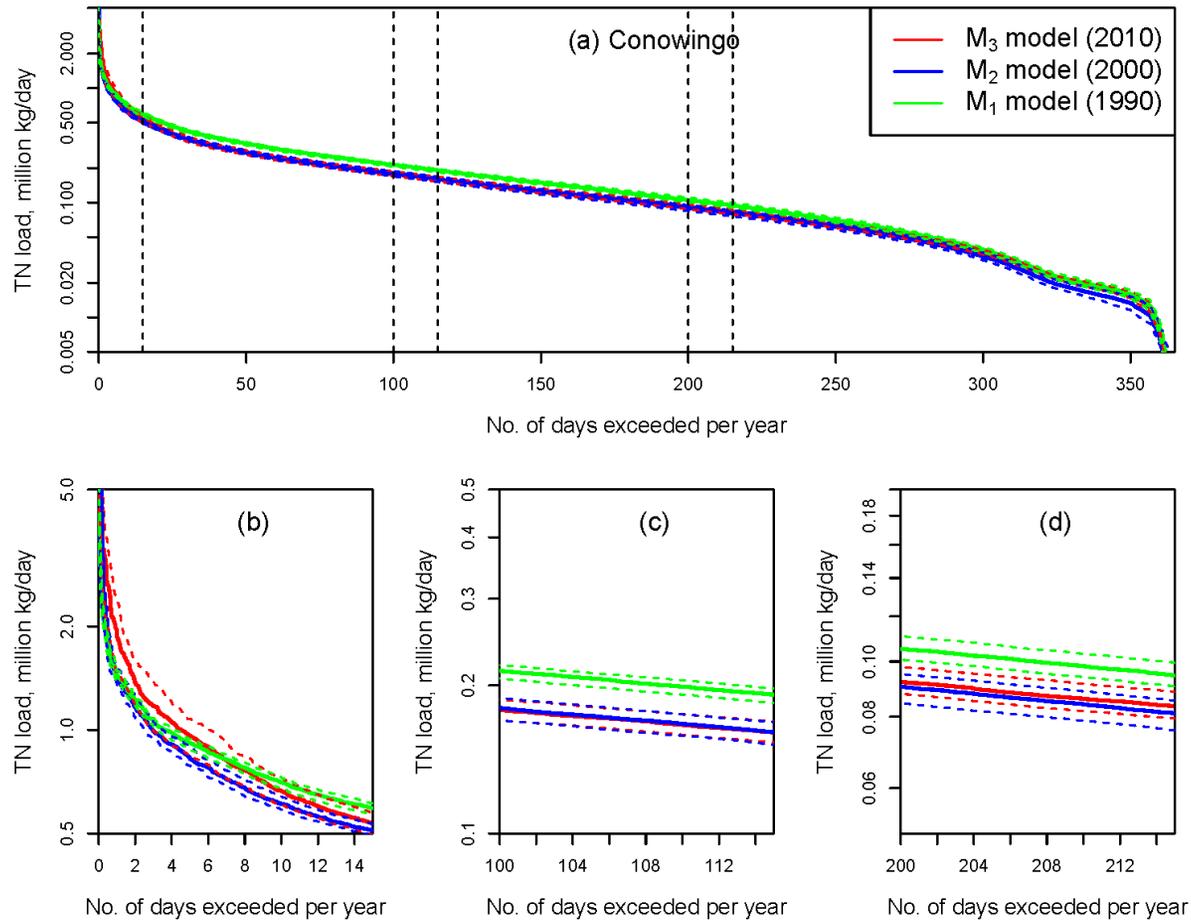


Figure B29. Frequency plots of ranked loadings for reservoir output loading of total nitrogen (TN) based on the three stationary WRTDS models. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Note that y-axis is on log-scale.

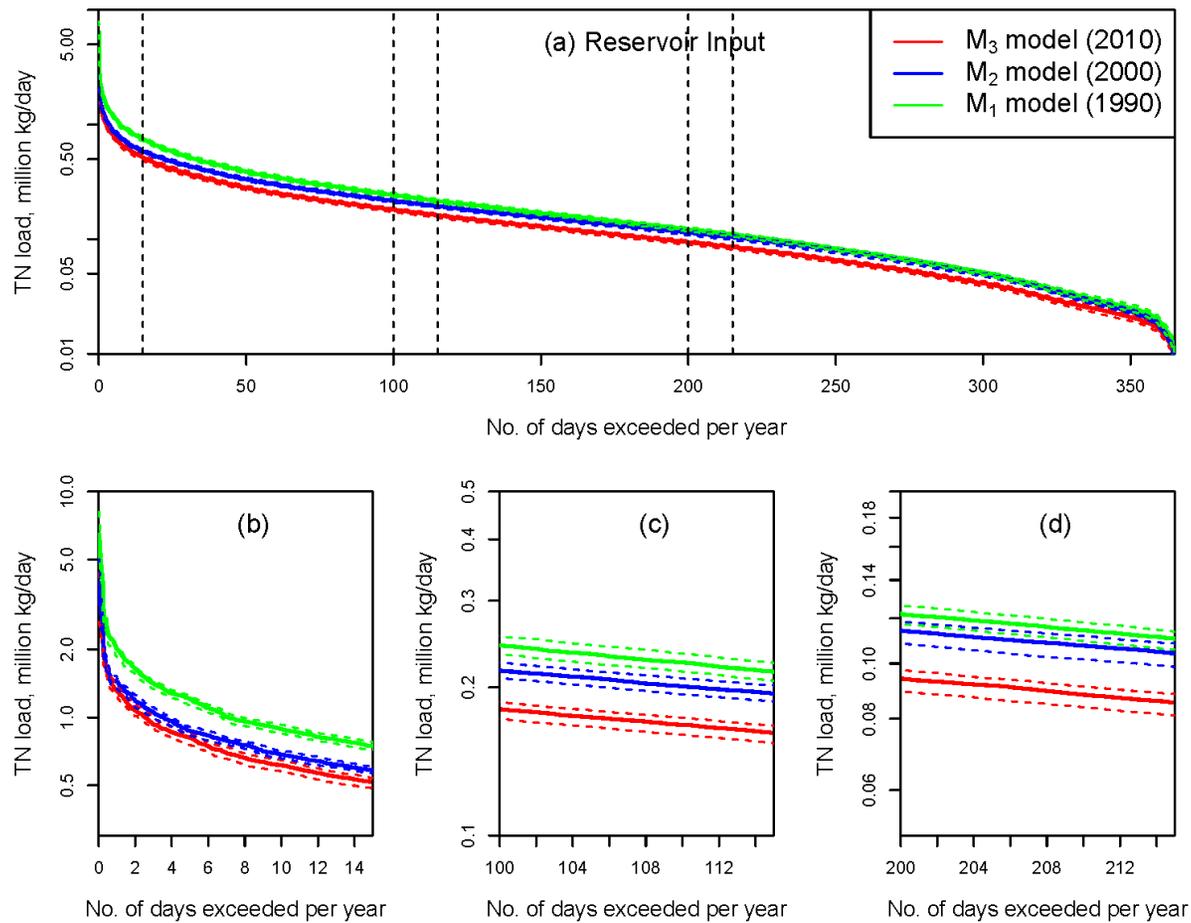


Figure B30. Frequency plots of ranked loadings for reservoir input loading of total nitrogen (TN) based on the three stationary WRTDS models. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Note that y-axis is on log-scale.

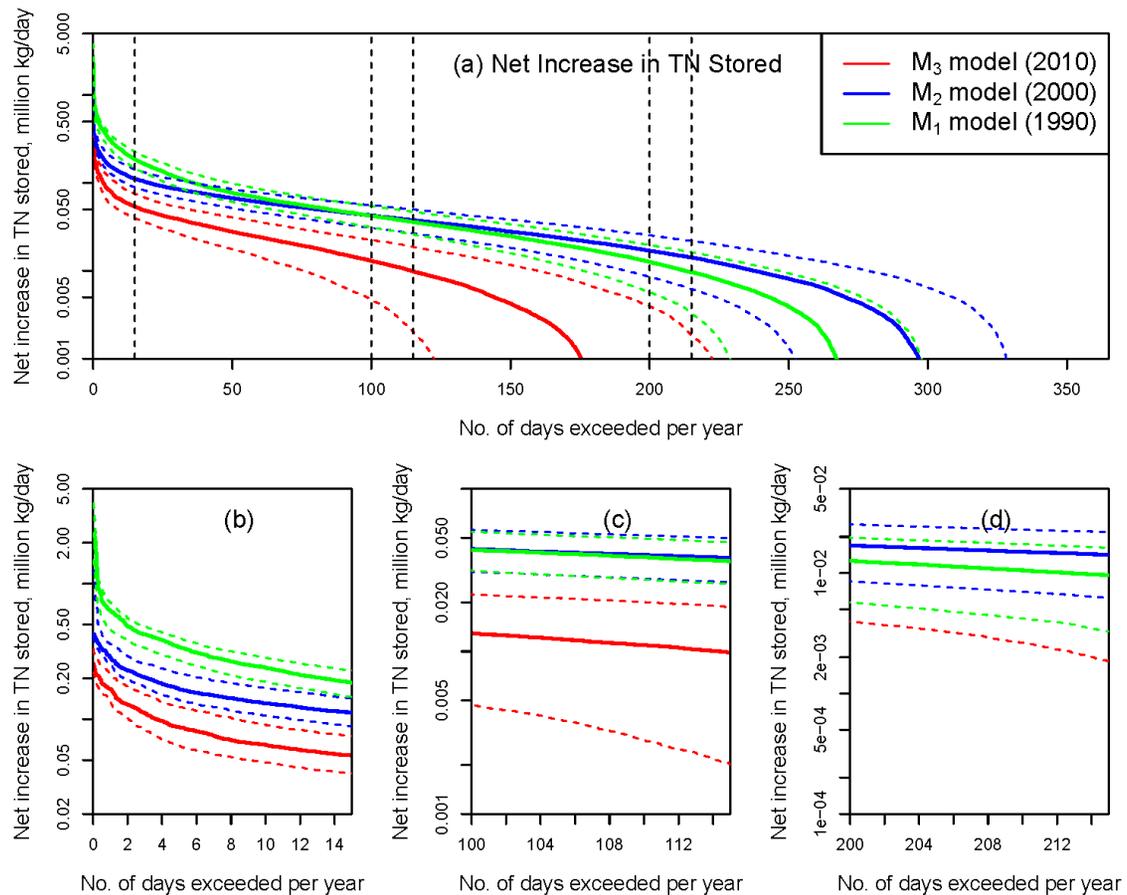


Figure B31. Frequency plots of ranked loadings for net increase in reservoir storage of total nitrogen (TN) based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. Note that y-axis is on log-scale.

## B5-II. Loading-discharge relationships for reservoir output, input, and net deposition based on stationary WRTDS models

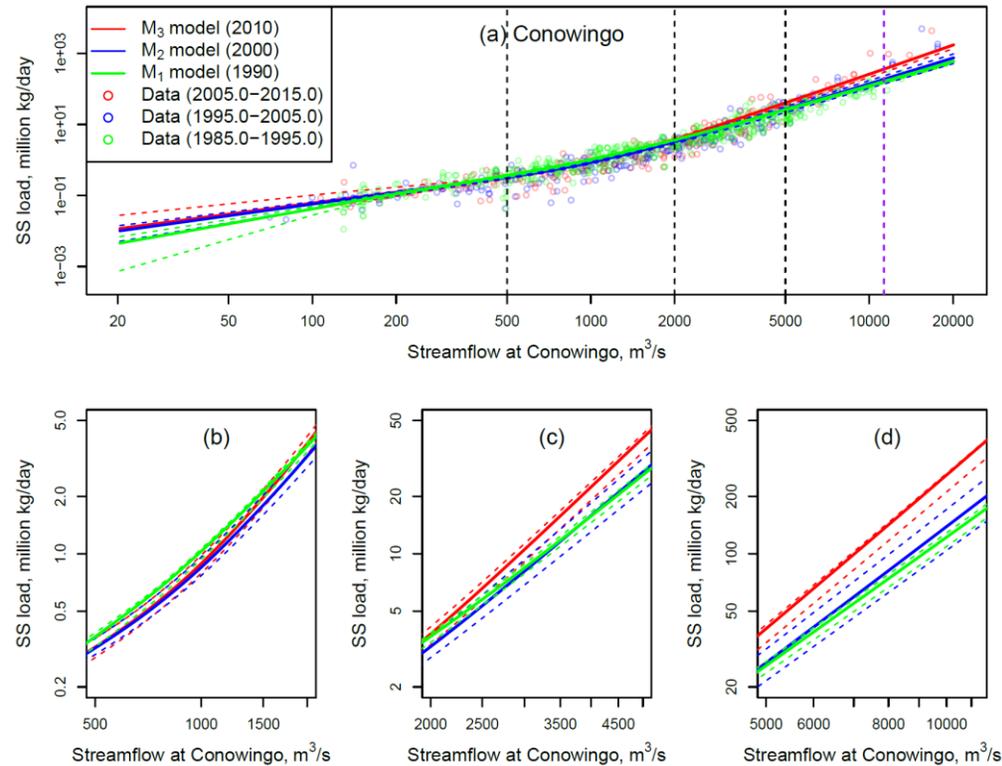


Figure B32. Loading-discharge relationships for reservoir output of suspended sediment (SS) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Plots (b)-(d) show results for three flow intervals. Observed loading data are plotted in (a) for reference purpose. Note that both axes are on log-scale and the curves were not obtained by fitting to the observed loadings. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

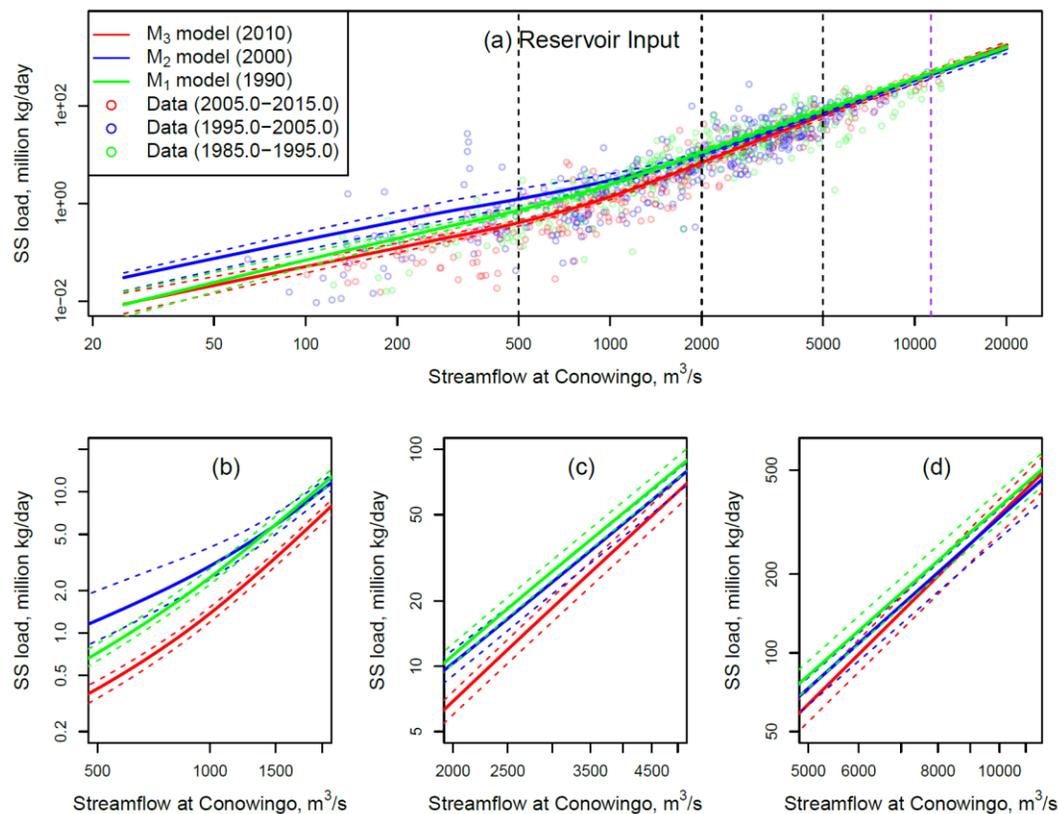


Figure B33. Loading-discharge relationships for reservoir input of suspended sediment (SS) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Plots (b)-(d) show results for three flow intervals. Observed loading data are plotted in (a) for reference purpose. Note that both axes are on log-scale and the curves were not obtained by fitting to the observed loadings. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

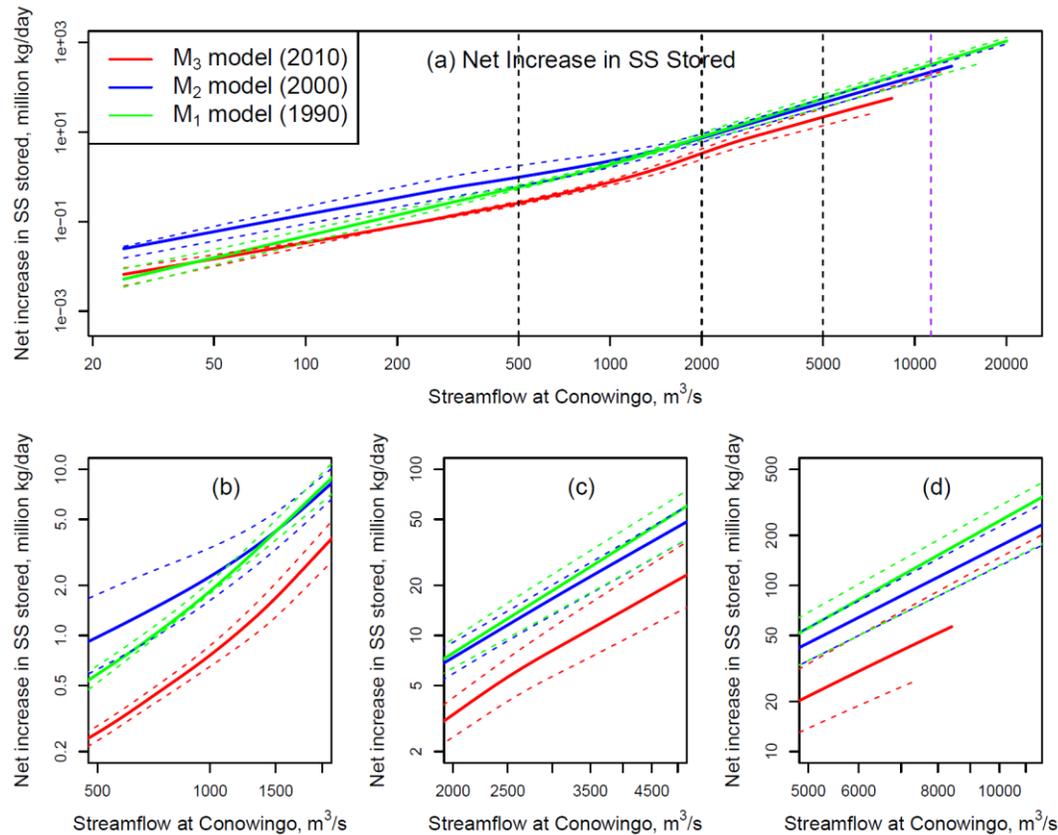


Figure B34. Loading-discharge relationships for net increase in reservoir storage of suspended sediment (SS) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. Plots (b)-(d) show results for three flow intervals. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

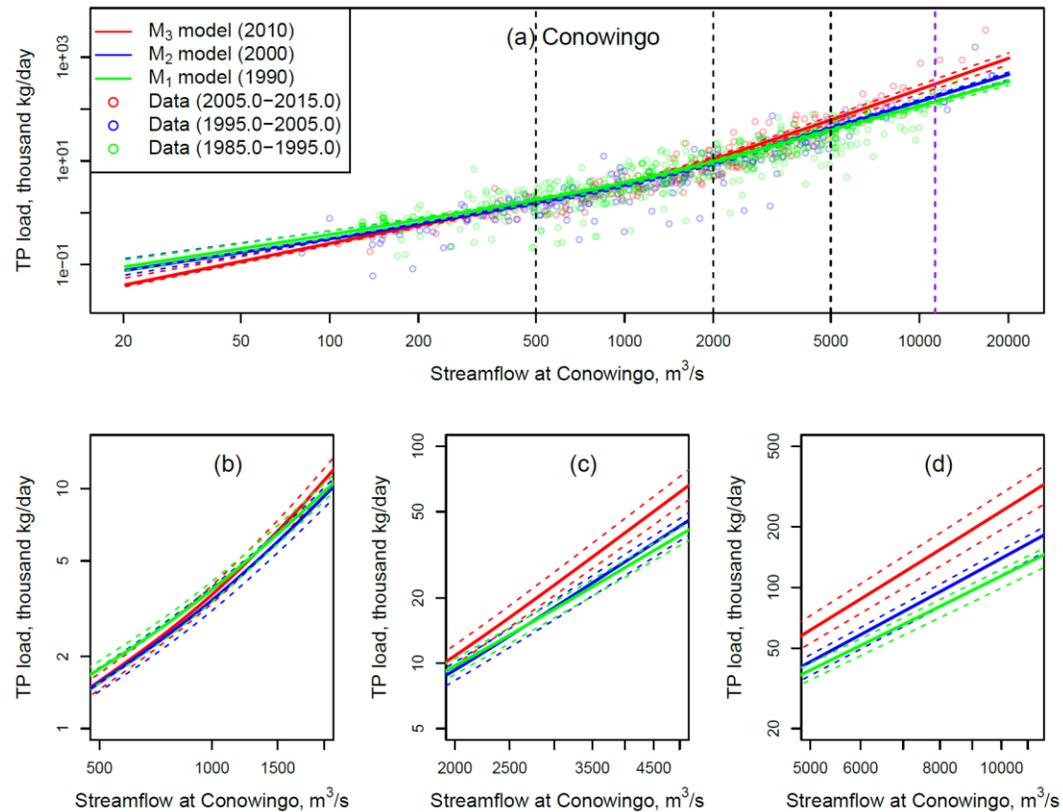


Figure B35. Loading-discharge relationships for reservoir output of total phosphorus (TP) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Plots (b)-(d) show results for three flow intervals. Observed loading data are plotted in (a) for reference purpose. Note that both axes are on log-scale and the curves were not obtained by fitting to the observed loadings. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

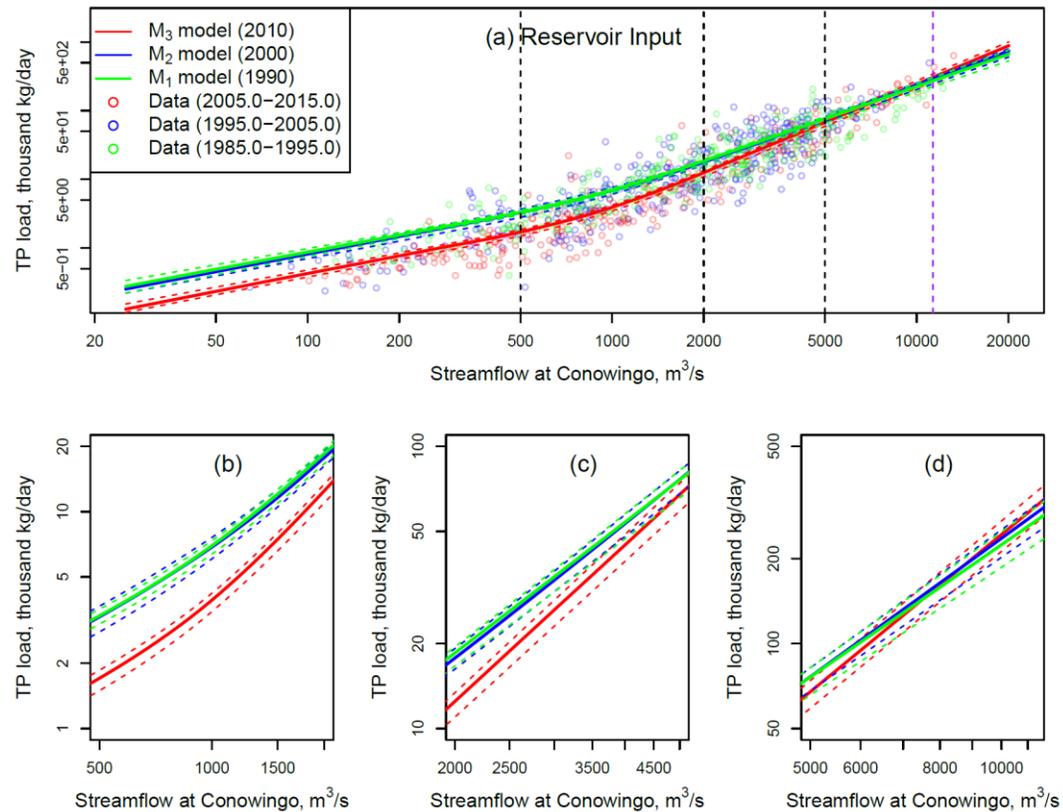


Figure B36. Loading-discharge relationships for reservoir input of total phosphorus (TP) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Plots (b)-(d) show results for three flow intervals. Observed loading data are plotted in (a) for reference purpose. Note that both axes are on log-scale and the curves were not obtained by fitting to the observed loadings. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

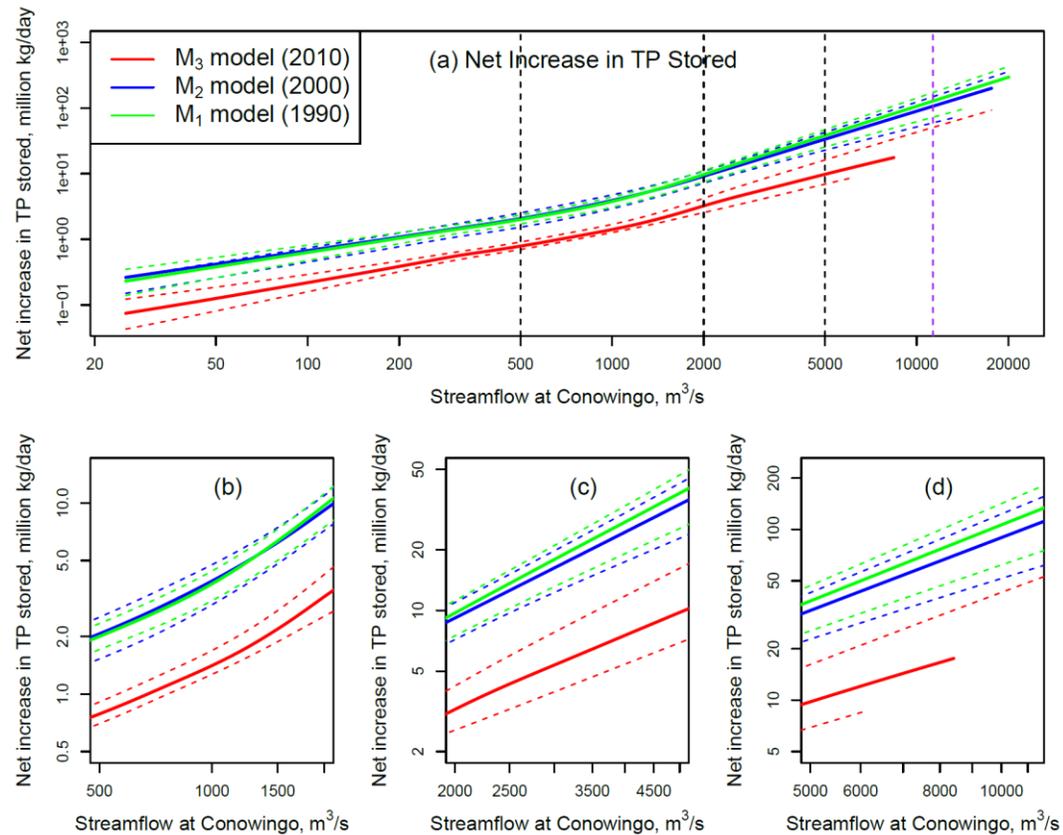


Figure B37. Loading-discharge relationships for net increase in reservoir storage of total phosphorus (TP) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. Plots (b)-(d) show results for three flow intervals. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

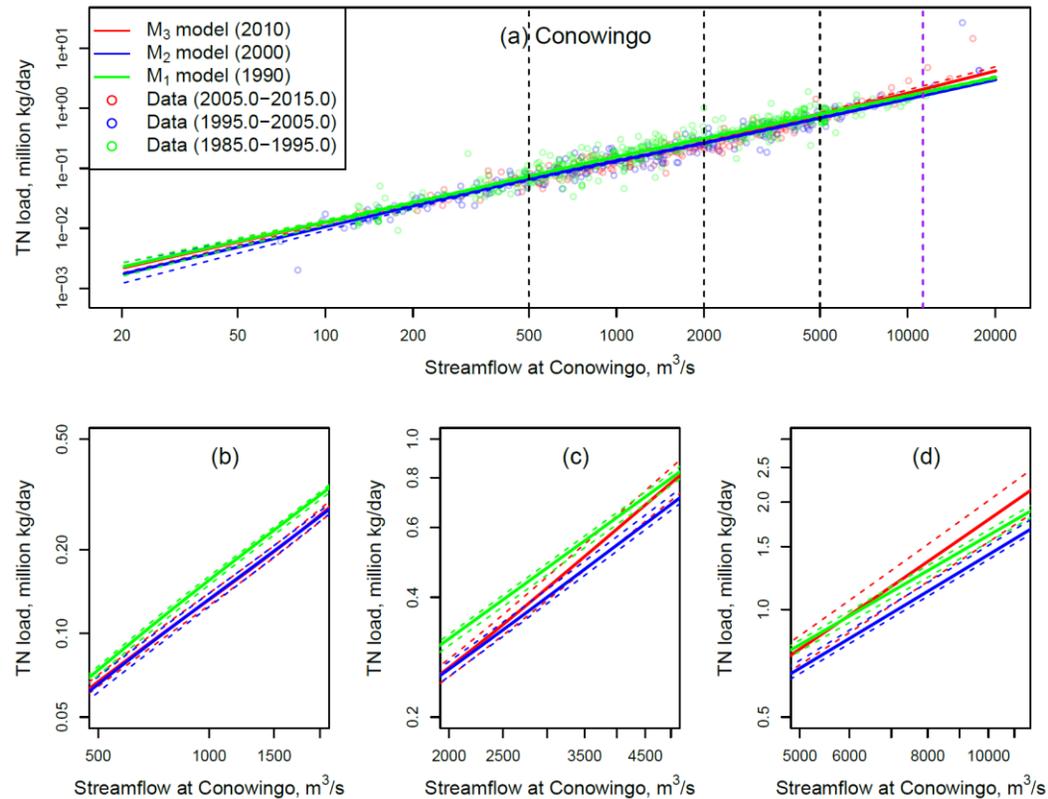


Figure B38. Loading-discharge relationships for reservoir output of total nitrogen (TN) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Plots (b)-(d) show results for three flow intervals. Observed loading data are plotted in (a) for reference purpose. Note that both axes are on log-scale and the curves were not obtained by fitting to the observed loadings. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

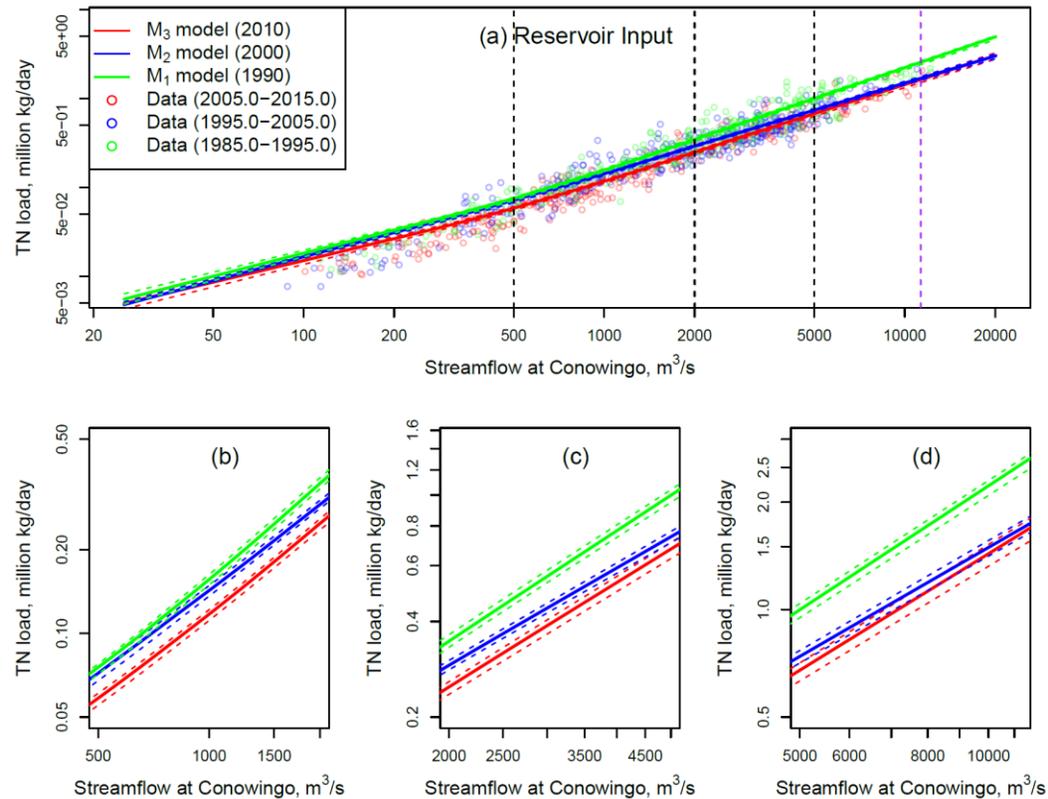


Figure B39. Loading-discharge relationships for reservoir input of total nitrogen (TN) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the upper and lower limits of model results derived from 10 replicates of each model that were based on resampled (with replacement) concentration data. Plots (b)-(d) show results for three flow intervals. Observed loading data are plotted in (a) for reference purpose. Note that both axes are on log-scale and the curves were not obtained by fitting to the observed loadings. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

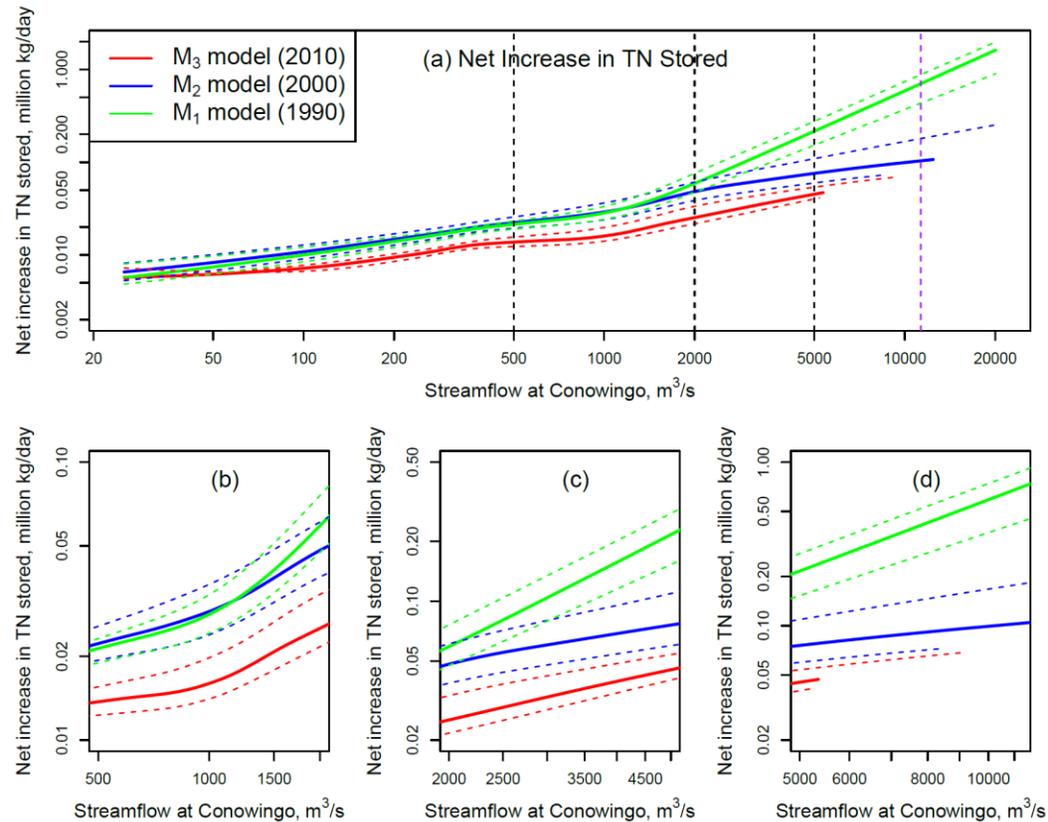


Figure B40. Loading-discharge relationships for net increase in reservoir storage of total nitrogen (TN) based on the three stationary WRTDS models. Lines shown are fitted LOWESS curves. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data. Plots (b)-(d) show results for three flow intervals. Vertical purple dashed line in (a) indicates the literature scour threshold of 11,300 m<sup>3</sup>/s.

### B5-III. Calendar-year cumulative loadings of reservoir output, input, and net storage based on stationary WRTDS models

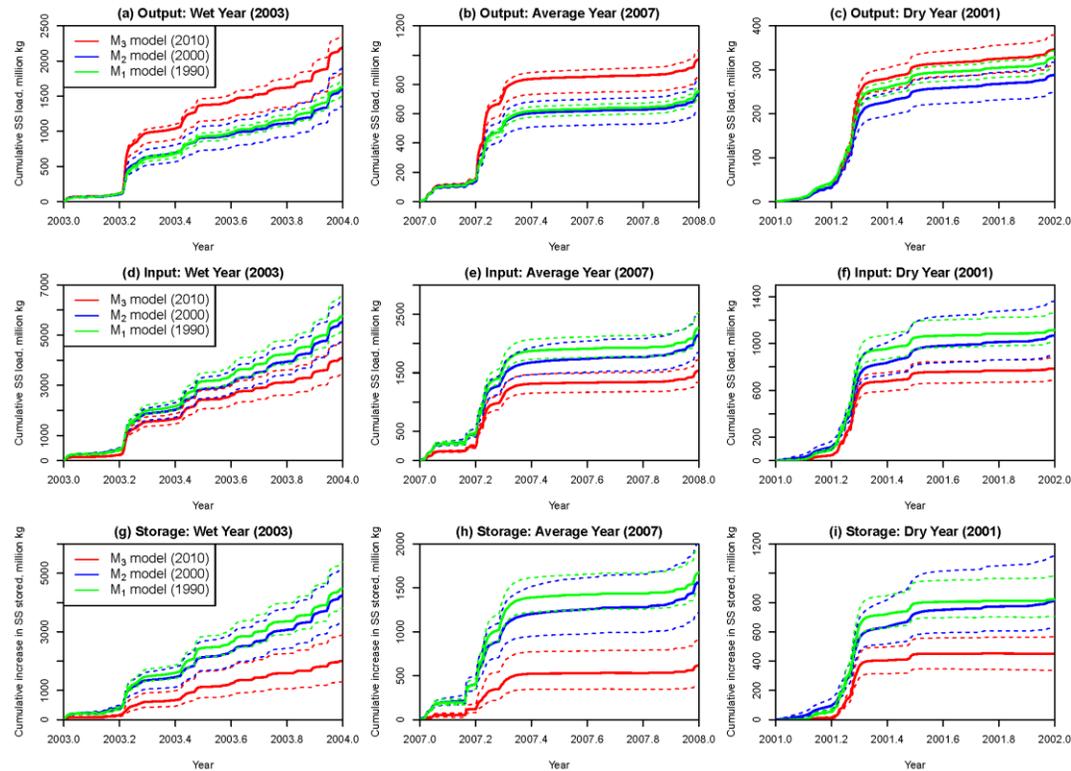


Figure B41. Modeled cumulative reservoir output, cumulative reservoir input, and cumulative reservoir storage of suspended sediment (SS) over the course of three selected wet, average, and dry calendar years (*i.e.*, 2003, 2007, and 2001) based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data.

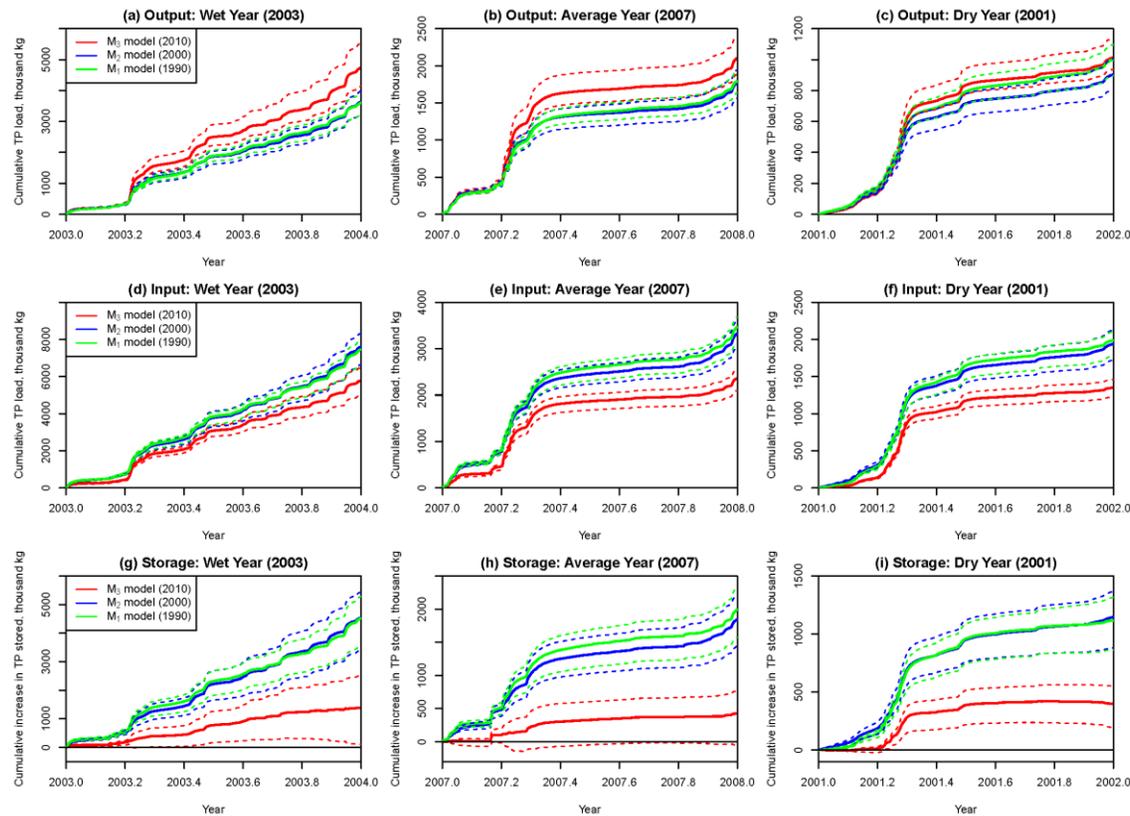


Figure B42. Modeled cumulative reservoir output, cumulative reservoir input, and cumulative reservoir storage of total phosphorus (TP) over the course of three selected wet, average, and dry calendar years (*i.e.*, 2003, 2007, and 2001) based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data.

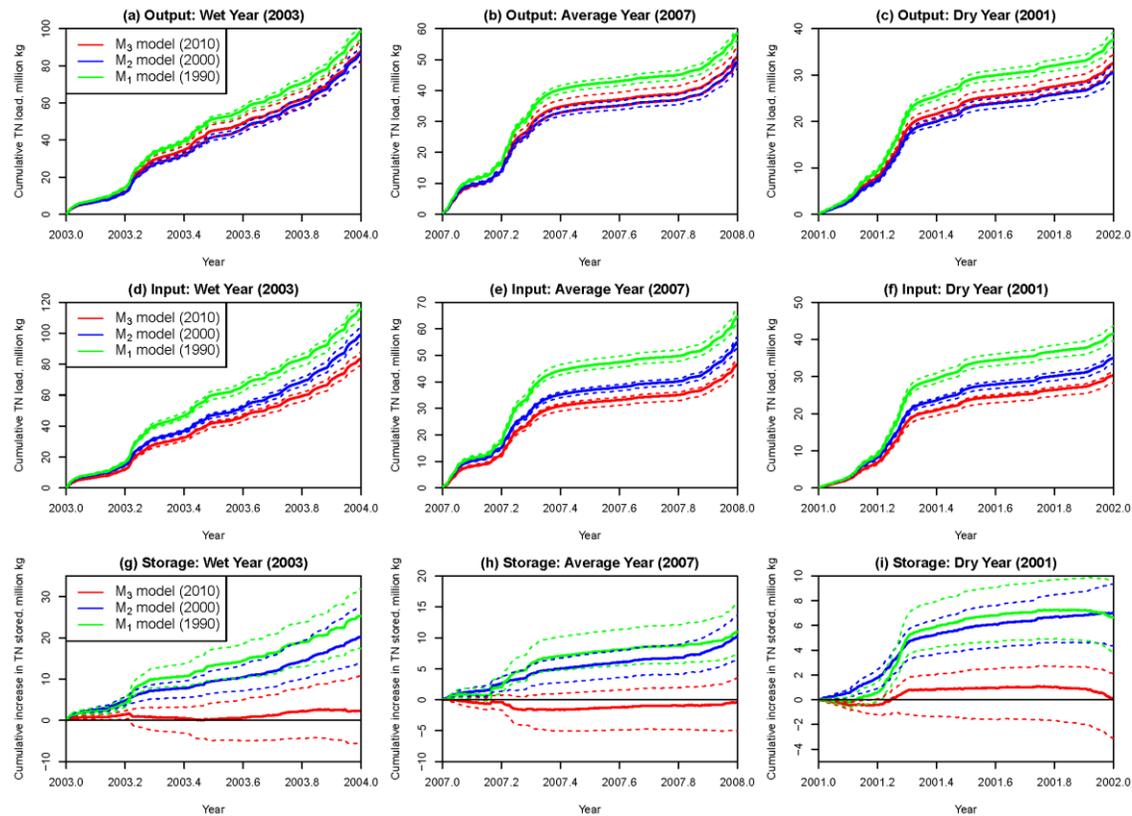


Figure B43. Modeled cumulative reservoir output, cumulative reservoir input, and cumulative reservoir storage of total nitrogen (TN) over the course of three selected wet, average, and dry calendar years (*i.e.*, 2003, 2007, and 2001) based on the three stationary WRTDS models. Dashed lines represent the 95% confidence intervals based on 100 sets of net storage estimates obtained from a 10x10 matrix created from 10 replicate runs of each model at both the inlet and outlet and based on random resampling with replacement of observed concentration data.

## Literature Cited

- Hainly, R. A., L. A. Reed, H. N. J. Flippo and G. J. Barton, 1995. Deposition and simulation of sediment transport in the Lower Susquehanna River reservoir system. U.S. Geological Survey U.S. Geological Survey Water-Resources Investigations Report 95-4122, Lemoyne, PA, p. 39. <http://pubs.er.usgs.gov/publication/wri954122>.
- Hirsch, R. M., S. A. Archfield and L. A. De Cicco, 2015. A bootstrap method for estimating uncertainty of water quality trends. *Journal of Environmental Modelling and Software* 73:148-166, DOI: 10.1016/j.envsoft.2015.07.017.
- Langland, M. J., 2009. Bathymetry and sediment-storage capacity change in three reservoirs on the Lower Susquehanna River, 1996-2008. U.S. Geological Survey Scientific Investigations Report 2009-5110, Reston, VA, p. 21. <http://pubs.usgs.gov/sir/2009/5110/>.
- Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18. <http://dx.doi.org/10.3133/ofr20141235>.
- Langland, M. J. and R. A. Hainly, 1997. Changes in bottom-surface elevations in three reservoirs on the lower Susquehanna River, Pennsylvania and Maryland, following the January 1996 flood - implications for nutrient and sediment loads to Chesapeake Bay. U.S. Geological Survey Water-Resources Investigations Report 97-4138, Lemoyne, PA, p. 34. <http://pa.water.usgs.gov/reports/wrir97-4138.pdf>.
- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Qian, S. S., 2010. *Environmental and Ecological Statistics with R*. Boca Raton, Florida, Chapman and Hall/CRC Press, ISBN 9781420062069.
- U.S. Geological Survey, 2014. Surface-water data for the nation.

**Appendix C. Supporting Information to Chapter 4**

**“Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the  
Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial  
Patterns”**

## Appendix C1. Default settings of WRTDS model runs

For this work, WRTDS was run using the *EGRET* (Exploration and Graphics for RivEr Trends) package version 2.2.0 (Hirsch and De Cicco, 2015). For each day of estimation, WRTDS pre-screens all concentration samples to select at least 100 samples that are sufficiently “close” to the estimation day. This proximity is evaluated with three distances, *i.e.*, time, discharge, and seasonal distances. Using the tri-cube weight function and the default half-window widths, these distances are converted to time, discharge, and seasonal weights, respectively, and their product is the total weight assigned to that sampled day. For all WRTDS model runs, we have used the default settings specified by the user guide (Hirsch and De Cicco, 2015), *i.e.*,  $\text{windowY} = 7$  (7 years for the time half-window),  $\text{windowS} = 0.5$  (0.5 year for the season half-window),  $\text{windowQ} = 2$  (2 natural log units for the discharge half-window), and  $\text{edgeAdjust} = \text{TRUE}$  (for alleviating edge effects near the start and end of the record). The selected samples are then used to fit the following equation:

$$\ln(C) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (\text{C1})$$

where  $C$  is concentration,  $Q$  is daily discharge,  $t$  is time in decimal years,  $\beta_i$  are fitted coefficients, and  $\varepsilon$  is the error term. The fitted coefficients in Equation (C1) are used to estimate the logarithmic  $C$  on the estimation day with known values of  $t$  and  $Q$ , which is then transformed back to concentration. Finally, this estimated daily concentration is multiplied by daily discharge to estimate the daily load.

### Literature Cited

Hirsch RM, De Cicco L. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). U.S. Geological Survey, Reston, VA, 2015, pp. 93. <http://dx.doi.org/10.3133/tm4A10>.

Table C1. Temporal coverage of observed water-quality data at the seven long-term monitoring sites in the Susquehanna River basin. <sup>a</sup>

Sites	Suspended Sediment (SS)			Total Phosphorus (TP)			Dissolved Phosphorus (DP)			Total Nitrogen (TN)			Dissolved Nitrogen (DN)		
	T <sub>Start</sub>	T <sub>End</sub>	#Sampled (f <sub>Sampling</sub> )	T <sub>Start</sub>	T <sub>End</sub>	#Sampled (f <sub>Sampling</sub> )	T <sub>Start</sub>	T <sub>End</sub>	#Sampled (f <sub>Sampling</sub> )	T <sub>Start</sub>	T <sub>End</sub>	#Sampled (f <sub>Sampling</sub> )	T <sub>Start</sub>	T <sub>End</sub>	#Sampled (f <sub>Sampling</sub> )
Susquehanna River at Towanda, PA	1988-10-05	2014-09-15	990 (38.1/yr)	1984-10-10	2014-09-15	1061 (35.4/yr)	1985-12-10	2014-09-30	895 (30.9/yr)	1984-10-10	2014-09-15	931 (31.0/yr)	1985-12-10	2014-09-30	893 (30.8/yr)
Susquehanna River at Danville, PA	1984-10-11	2014-09-30	1183 (39.4/yr)	1984-10-11	2014-09-30	1213 (40.4/yr)	1984-10-22	2014-09-30	1002 (33.4/yr)	1984-10-11	2014-09-30	1082 (36.1/yr)	1985-12-10	2014-09-30	991 (34.2/yr)
West Branch Susquehanna River at Lewisburg, PA	1984-10-11	2014-09-30	1130 (37.7/yr)	1984-10-11	2014-09-30	1165 (38.8/yr)	1984-10-22	2014-09-30	990 (33.0/yr)	1984-10-11	2014-09-30	1031 (34.4/yr)	1985-12-10	2014-09-30	977 (33.7/yr)
Juniata River at Newport, PA	1984-10-10	2014-09-17	1006 (33.5/yr)	1984-10-10	2014-09-17	1050 (35.0/yr)	1984-10-22	2014-09-29	847 (28.2/yr)	1984-10-10	2014-09-17	914 (30.5/yr)	1985-12-11	2014-09-29	832 (28.7/yr)
Susquehanna River at Marietta, PA	1986-10-07	2014-09-29	1037 (37.0/yr)	1986-10-07	2014-09-29	1044 (37.3/yr)	1986-10-07	2014-09-29	957 (34.2/yr)	1986-10-07	2014-09-29	1043 (37.3/yr)	1986-10-07	2014-09-29	946 (33.8/yr)
Conestoga River at Conestoga, PA	1984-10-18	2014-09-29	1076 (35.9/yr)	1984-10-18	2014-09-29	1065 (35.5/yr)	1984-10-22	2014-09-29	904 (30.1/yr)	1984-10-18	2014-09-29	1014 (33.8/yr)	1985-12-10	2014-09-29	884 (30.5/yr)
Susquehanna River near Conowingo, MD	1984-10-25	2014-09-03	799 (26.6/yr)	1984-10-25	2014-09-03	808 (26.9/yr)	1984-10-25	2014-09-03	784 (26.1/yr)	1984-10-25	2014-09-03	804 (26.8/yr)	1984-10-25	2014-08-05	766 (25.5/yr)

<sup>a</sup> T<sub>Start</sub>: the first sampled day; T<sub>End</sub>: the last sampled day; #<sub>Sampled</sub>: the total number of sampled days; f<sub>Sampling</sub>: the average number of sampled days per year.

Table C2. Summary of long-term median values of flow-normalized annual loads and yields at the seven Susquehanna sites. <sup>a</sup>

Sites	Drainage area (km <sup>2</sup> )	Long-term median of annual load							Long-term median of annual yield							
		SS	TP	DP	PP	TN	DN	PN	SS	TP	DP	PP	TN	DN	PN	
		(kg/day)	(kg P/day)			(kg N/day)				(kg km <sup>-2</sup> yr <sup>-1</sup> )	(kg P km <sup>-2</sup> yr <sup>-1</sup> )	(kg N km <sup>-2</sup> yr <sup>-1</sup> )				
Towanda	20,194	3.0 x 10 <sup>6</sup>	3.0 x 10 <sup>3</sup>	1.0 x 10 <sup>3</sup>	1.9 x 10 <sup>3</sup>	3.4 x 10 <sup>4</sup>	3.0 x 10 <sup>4</sup>	3.6 x 10 <sup>3</sup>	5.4 x 10 <sup>4</sup>	54	18	35	610	541	66	
Danville	29,060	4.4 x 10 <sup>6</sup>	5.0 x 10 <sup>3</sup>	1.3 x 10 <sup>3</sup>	3.7 x 10 <sup>3</sup>	5.2 x 10 <sup>4</sup>	4.6 x 10 <sup>4</sup>	7.5 x 10 <sup>3</sup>	5.5 x 10 <sup>4</sup>	63	16	46	660	573	95	
Lewisburg	17,734	1.3 x 10 <sup>6</sup>	1.7 x 10 <sup>3</sup>	5.5 x 10 <sup>2</sup>	1.1 x 10 <sup>3</sup>	2.8 x 10 <sup>4</sup>	2.5 x 10 <sup>4</sup>	3.2 x 10 <sup>3</sup>	2.8 x 10 <sup>4</sup>	35	11	22	583	517	66	
Newport	8,687	6.3 x 10 <sup>5</sup>	9.4 x 10 <sup>2</sup>	4.5 x 10 <sup>2</sup>	5.2 x 10 <sup>2</sup>	1.9 x 10 <sup>4</sup>	1.8 x 10 <sup>4</sup>	1.9 x 10 <sup>3</sup>	2.6 x 10 <sup>4</sup>	40	19	22	818	739	79	
Marietta	67,314	8.4 x 10 <sup>6</sup>	1.0 x 10 <sup>4</sup>	2.9 x 10 <sup>3</sup>	7.5 x 10 <sup>3</sup>	1.6 x 10 <sup>5</sup>	1.4 x 10 <sup>5</sup>	2.2 x 10 <sup>4</sup>	4.5 x 10 <sup>4</sup>	56	16	40	879	760	118	
Conestoga	1,217	3.4 x 10 <sup>5</sup>	7.7 x 10 <sup>2</sup>	3.2 x 10 <sup>2</sup>	4.4 x 10 <sup>2</sup>	1.3 x 10 <sup>4</sup>	1.2 x 10 <sup>4</sup>	9.7 x 10 <sup>2</sup>	1.0 x 10 <sup>5</sup>	231	96	132	3999	3683	292	
Conowingo	70,189	4.2 x 10 <sup>6</sup>	6.8 x 10 <sup>3</sup>	1.6 x 10 <sup>3</sup>	4.8 x 10 <sup>3</sup>	1.7 x 10 <sup>5</sup>	1.4 x 10 <sup>5</sup>	2.5 x 10 <sup>4</sup>	2.2 x 10 <sup>4</sup>	36	8	25	875	729	129	

<sup>a</sup> study period varies with species; see Table C1 for data coverage for each species at each site.

Table C3. Period-of-record averages of annual riverine conditions (river flow, concentration, and yield) and annual source inputs (atmospheric deposition, fertilizer, manure, point source, and total) at the seven Susquehanna sites for the period between 1987 and 2011. <sup>a</sup>

Sites	Period-average Riverine Conditions					Period-average Source Input Yields <sup>b</sup> (kg km <sup>-2</sup> yr <sup>-1</sup> )				
	Flow Yield, (m yr <sup>-1</sup> )	Conc., (mg L <sup>-1</sup> )	FN Conc. <sup>c</sup> (mg L <sup>-1</sup> )	Yield <sup>d</sup> (kg km <sup>-2</sup> yr <sup>-1</sup> )	FN Yield <sup>e</sup> (kg km <sup>-2</sup> yr <sup>-1</sup> )	Atmospheric Deposition	Fertilizer	Manure	Point Source	Total Input
<b>Total Phosphorus (TP)</b>										
Towanda	0.52	0.078	0.077	56	54	NA <sup>f</sup>	93	207	15	316
Danville	0.53	0.083	0.082	65	62	NA	102	199	22	326
Lewisburg	0.56	0.050	0.049	36	33	NA	69	97	8	175
Newport	0.47	0.068	0.067	42	40	NA	162	338	15	516
Marietta	0.53	0.074	0.074	58	56	NA	132	255	20	409
Conestoga	0.52	0.294	0.294	208	215	NA	233	2283	69	2587
Conowingo	0.53	0.047	0.046	42	36	NA	140	335	21	498
<b>Total Nitrogen (TN)</b>										
Towanda	0.52	1.11	1.12	626	624	1164	1172	860	115	3336
Danville	0.53	1.16	1.16	674	666	1221	1299	826	137	3523
Lewisburg	0.56	1.03	1.04	585	579	1256	676	384	63	2400
Newport	0.47	1.60	1.59	843	822	1293	1503	1349	80	4254
Marietta	0.53	1.53	1.54	879	879	1283	1381	995	136	3831
Conestoga	0.52	7.84	7.86	3918	3907	2142	4282	8146	495	15151
Conowingo	0.53	1.63	1.62	926	888	1316	1510	1278	142	4285

<sup>a</sup> The average condition is calculated from annual averages for the period of 1987-2011, which is the longest period that has data at all sites.

Annual average values are calculated as the average of daily estimates.

<sup>b</sup> Source Input Yield: source input load per unit of basin area.

<sup>c</sup> FN Conc.: flow-normalized concentration.

<sup>d</sup> Yield: true-condition riverine yield, *i.e.*, true-condition load per unit of basin area.

<sup>e</sup> FN Yield: flow-normalized riverine yield, *i.e.*, flow-normalized load per unit of basin area.

<sup>f</sup> NA: not applicable; atmospheric deposition is not considered a source input for total phosphorus.

Table C4. Initial conditions (year 1987), final conditions (year 2011), and period-of-record (1987 to 2011 <sup>a</sup>) changes in flow-normalized annual riverine yield <sup>b</sup> and annual source input yields <sup>c</sup> (*i.e.*, atmospheric deposition, fertilizer, manure, point source, and total) for the seven Susquehanna sites.

Sites	Value	Total Phosphorus, TP (kg km <sup>-2</sup> )					Total Nitrogen, TN (kg km <sup>-2</sup> )					
		FN Yield	Fertilizer	Manure	Point Source	Total Input <sup>c</sup>	FN Yield	Atmospheric Deposition	Fertilizer	Manure	Point Source	Total Input
Towanda	1987	58	108	247	18	373	839	1365	1299	1042	121	3846
	2011	52	75	189	10	275	486	1010	972	762	74	2843
	$\Delta^d$	-6	-33	-57	-8	-99	-353	-356	-327	-280	-47	-1003
Danville	1987	67	117	236	30	387	862	1455	1428	1003	147	4068
	2011	56	74	172	16	265	529	1061	1029	687	106	2923
	$\Delta$	-11	-44	-64	-14	-122	-333	-394	-400	-315	-41	-1146
Lewisburg	1987	38	73	97	12	182	676	1542	717	390	53	2718
	2011	23	53	92	7	153	457	992	573	359	64	2008
	$\Delta$	-15	-20	-5	-5	-29	-219	-550	-144	-30	+11	-710
Newport	1987	50	166	293	17	476	927	1571	1526	1204	70	4392
	2011	32	136	377	13	527	750	1020	1322	1470	83	3923
	$\Delta$	-18	-29	+83	-4	+51	-177	-552	-204	+266	+13	-469
Marietta	1987	63	146	249	26	423	1063	1533	1496	1000	130	4189
	2011	48	106	260	16	383	735	1086	1164	990	119	3396
	$\Delta$	-14	-40	+10	-11	-39	-328	-447	-331	-10	-11	-794
Conestoga	1987	295	216	2418	110	2745	4378	2277	4409	8261	717	15737
	2011	149	238	2145	56	2441	3300	2191	4064	8008	415	14773
	$\Delta$	-146	+23	-273	-53	-304	-1078	-87	-345	-253	-301	-964
Conowingo	1987	34	152	333	28	515	1021	1562	1626	1283	140	4643
	2011	50	116	336	16	470	868	1128	1297	1274	125	3863
	$\Delta$	+16	-36	+3	-12	-45	-153	-434	-329	-9	-15	-780

<sup>a</sup> The period of 1987-2011 is the longest period that has data at all sites.

<sup>b</sup> FN yield: flow-normalized riverine yield, *i.e.*, flow-normalized load per unit of basin area.

<sup>c</sup> Source input yield: source input load per unit of basin area.

<sup>d</sup>  $\Delta$  denotes period-of-record change in yield, which is calculated as:  $\Delta = (\text{yield in 2011}) - (\text{yield in 1987})$ . Note that negative values (*i.e.*, declines between 1987 and 2011) are highlighted in green, whereas positive values are highlighted in red.

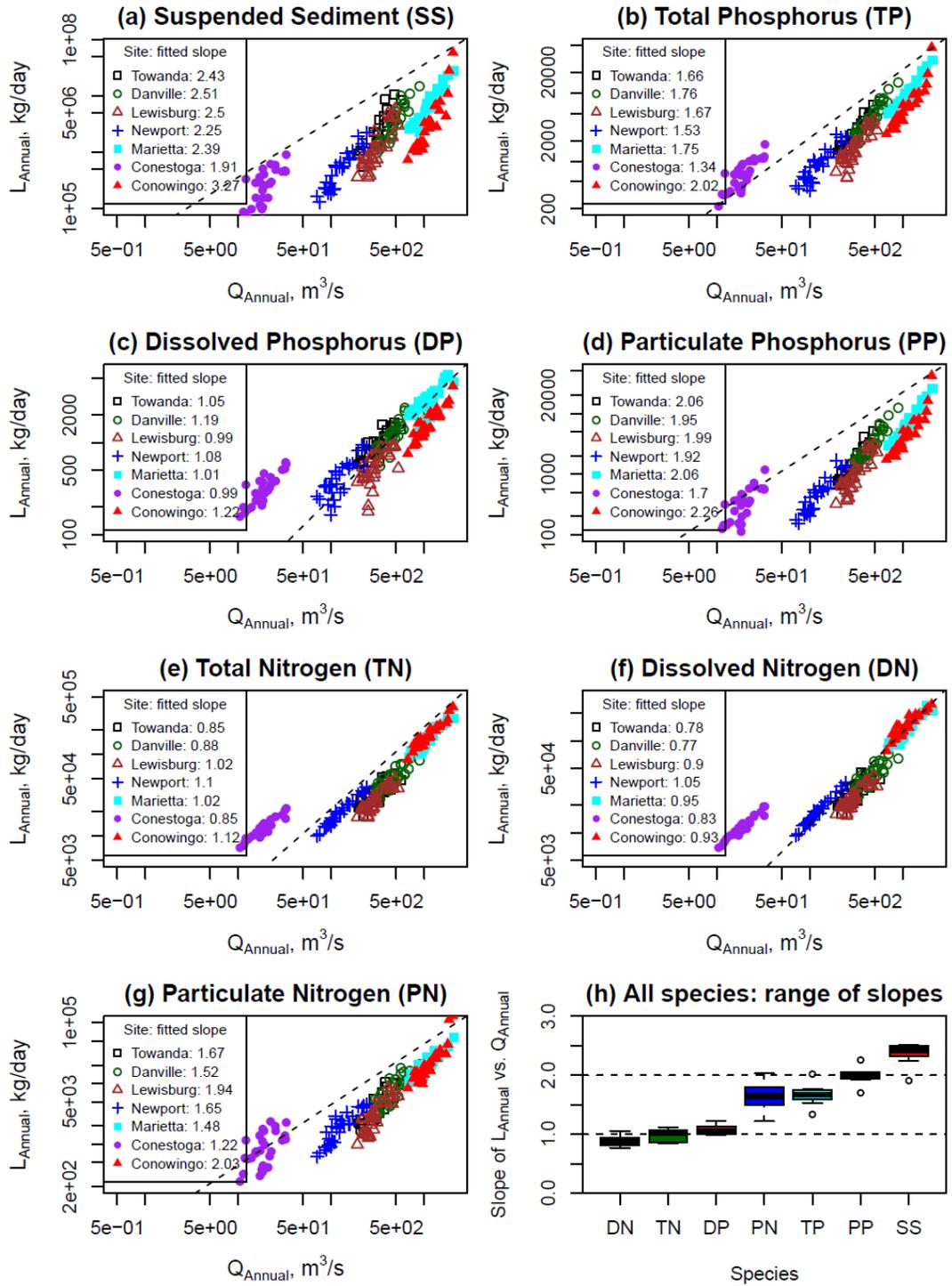


Figure C1. Relations between annual discharge ( $Q_{\text{Annual}}$ ) and annual loadings ( $L_{\text{Annual}}$ ) of (a) TN, (b) DN, (c) PN, (d) TP, (e) DP, (f) PP, and (g) SS at the seven Susquehanna sites. Dashed lines in (a)-(g) have the reference slope of 1.0 on log-log scale. Inset shows linear regression slopes for each site, which are summarized with boxplots in (h).

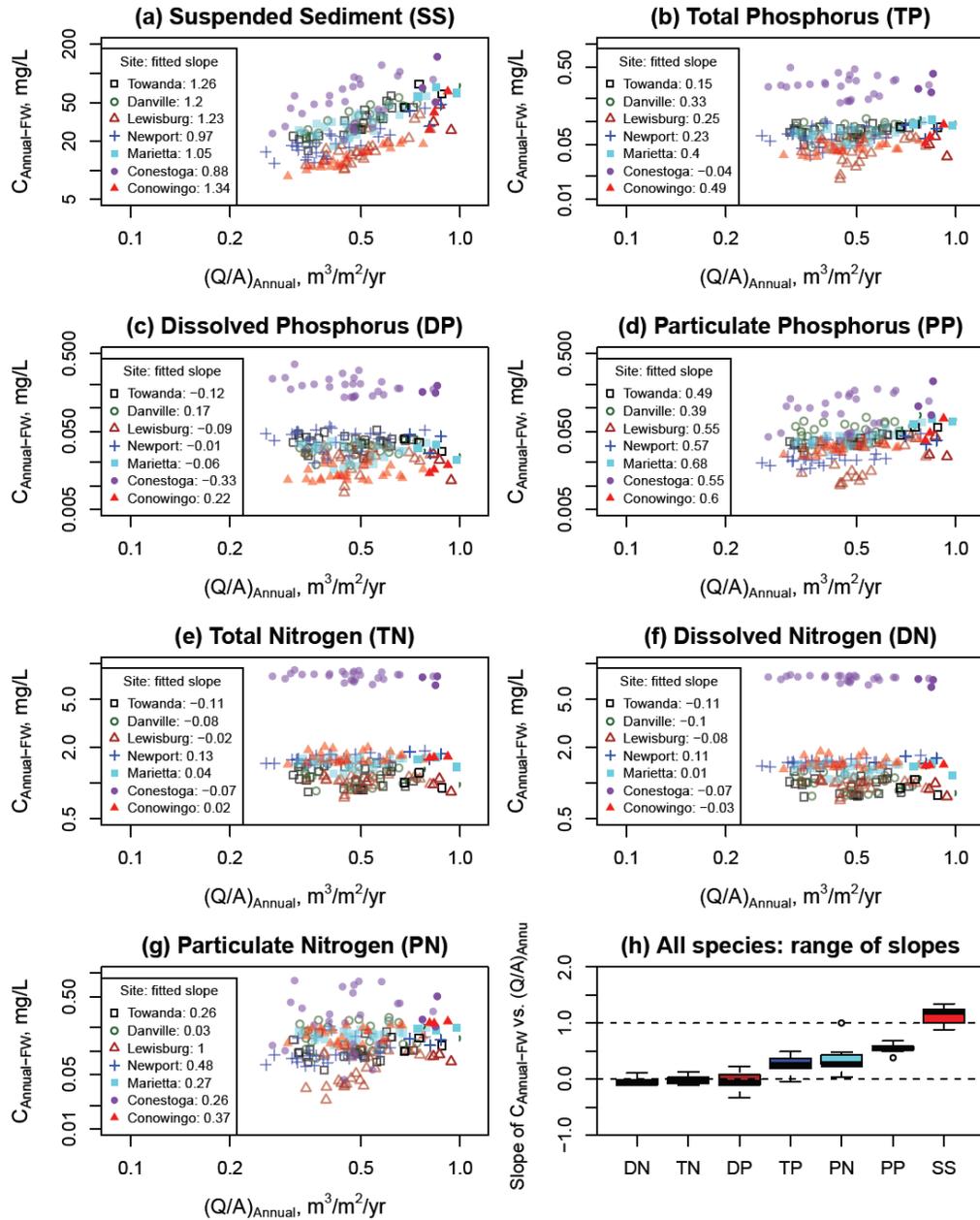


Figure C2. Relations between area-normalized annual discharge ( $[Q/A]_{\text{Annual}}$ ) and annual flow-weighted concentration ( $C_{\text{Annual-FW}}$ ) for seven species: (a) TN, (b) DN, (c) PN, (d) TP, (e) DP, (f) PP, and (g) SS at the seven Susquehanna sites. Inset shows linear regression slopes for each site. These slopes are summarized for each species with boxplots in (h). For the study period of 1985-2013, three years are associated with extreme-flow events, *i.e.*, 2011 (Tropical Storm Lee), 2004 (Hurricane Ivan), and 1996 (ice-jam release in the Lower Susquehanna River Reservoir System). These years are located at the far right side in (a)-(g) and are shown with non-transparent colors.

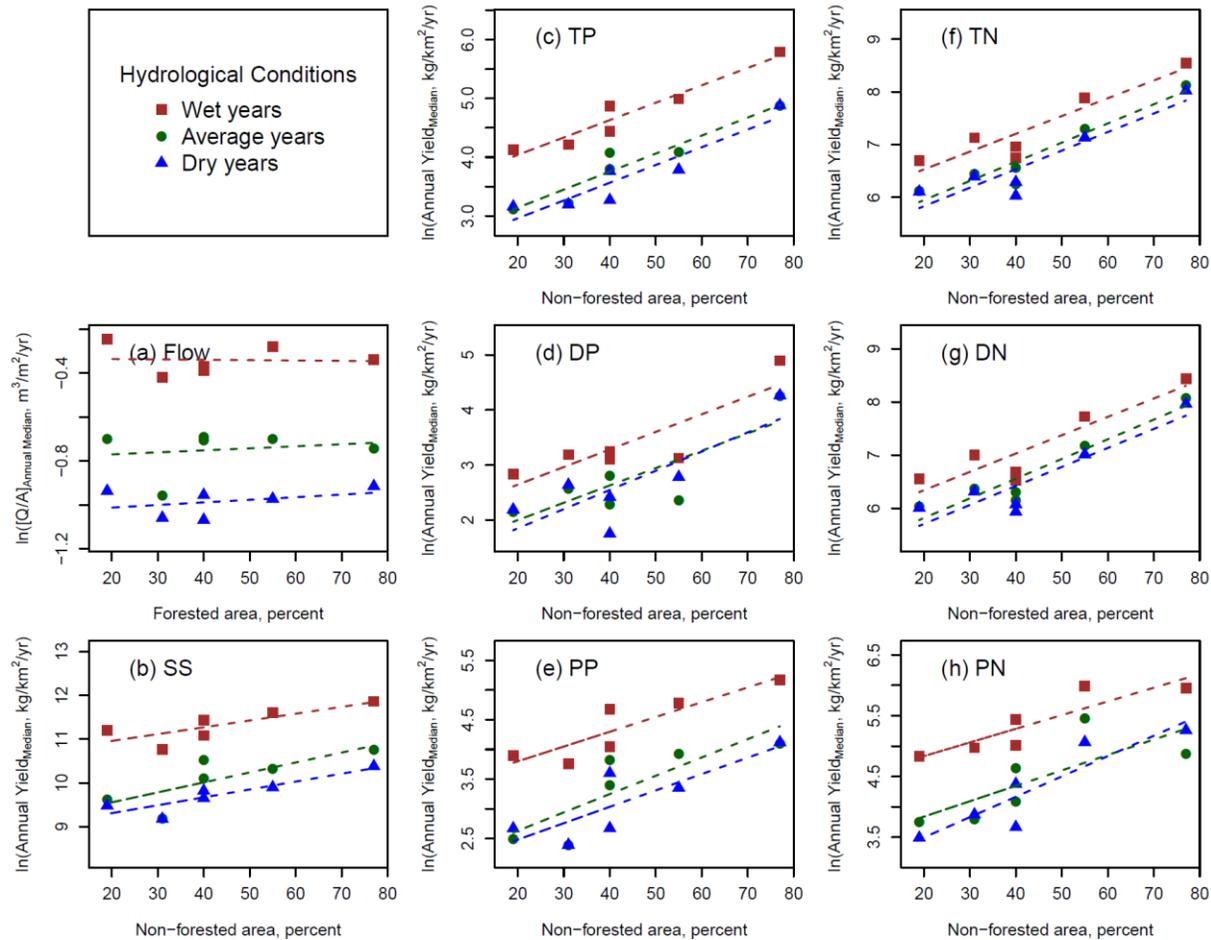


Figure C3. Relations between area fraction of *non-forested land* and log-transformed period-of-record medians of (a) area-normalized annual discharge ( $[Q/A]_{\text{Annual Median}}$ ) and annual *true-condition* loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN in the seven Susquehanna sub-basins under different flow conditions. Each point represents one sub-basin. Dashed lines are linear fits.

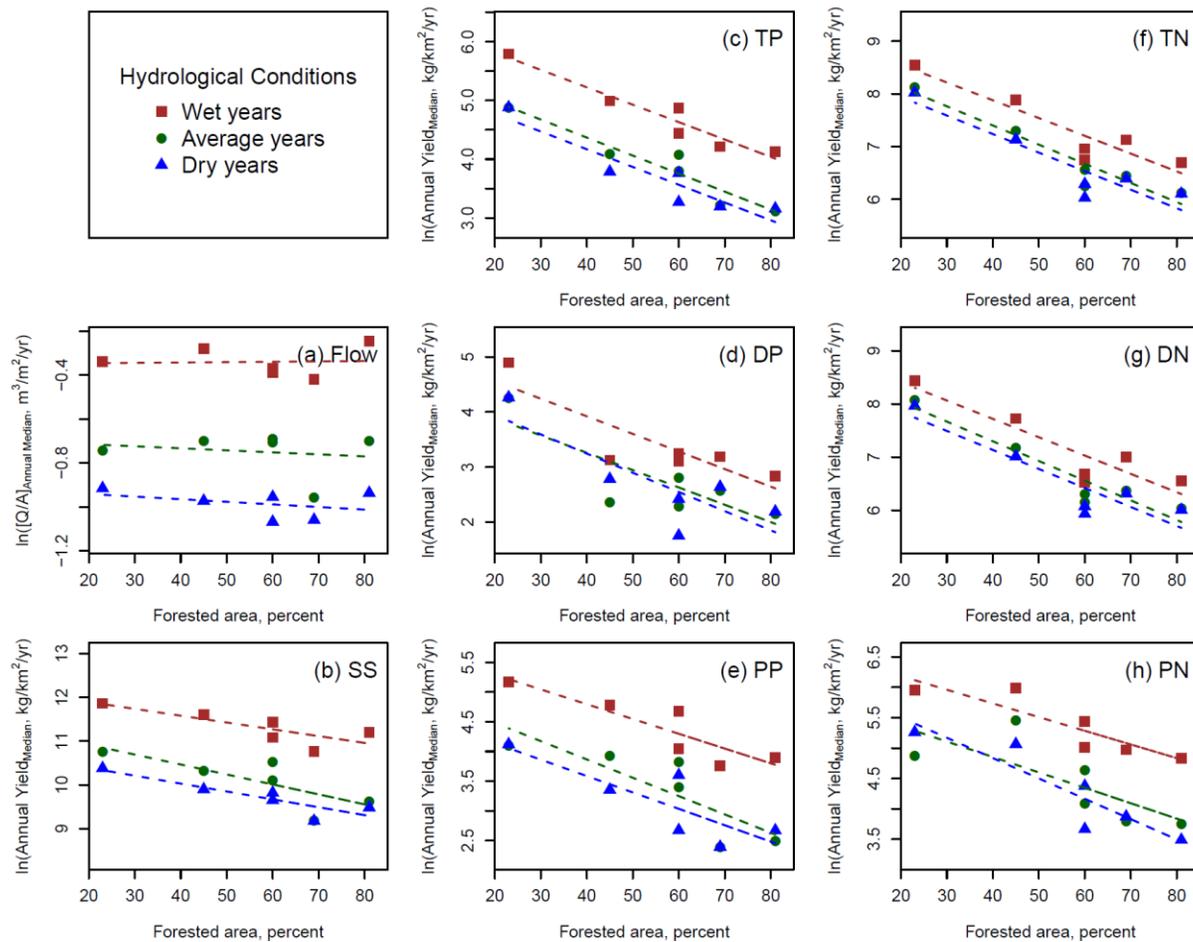


Figure C4. Relations between area fraction of *forested land* and log-transformed period-of-record medians of (a) area-normalized annual discharge ( $[Q/A]_{\text{Annual Median}}$ ) and annual *true-condition* loadings of (b) SS, (c) TP, (d) DP, (e) PP, (f) TN, (g) DN, and (h) PN in the seven Susquehanna sub-basins under different flow conditions. Each point represents one sub-basin. Dashed lines are linear fits.

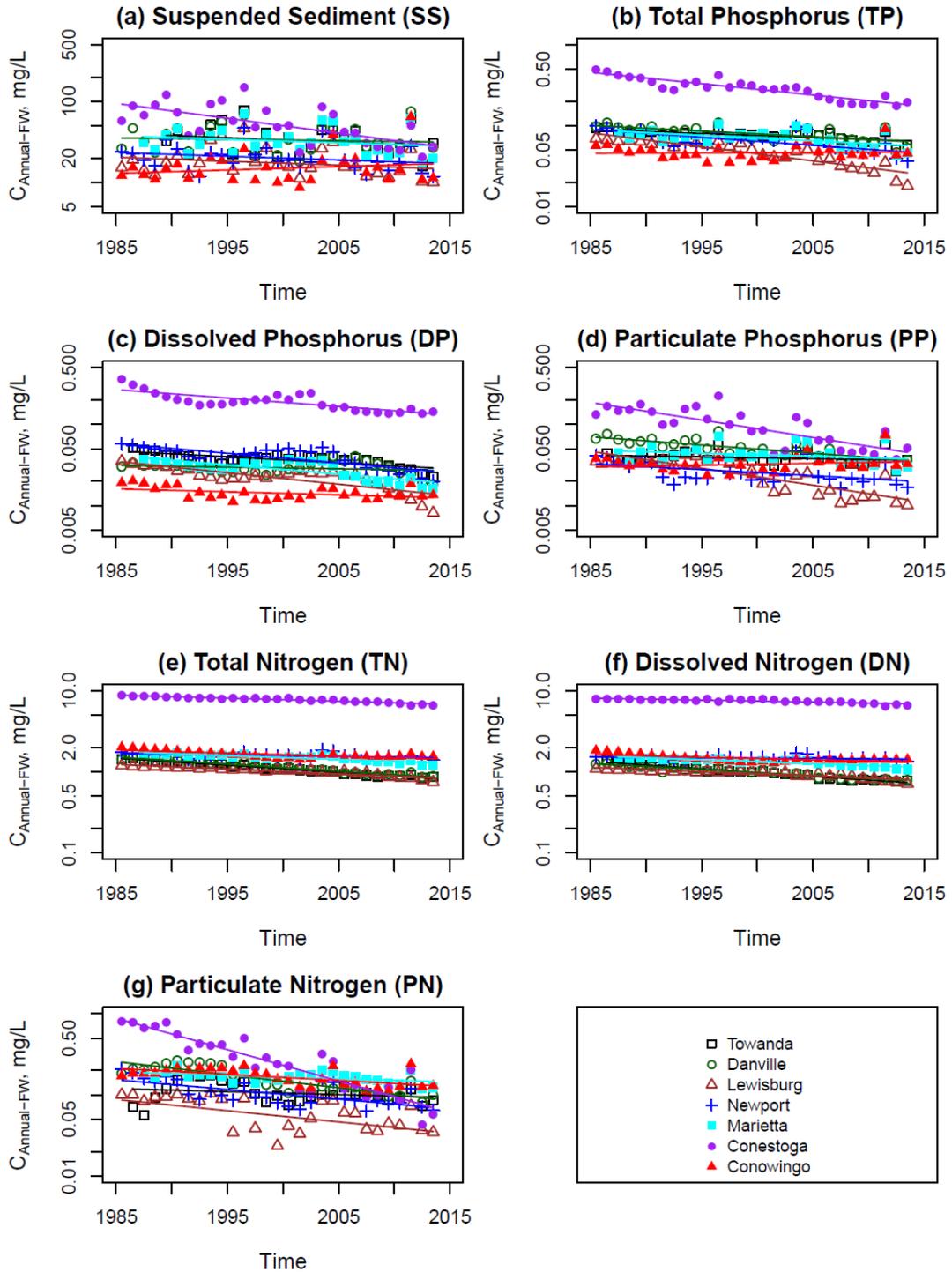


Figure C5. Time series of annual flow-weighted concentration ( $C_{\text{Annual-FW}}$ ) for seven species: (a) TN, (b) DN, (c) PN, (d) TP, (e) DP, (f) PP, and (g) SS at the seven Susquehanna sites. Despite a few exceptions,  $C_{\text{Annual-FW}}$  does not vary significantly over time. Solid lines are linear fits.

*Page intentionally left blank*

**Appendix D. Supporting Information to Chapter 5**

**“Long-term Trends of Nutrients and Sediment from the Nontidal Chesapeake  
Watershed: An Assessment of Progress by River and Season”**

## Appendix D1: Details of monitoring data in the nine Chesapeake tributaries

This appendix documents the monitoring data analyzed in this work. Table D1 summarizes the nine USGS River Input Monitoring stations, which are located at the fall-line of nine major tributaries to Chesapeake Bay. Table D2 Summarizes details of water-quality parameters, including codes, names, abbreviations, and units. At each monitoring site, seven nutrient and sediment parameters have been analyzed, namely, suspended sediment (SS), dissolved phosphorus (DP), particulate phosphorus (PP), total phosphorus (TP), dissolved nitrogen (DN), particulate nitrogen (PN), and total nitrogen (TN). Details of the monitoring data for these parameters are given in Table D3, Table D4, Table D5, Table D6, Table D7, Table D8, and Table D9, respectively.

Table D1. Details of the monitoring stations, adapted from Table 3 in Sprague *et al.* (2000).

USGS site ID	River and station name	Abbreviation	Drainage area (km <sup>2</sup> )
01578310	Susquehanna River at Conowingo, MD	SUS	70,189
01646580	Potomac River at Chain Bridge, Washington D.C.	POT	30,044
02035000	James River at Cartersville, VA	JAM	16,213
01668000	Rappahannock River near Fredericksburg, VA	RAP	4,144
02041650	Appomattox River at Matoaca, VA	APP	3,471
01673000	Pamunkey River near Hanover, VA	PAM	2,774
01674500	Mattaponi River near Beulahville, VA	MAT	1,557
01594440	Patuxent River at Bowie, MD	PAT	901
01491000	Choptank River near Greensboro, MD	CHO	293

Table D2. Details of water-quality parameters.

Parameter code	Parameter name	Abbreviation	Unit
P00600	Total nitrogen, water, unfiltered	TN	mg N/L
P00601	Total nitrogen, suspended sediment, total	PN	mg N/L
P00602	Total nitrogen, water, filtered	DN	mg N/L
P62855 <sup>a</sup>	Total nitrogen, water, unfiltered, analytically determined	TN	mg N/L
P49570 <sup>a</sup>	Particulate nitrogen, suspended in water	PN	mg N/L
P62854 <sup>a</sup>	Total nitrogen, water, filtered, analytically determined	DN	mg N/L
P00625	Ammonia plus organic nitrogen, water, unfiltered	TKN	mg N/L
P00623	Ammonia plus organic nitrogen, water, filtered	DKN	mg N/L
P00631	Nitrate plus nitrite, water, filtered	DNO <sub>x</sub>	mg N/L
P00665	Phosphorus, water, unfiltered	TP	mg P/L
P00666	Phosphorus, water, filtered	DP	mg P/L
P00667	Phosphorus, suspended sediment, total	PP	mg P/L
P80154	Suspended sediment concentration	SS	mg/L

<sup>a</sup> P49570, P62854, and P62855 have been monitored at Maryland sites only.

Table D3. Details of monitoring data for suspended sediment (SS).<sup>a</sup>

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	1978-01-29	2013-10-23	None	1031 (0)	MD
POT	1973-02-05	2013-02-05	None	821 (1)	MD
PAT	1978-01-30	2013-02-27	None	908 (0)	MD
CHO	1974-11-04	2012-11-01	None	760 (0)	MD
JAM	1974-03-12	2013-02-28	1995-08-16 to 1999-10-01	671 (2)	VA
RAP	1977-12-30	2013-02-28	1994-11-09 to 2000-10-24	410 (0)	VA
APP	1977-12-28	2013-03-07	1994-08-24 to 1999-10-12	411 (0)	VA
PAM	1974-10-16	2013-03-07	1994-11-08 to 1999-10-01	458 (0)	VA
MAT	1979-04-10	2013-03-07	1994-11-08 to 2000-02-16, with 1 sample on 1997-11-10	387 (0)	VA

<sup>a</sup> For all sites, SS sample concentrations were reported as parameter P80154.

Table D4. Details of monitoring data for dissolved phosphorus (DP).<sup>a</sup>

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	1978-08-03	2013-04-17	None	952 (43)	MD
POT	1977-10-11	2013-04-20	None	487 (22)	MD
PAT	1978-01-30	2013-04-20	None	889 (21)	MD
CHO	1977-12-27	2013-03-27	None	647 (16)	MD
JAM	1977-10-18	2013-04-02	None	878 (3)	VA
RAP	1977-12-30	2013-04-09	None	755 (40)	VA
APP	1977-12-28	2013-04-03	None	828 (60)	VA
PAM	1977-10-19	2013-03-19	None	901 (28)	VA
MAT	1979-04-10	2013-03-19	None	869 (41)	VA

<sup>a</sup> For all sites, DP sample concentrations were reported as parameter P00666.

Table D5. Details of monitoring data for particulate phosphorus (PP).<sup>a</sup>

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	2003-11-13	2004-09-22	None	15 (0)	MD
POT	2003-12-08	2004-09-30	None	14 (0)	MD
PAT	2004-01-12	2004-09-29	None	14 (0)	MD
CHO	2003-04-16	2004-09-07	None	16 (0)	MD
JAM	1995-09-21	2013-04-02	None	441 (0)	VA
RAP	1996-02-20	2013-04-09	None	422 (1)	VA
APP	1996-02-22	2013-04-03	None	436 (1)	VA
PAM	1995-07-06	2013-04-16	None	478 (0)	VA
MAT	1995-07-06	2013-04-16	None	475 (2)	VA

<sup>a</sup> For all sites, PP sample concentrations were reported as parameter P00667.

Table D6. Details of monitoring data for total phosphorus (TP). <sup>a</sup>

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	1978-08-03	2013-04-17	None	986 (23)	MD
POT	1973-02-05	2013-04-20	None	632 (4)	MD
PAT	1978-01-30	2013-04-20	None	936 (1)	MD
CHO	1972-09-26	2013-03-27	None	726 (11)	MD
JAM	1974-02-21	2013-04-02	None	1018 (3)	VA <sup>b</sup>
RAP	1977-12-30	2013-04-09	None	824 (11)	VA <sup>b</sup>
APP	1977-12-28	2013-04-03	None	846 (8)	VA <sup>b</sup>
PAM	1974-10-16	2013-04-16	None	923 (3)	VA <sup>b</sup>
MAT	1979-04-10	2013-04-16	None	864 (5)	VA <sup>b</sup>

<sup>a</sup> For all sites, TP sample concentrations were reported as parameter P00665.

<sup>b</sup> TP sample concentrations were calculated as the sum of dissolved phosphorus (P00666) and particulate phosphorus (P00667) concentrations for the Virginia sites for specific periods: James (2/22/1996 to 10/19/2004); Rappahannock (2/20/1996 to 11/04/2004); Appomattox (2/22/1996 to 10/19/2004); Pamunkey (2/21/1996 to 11/02/2004); and Mattaponi (2/23/1996 to 11/02/2004).

Table D7. Details of monitoring data for dissolved nitrogen (DN). <sup>a</sup>

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	1978-08-03	2013-04-17	None	952 (43)	MD <sup>b</sup>
POT	1977-10-11	2013-04-20	None	487 (22)	MD <sup>b</sup>
PAT	1978-01-30	2013-04-20	None	889 (21)	MD <sup>b</sup>
CHO	1977-12-27	2013-03-27	None	647 (16)	MD <sup>b</sup>
JAM	1995-08-23	2013-04-02	None	444 (0)	VA <sup>c</sup>
RAP	1996-02-20	2013-04-09	None	423 (0)	VA <sup>c</sup>
APP	1996-02-22	2013-04-03	None	437 (0)	VA <sup>c</sup>
PAM	1995-07-06	2013-04-16	None	482 (0)	VA <sup>c</sup>
MAT	1995-07-06	2013-04-16	None	479 (0)	VA <sup>c</sup>

<sup>a</sup> For all sites, DN sample concentrations were reported as parameter P00602.

<sup>b</sup> For all Maryland sites, in the absence of DN sampling, DN concentrations were calculated as one of the following scenarios, whichever available: (1) P62854 (DN); (2) P00623 (DKN) + P00631 (DNO<sub>x</sub>).

<sup>c</sup> For all Virginia sites, DN concentrations were generally not available prior to 1995 or 1996 at these sites.

Table D8. Details of monitoring data for particulate nitrogen (PN)

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	2003-01-09	2013-04-17	None	192 (5)	MD <sup>a</sup>
POT	2001-10-30	2013-04-20	None	239 (3)	MD <sup>a</sup>
PAT	2003-01-08	2013-04-20	None	186 (5)	MD <sup>a</sup>
CHO	2003-01-13	2013-03-27	None	164 (15)	MD <sup>a</sup>
JAM	1995-08-23	2013-04-02	None	433 (19)	VA <sup>b</sup>
RAP	1996-02-20	2013-04-09	None	410 (28)	VA <sup>b</sup>
APP	1996-02-22	2013-04-03	None	431 (12)	VA <sup>b</sup>
PAM	1995-08-15	2013-04-16	None	471 (13)	VA <sup>b</sup>
MAT	1995-08-15	2013-04-16	None	468 (16)	VA <sup>b</sup>

<sup>a</sup> For all Maryland sites, PN concentrations were reported as parameter P49570.

<sup>b</sup> For all Virginia sites, PN concentrations were reported as parameter P00601. In addition, PN concentrations were generally not available prior to 1995 or 1996 at these sites.

Table D9. Details of monitoring data for total nitrogen (TN).<sup>a</sup>

Site	First Sample	Last Sample	Sample gap	No. of samples (censored)	State
SUS	1978-08-03	2013-04-17	None	1020 (29)	MD <sup>b</sup>
POT	1973-06-19	2013-04-16	None	1609 (8)	MD <sup>b</sup>
PAT	1978-01-30	2013-03-04	None	855 (7)	MD <sup>b</sup>
CHO	1974-11-04	2013-03-07	None	721 (27)	MD <sup>b</sup>
JAM	1974-02-21	2013-04-02	None	976 (26)	VA <sup>c</sup>
RAP	1978-01-31	2013-04-09	None	778 (41)	VA <sup>c</sup>
APP	1978-01-25	2013-04-03	None	798 (29)	VA <sup>c</sup>
PAM	1974-10-16	2013-03-19	None	876 (5)	VA <sup>c</sup>
MAT	1979-04-10	2013-03-19	None	815 (26)	VA <sup>c</sup>

<sup>a</sup> For all sites, TN sample concentrations were reported as parameter P00600.

<sup>b</sup> For all Maryland sites, in the absence of TN sampling, TN concentrations were calculated as one of the following scenarios, whichever available: (1) P00625 (TKN) + P00631 (DNO<sub>x</sub>); (2) P49570 (PN) + P62854 (DN); (3) P62855 (TN).

<sup>c</sup> For all Virginia sites, in the absence of TN sampling, TN concentrations were calculated as one of the following scenarios, whichever available: (1) P00601 (PN) + P00602 (DN); (2) P00625 (TKN) + P00631 (DNO<sub>x</sub>).

## **Appendix D2. Seasonal trend plots for the nine Chesapeake tributaries**

This appendix documents the complete set of long-term *seasonal* trends of *flow-normalized* nutrient and sediment loadings in the nine major tributaries to Chesapeake Bay. For each river, trend plots for seven constituents are presented, namely, total phosphorus (TP), particulate phosphorus (PP), dissolved phosphorus (DP), suspended sediment (SS), total nitrogen (TN), particulate nitrogen (PN), and dissolved nitrogen (DN). Details of the monitoring sites are summarized in Table 5.1 in the main article. Details of the monitoring data are summarized in Appendix D1. For completeness, trend plots for the nontidal Chesapeake Bay watershed (*i.e.*, sum of the nine rivers) are also enclosed.

Each of the following pages show seasonal trends for all seven constituents in one river. There are 10 pages of plots, which correspond to NTCBW (Figure D1), Susquehanna (Figure D2), Potomac (Figure D3), James (Figure D4), Rappahannock (Figure D5), Appomattox (Figure D6), Pamunkey (Figure D7), Mattaponi (Figure D8), Patuxent (Figure D9), and Choptank (Figure D10), respectively. The last figure (Figure D11) compares the long-term median yields among the nine tributaries.

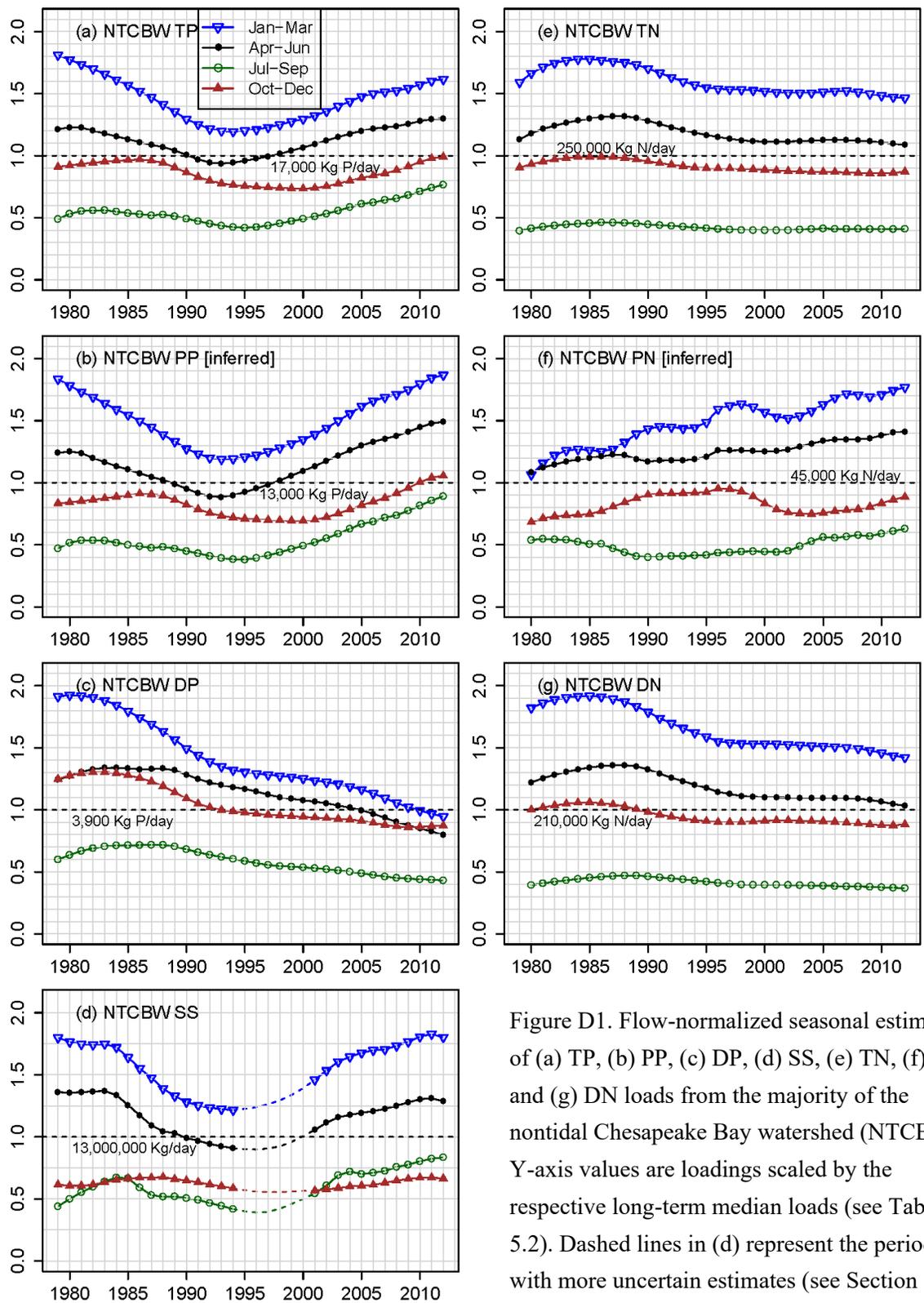


Figure D1. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the majority of the nontidal Chesapeake Bay watershed (NTCBW). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

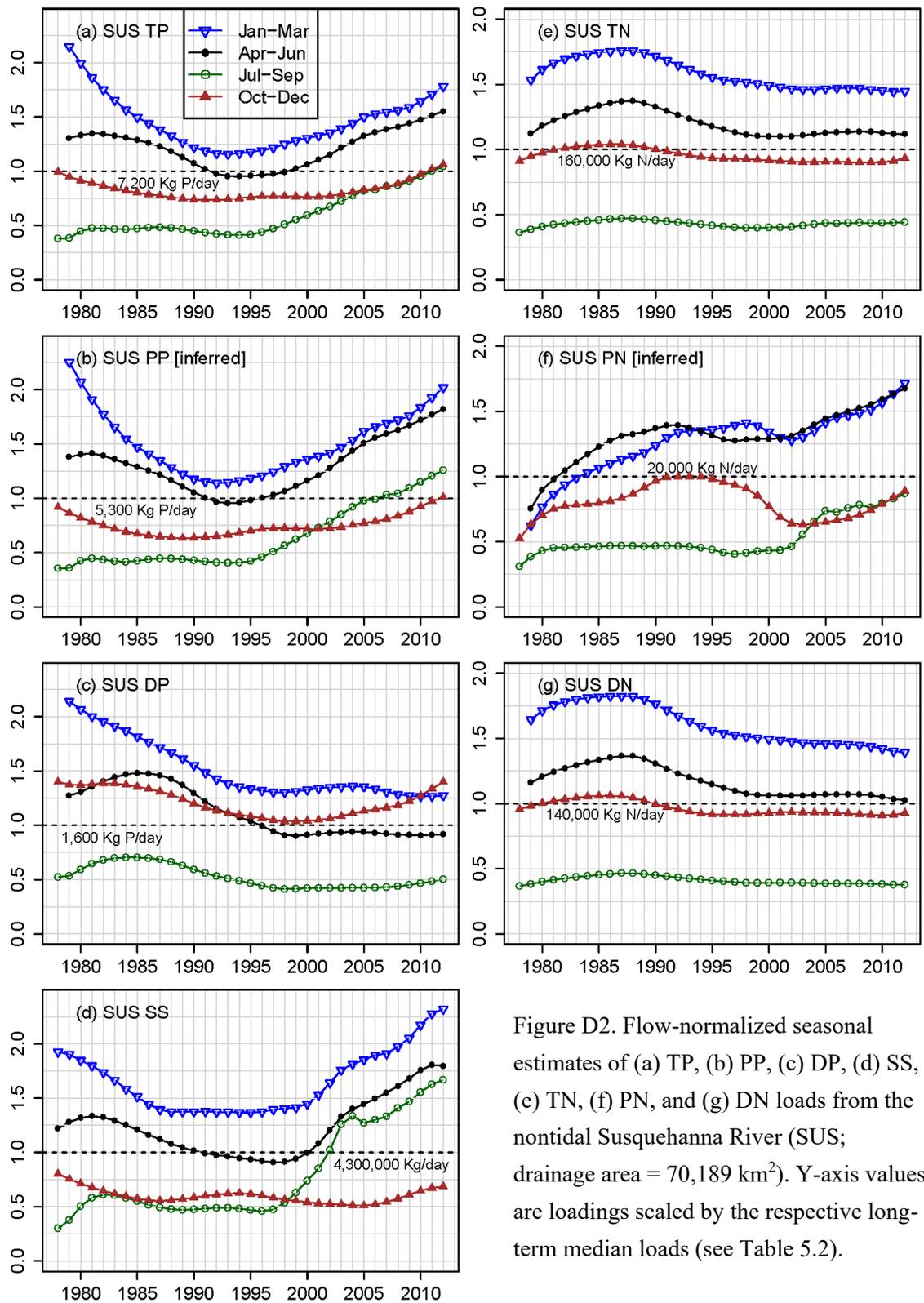


Figure D2. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Susquehanna River (SUS; drainage area = 70,189 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2).

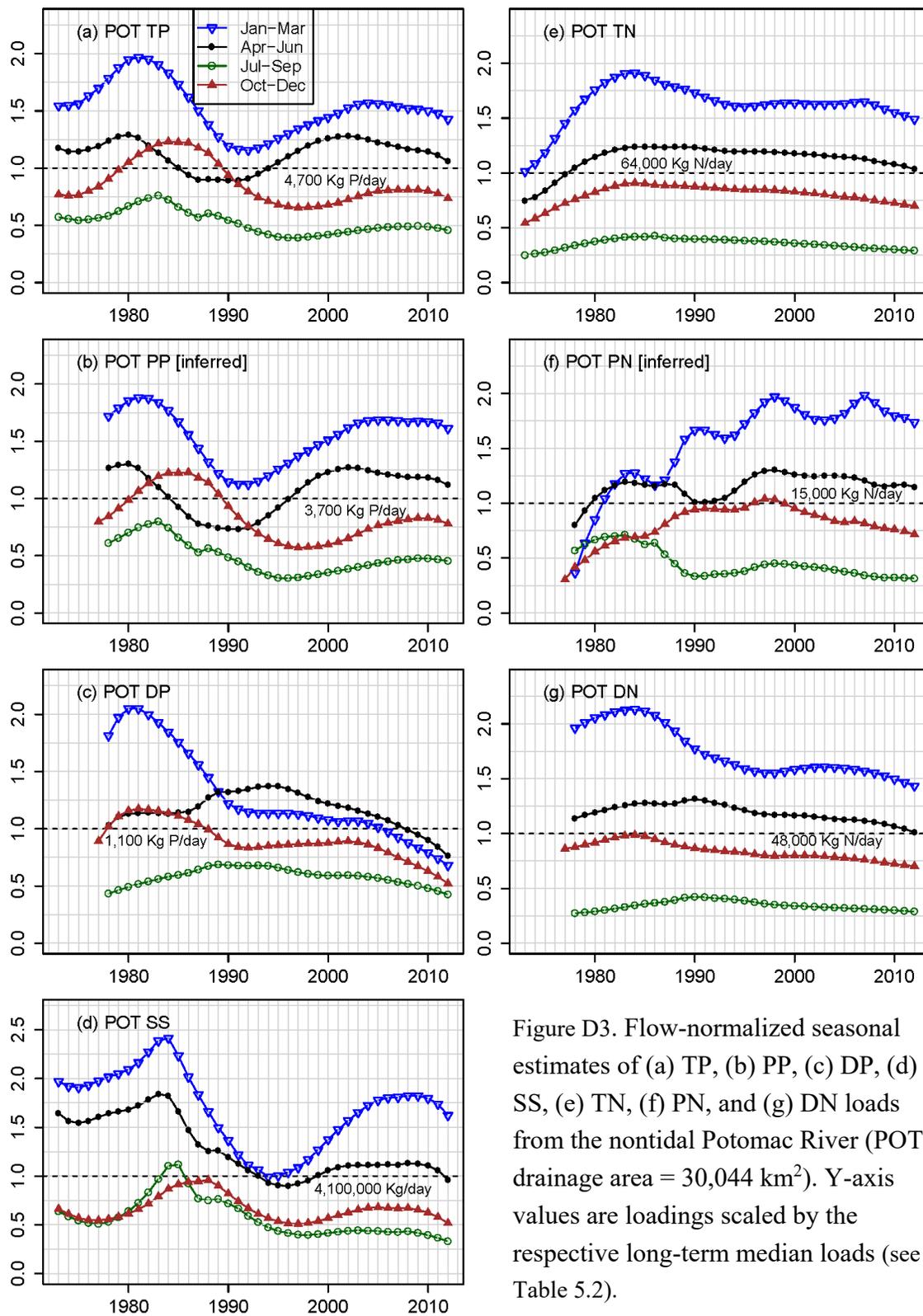


Figure D3. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Potomac River (POT; drainage area = 30,044 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2).

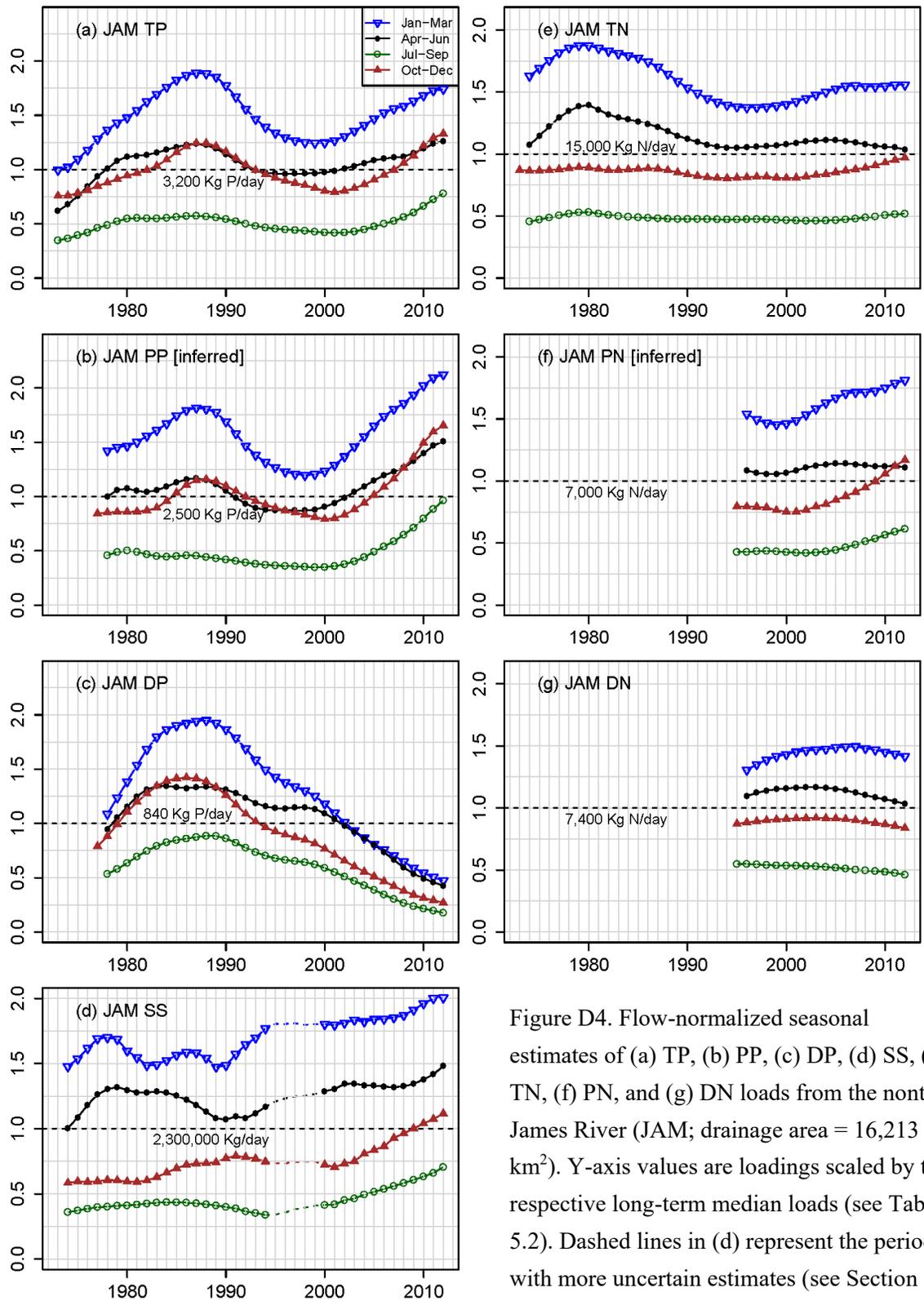


Figure D4. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal James River (JAM; drainage area = 16,213 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

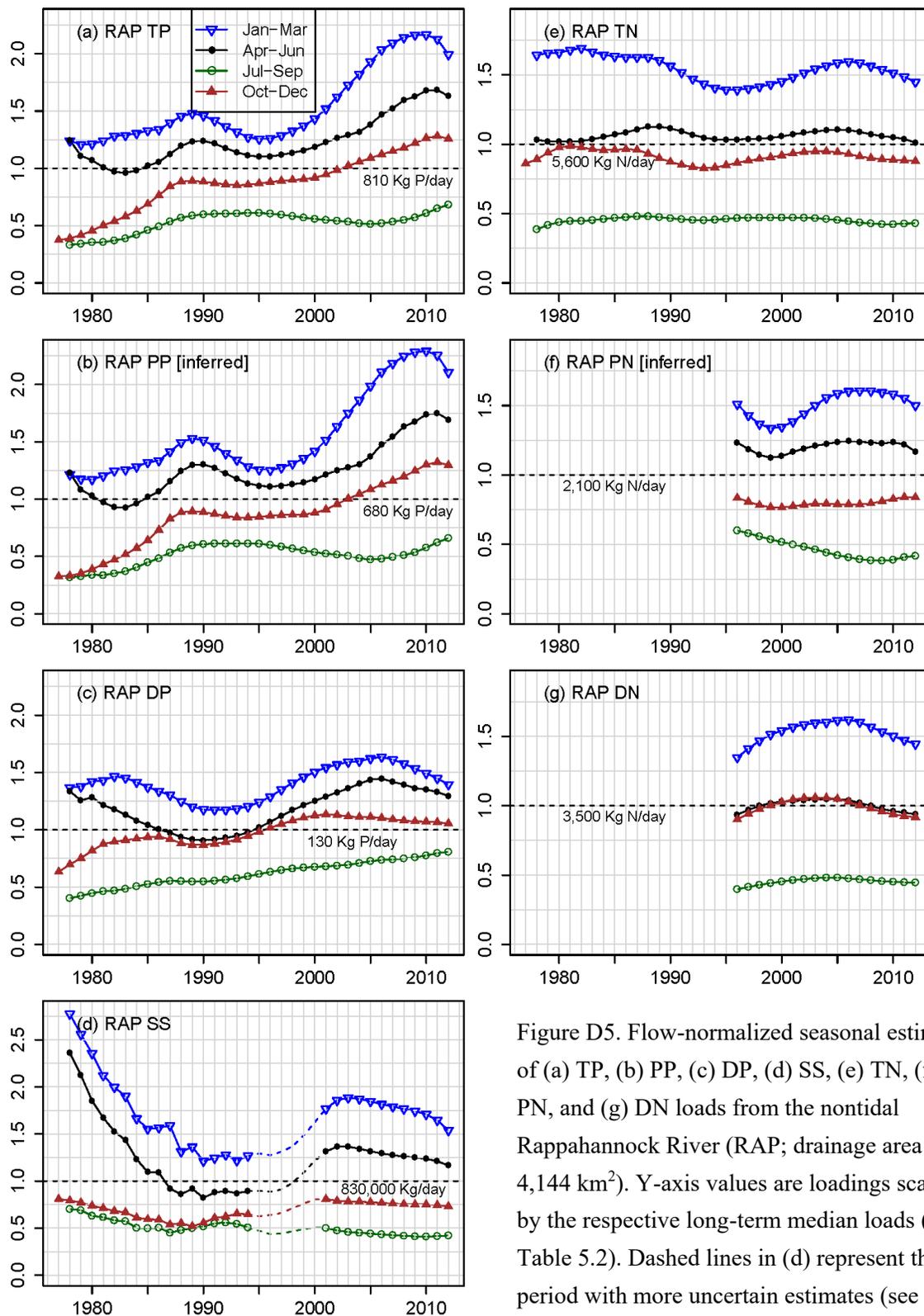


Figure D5. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Rappahannock River (RAP; drainage area = 4,144 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

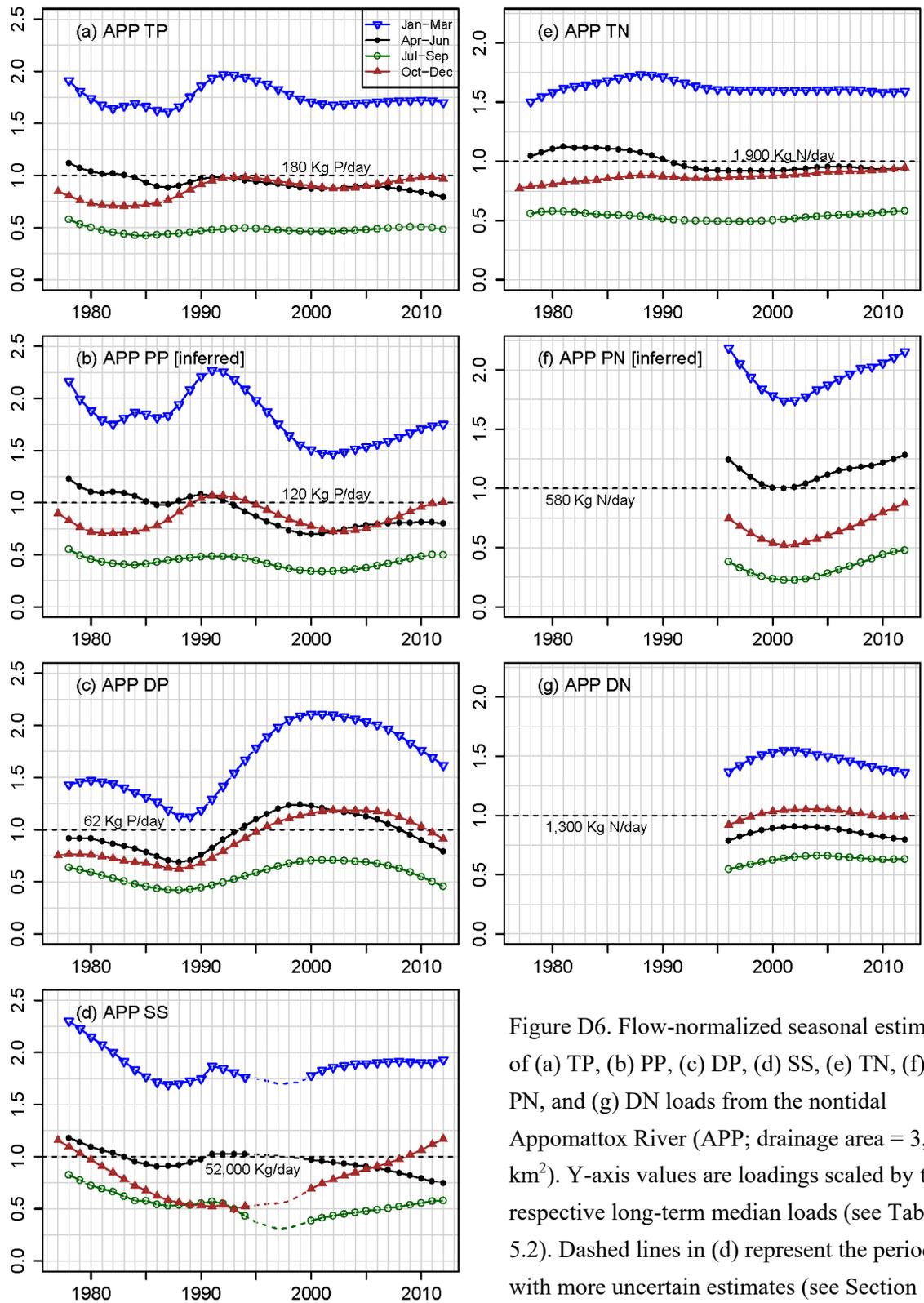


Figure D6. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Appomattox River (APP; drainage area = 3,471 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

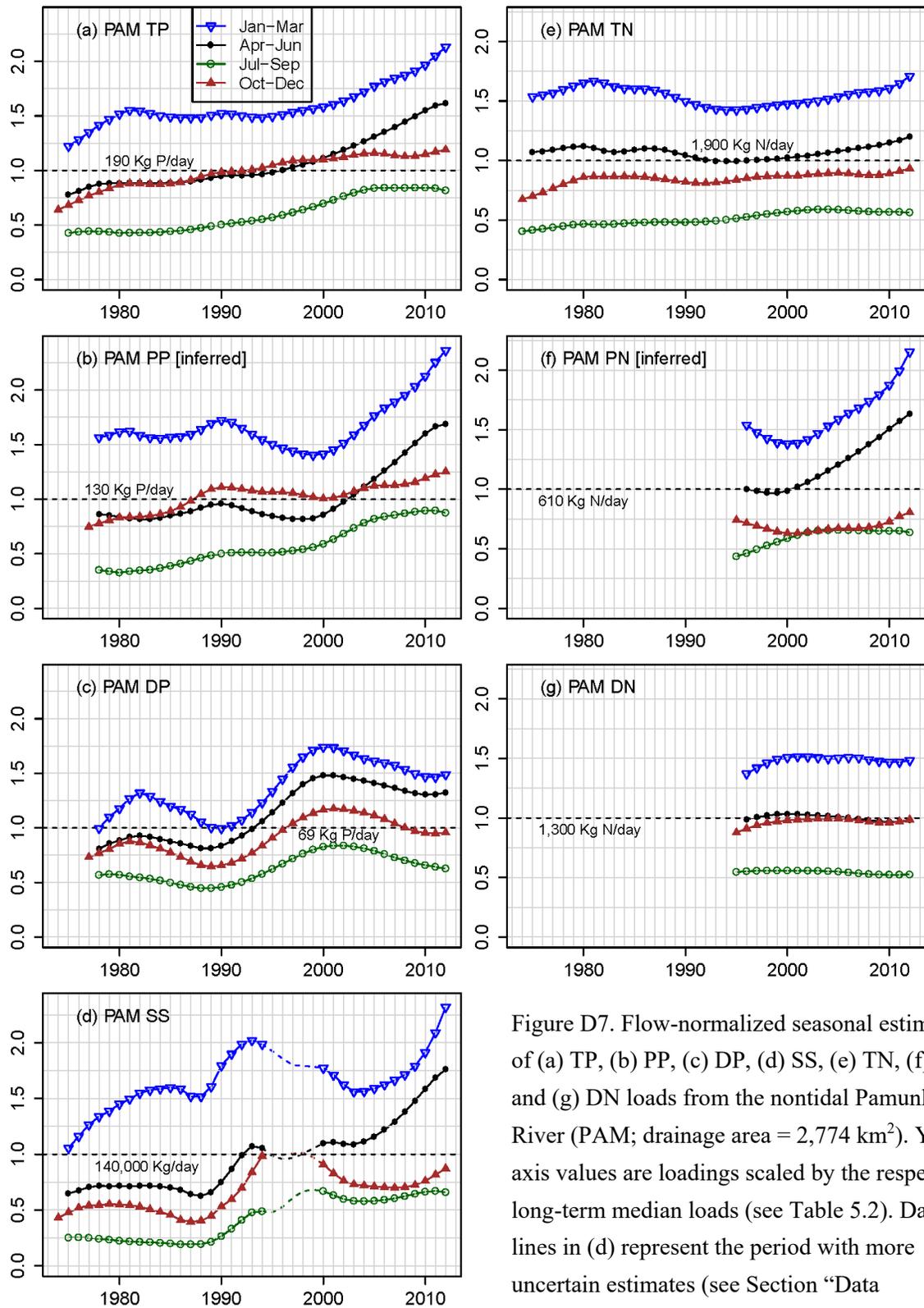


Figure D7. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Pamunkey River (PAM; drainage area = 2,774 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2). Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

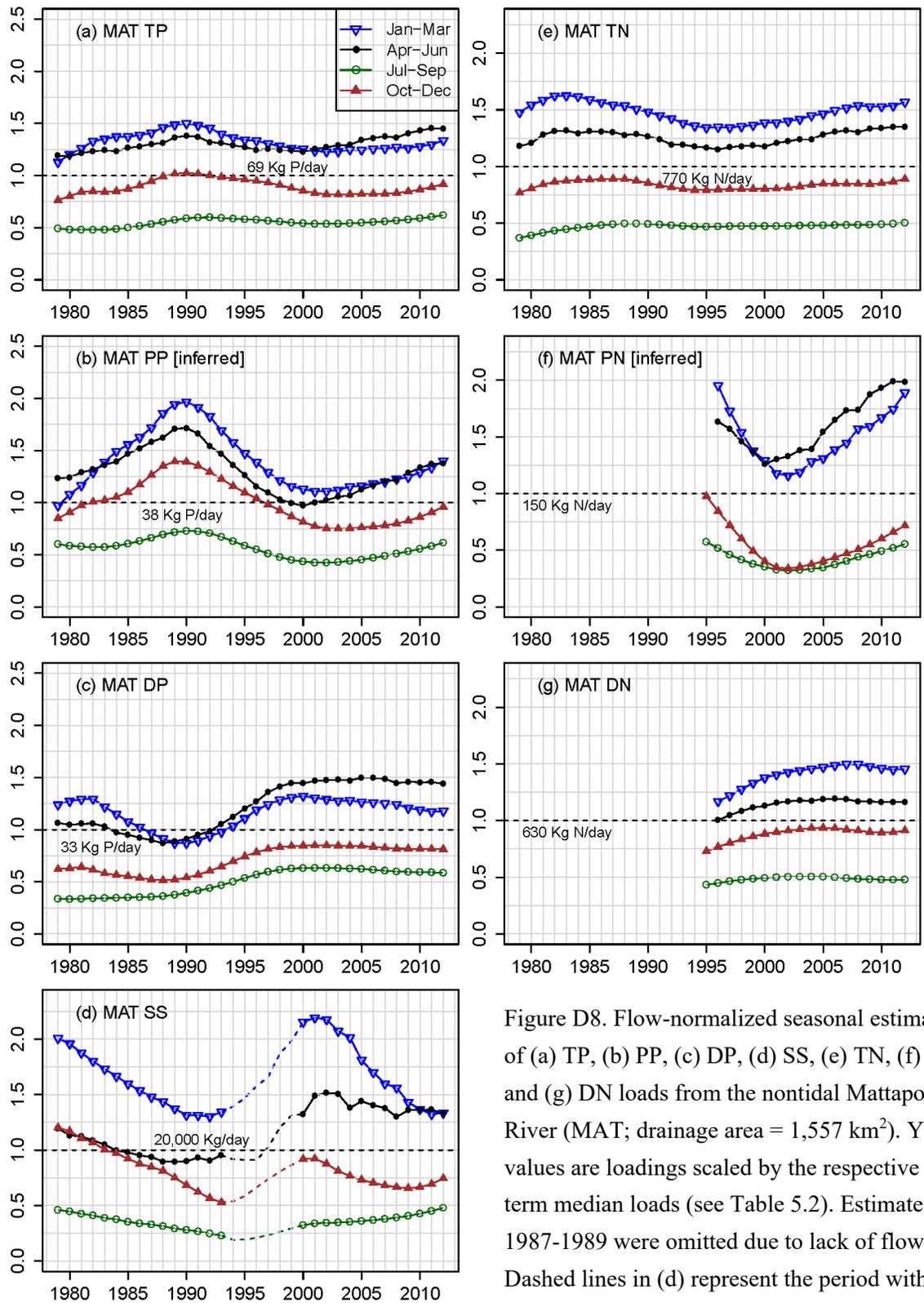


Figure D8. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Mattaponi River (MAT; drainage area = 1,557 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2). Estimates for 1987-1989 were omitted due to lack of flow data. Dashed lines in (d) represent the period with more uncertain estimates (see Section “Data compilation and analyses”).

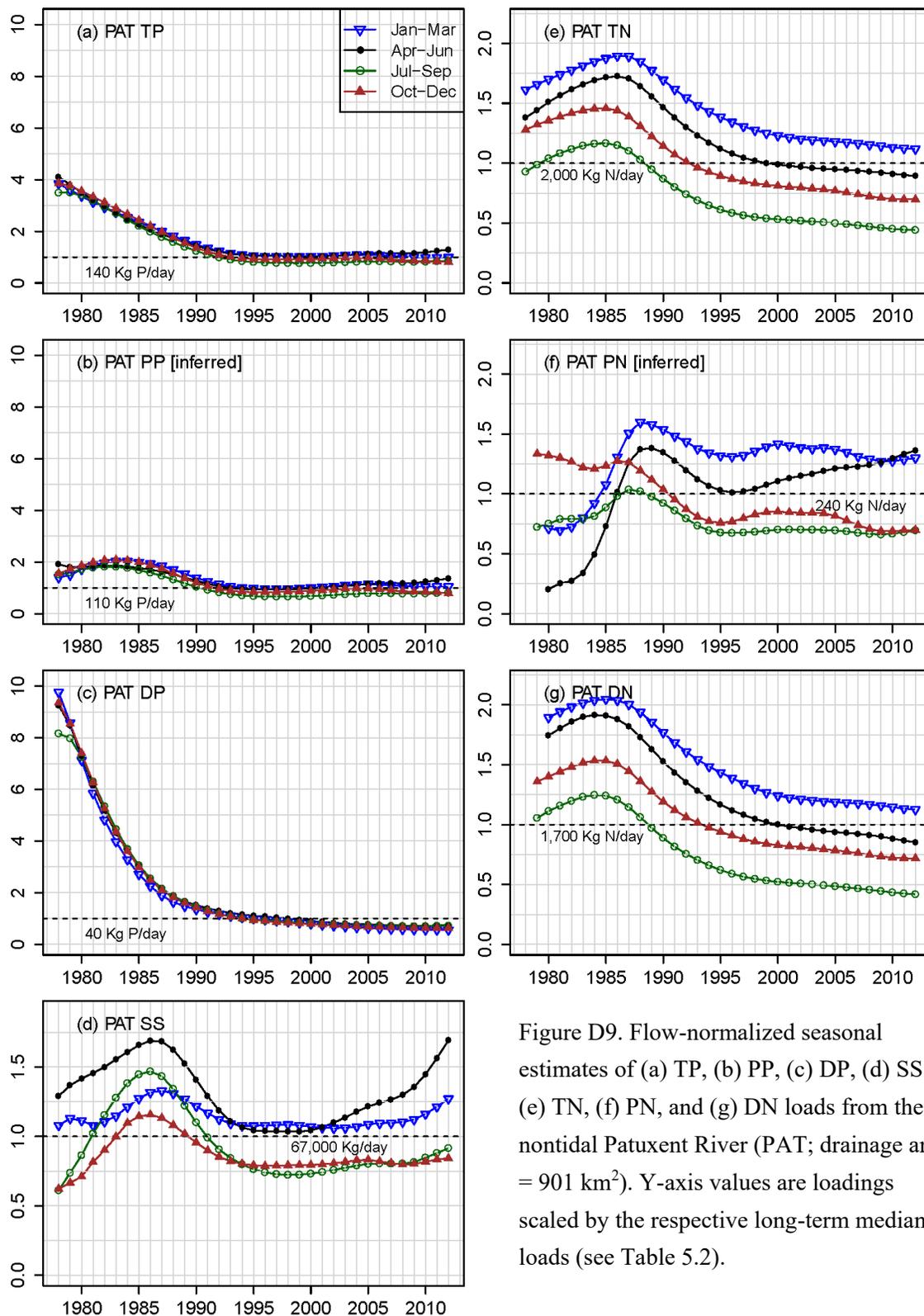


Figure D9. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Patuxent River (PAT; drainage area = 901 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2).

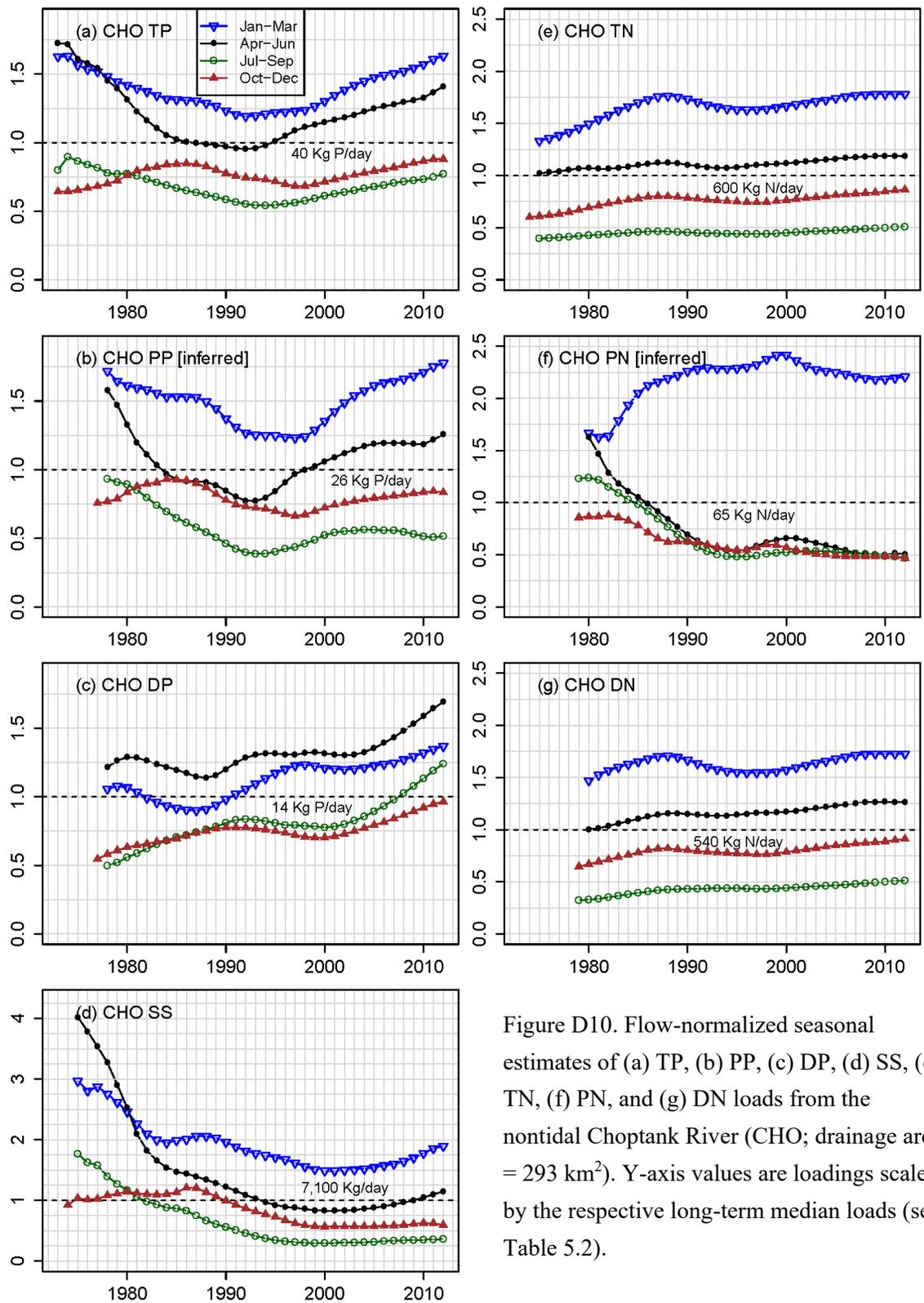


Figure D10. Flow-normalized seasonal estimates of (a) TP, (b) PP, (c) DP, (d) SS, (e) TN, (f) PN, and (g) DN loads from the nontidal Choptank River (CHO; drainage area = 293 km<sup>2</sup>). Y-axis values are loadings scaled by the respective long-term median loads (see Table 5.2).

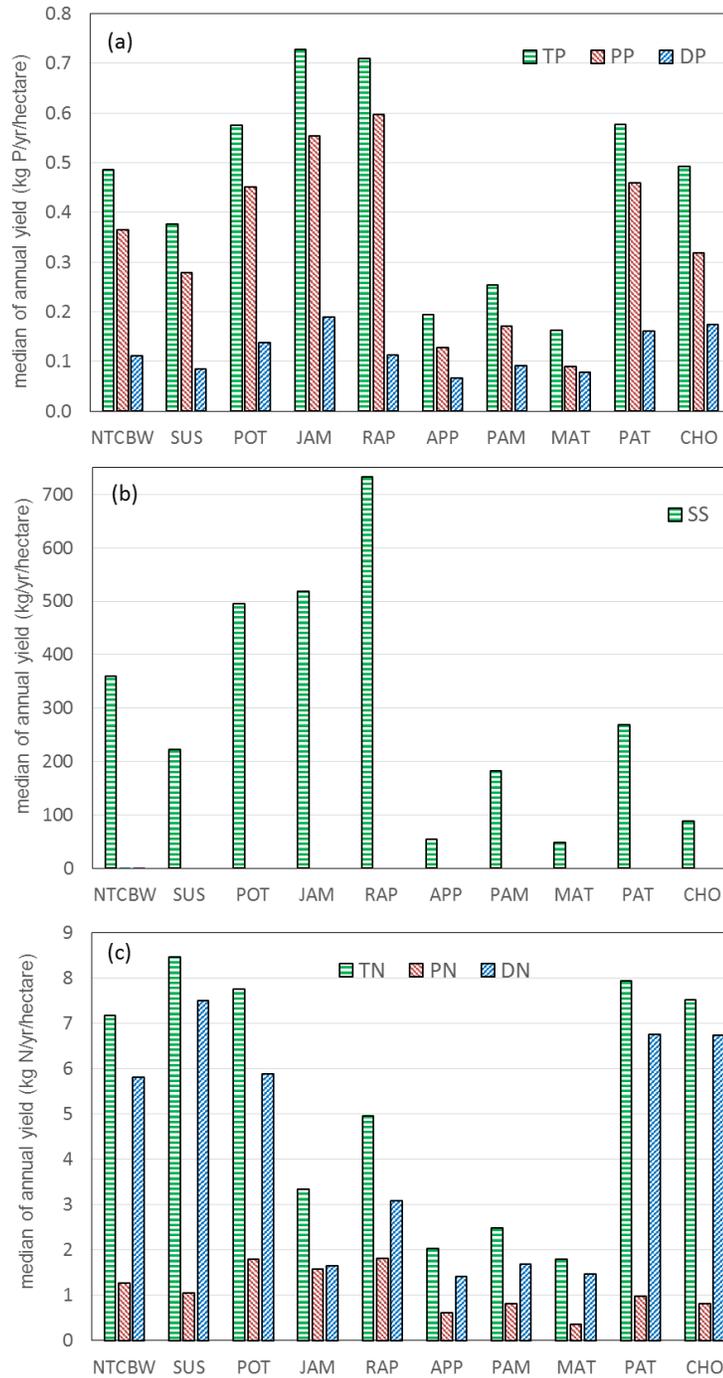


Figure D11. Comparison of long-term median values of annual yields for (a) TP, PP, DP, (b) SS, (c) TN, PN, DN among the nontidal Chesapeake Bay watershed (NTCBW) and the nine tributaries (SUS: Susquehanna, POT: Potomac, JAM: James, RAP: Rappahannock, APP: Appomattox, PAM: Pamunkey, MAT: Mattaponi, PAT: Patuxent, CHO: Choptank). See Table 5.2 for exact numbers.

### Appendix D3. Monthly trends of phosphorus loading in the three Chesapeake tributaries that show cross-over effects <sup>a</sup>

Table D10. Changes in the relative ranking of phosphorus loading in Season 2 (S2) and Season 4 (S4) in the three tributaries. <sup>a</sup>

River	Species	Figure No.	Timing of cross-over	Period with S2 > S4	Period with S2 ≈ S4	Period with S2 < S4	Causative monthly trends
Appomattox	TP	B6a	1993 & 2007	1978-1992 (F <sub>S2</sub> /F <sub>S4</sub> = 1.24) <sup>b</sup> (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.44) <sup>c</sup>	1993-2006 (F <sub>S2</sub> /F <sub>S4</sub> = 0.98) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.15)	2007-2012 (F <sub>S2</sub> /F <sub>S4</sub> = 0.88) (Q <sub>S2</sub> /Q <sub>S4</sub> = 0.95)	Strong declines in March, April, and May since 1991; relatively flat in other months (Figure D12).
Appomattox	PP	B6b	1991	1978-1990 (F <sub>S2</sub> /F <sub>S4</sub> = 1.34) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.45)	-	1991-2012 (F <sub>S2</sub> /F <sub>S4</sub> = 0.92) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.11)	Strong declines in January, February, March, April, and May since 1991; relatively flat in other months (Figure D13).
Appomattox	DP	B6c	2002	1978-2001 (F <sub>S2</sub> /F <sub>S4</sub> = 1.13) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.48)	-	2002-2012 (F <sub>S2</sub> /F <sub>S4</sub> = 0.93) (Q <sub>S2</sub> /Q <sub>S4</sub> = 0.86)	Much stronger declines in March, April, and May than in the other months since 2002 (Figure D14).
Susquehanna	DP	B2c	1993	1982-1992 (F <sub>S2</sub> /F <sub>S4</sub> = 1.07) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.60)	-	1993-2012 (F <sub>S2</sub> /F <sub>S4</sub> = 0.84) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.30)	Strong declines in February, March, April, May, and June since 1993; slight rises in September, October, November, and December (Figure D15).
Pamunkey	PP	B7b	2004	2004-2012 (F <sub>S2</sub> /F <sub>S4</sub> = 1.22) (Q <sub>S2</sub> /Q <sub>S4</sub> = 0.89)	-	1981-2003 (F <sub>S2</sub> /F <sub>S4</sub> = 0.87) (Q <sub>S2</sub> /Q <sub>S4</sub> = 1.49)	Much stronger rises in January, February, March, April, May, and June than in the other months since 2003 (Figure D16).

<sup>a</sup> Table shows only those cases where dominant season for loading in Season 2 and Season 4 showed a mid-term switch (*i.e.*, “cross-over” effect).

<sup>b</sup> F<sub>S2</sub>/F<sub>S4</sub> represents the ratio between average flow-normalized loads in S2 and S4 for the corresponding periods.

<sup>c</sup> Q<sub>S2</sub>/Q<sub>S4</sub> represents the ratio between average streamflow discharges in S2 and S4 for the corresponding periods.

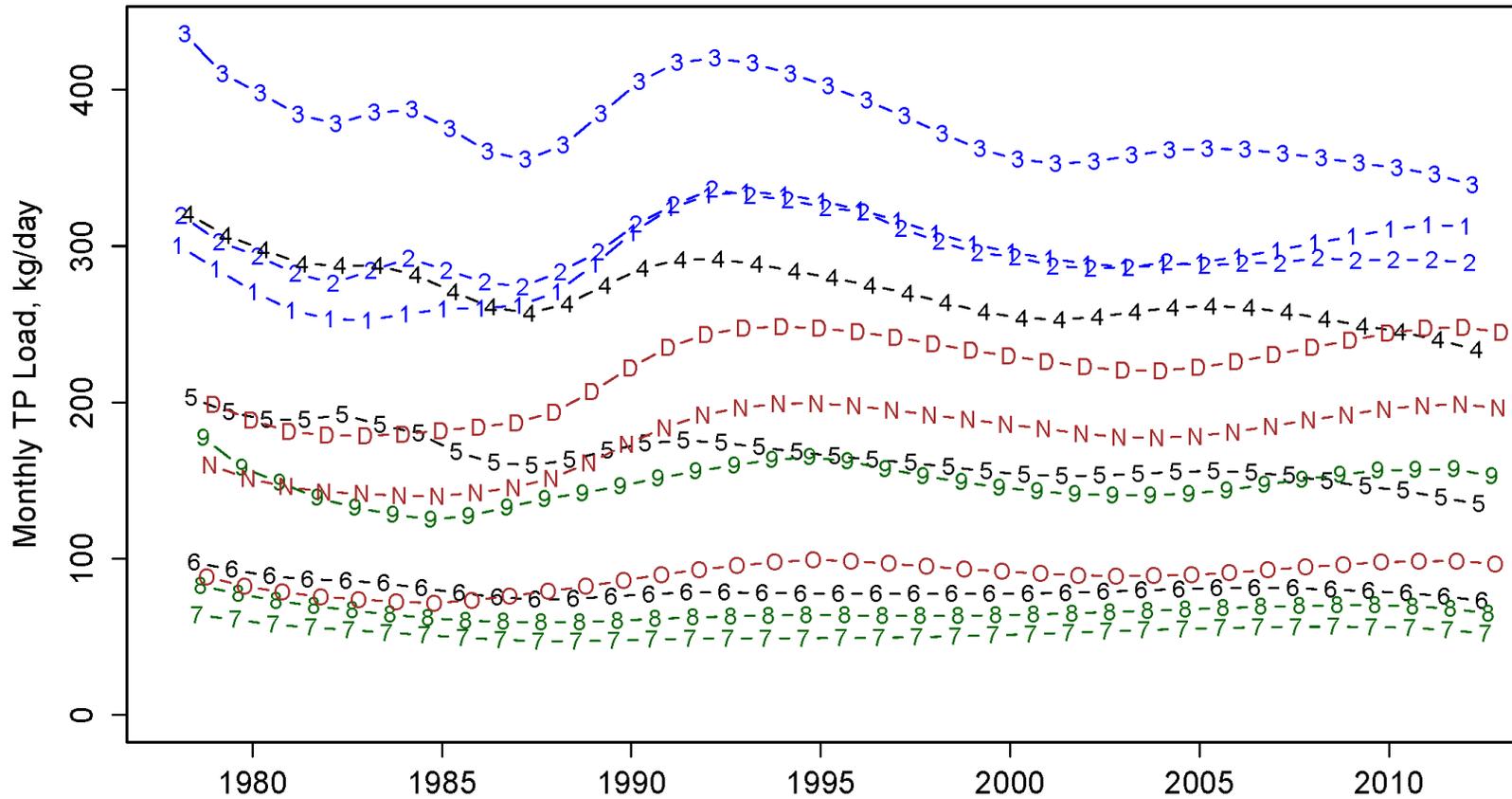


Figure D12. Flow-normalized monthly TP load in the Appomattox River. (Legend: 1=January; 2=February; 3=March; 4=April; 5=May; 6=June; 7=July; 8=August; 9=September; O=October; N=November; D=December.)

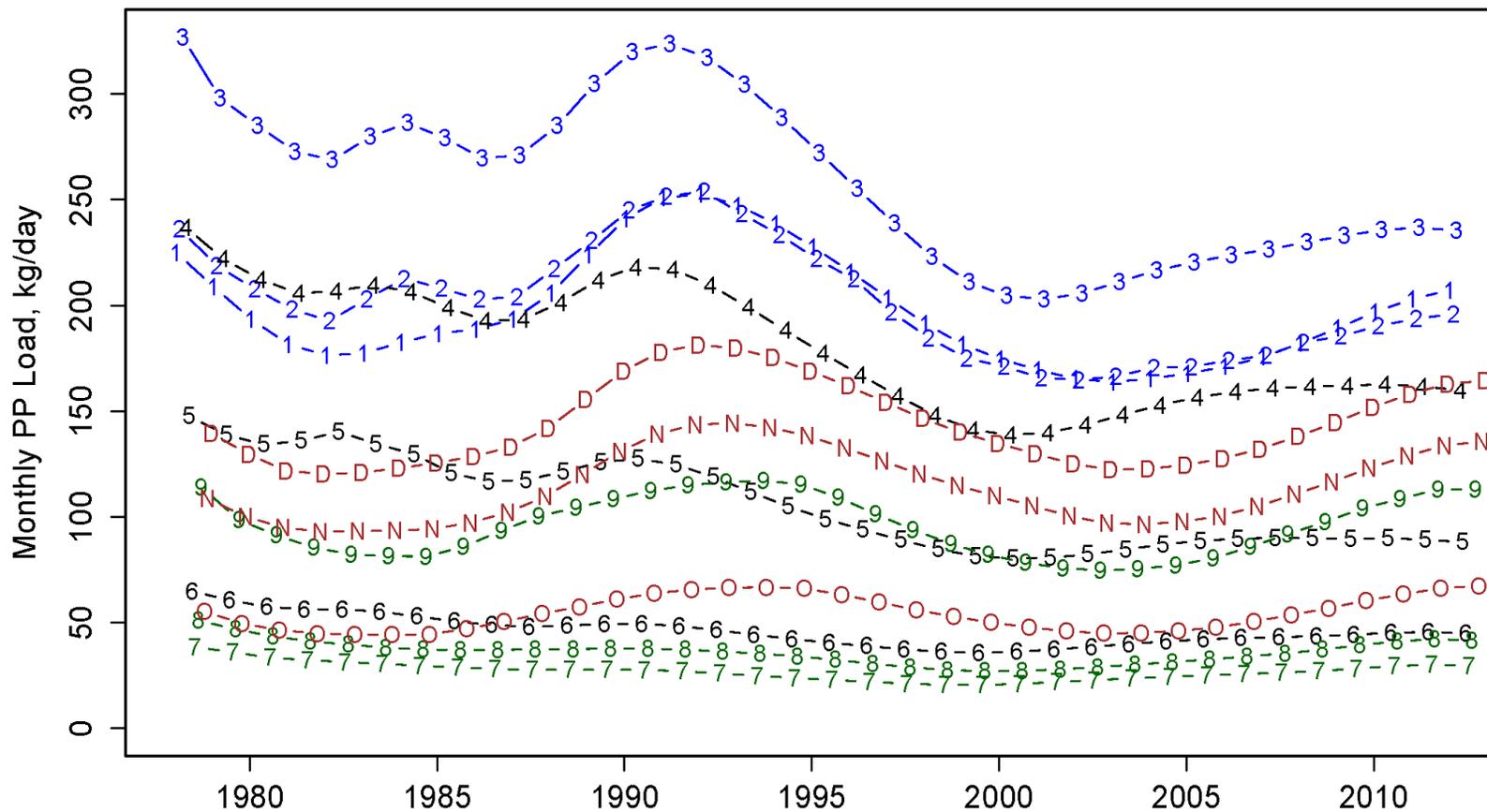


Figure D13. Flow-normalized monthly PP load in the Appomattox River. (Legend: 1=January; 2=February; 3=March; 4=April; 5=May; 6=June; 7=July; 8=August; 9=September; O=October; N=November; D=December.)

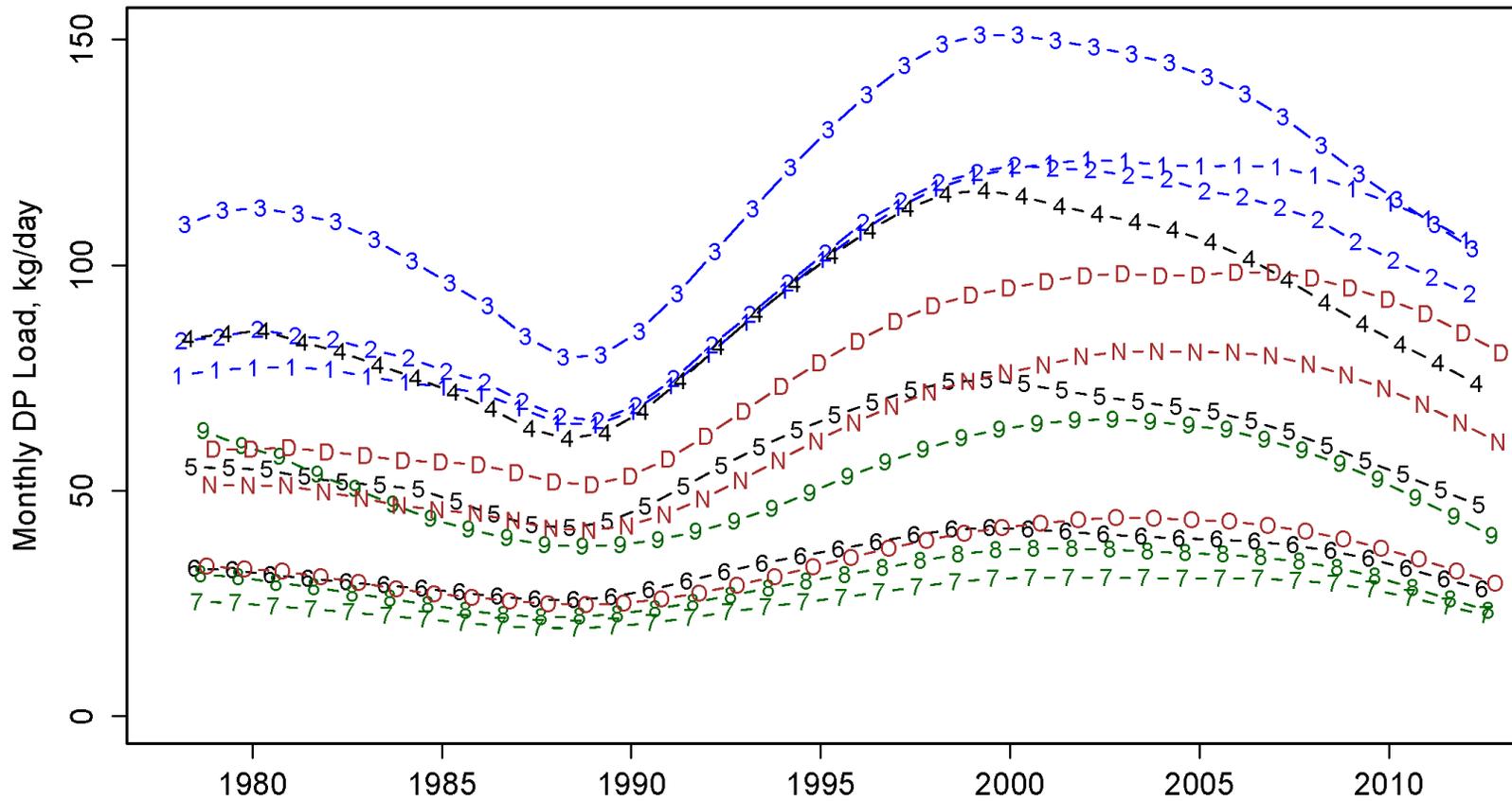


Figure D14. Flow-normalized monthly DP load in the Appomattox River. (Legend: 1=January; 2=February; 3=March; 4=April; 5=May; 6=June; 7=July; 8=August; 9=September; O=October; N=November; D=December.)

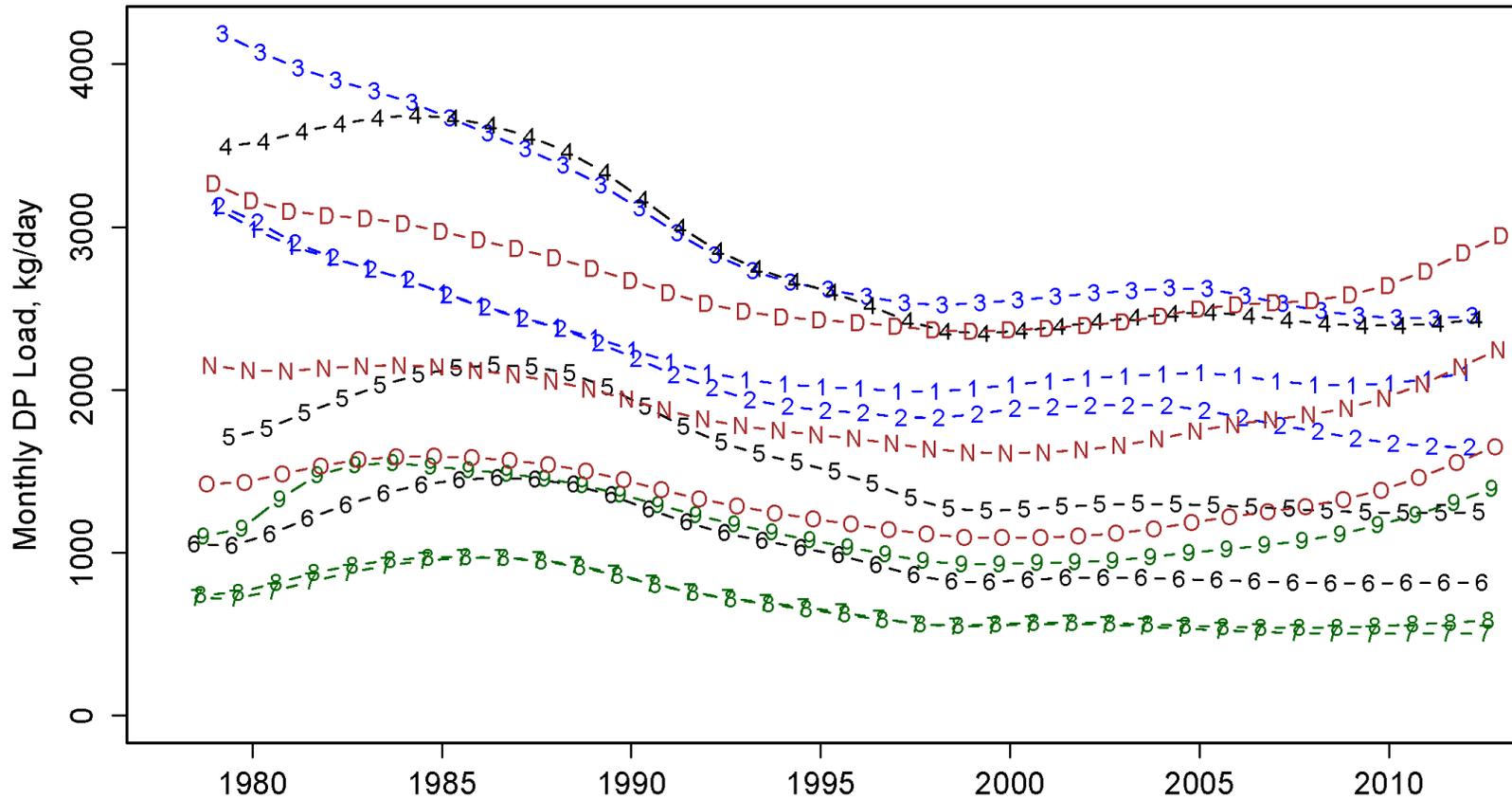


Figure D15. Flow-normalized monthly DP load in the Susquehanna River. (Legend: 1=January; 2=February; 3=March; 4=April; 5=May; 6=June; 7=July; 8=August; 9=September; O=October; N=November; D=December.)

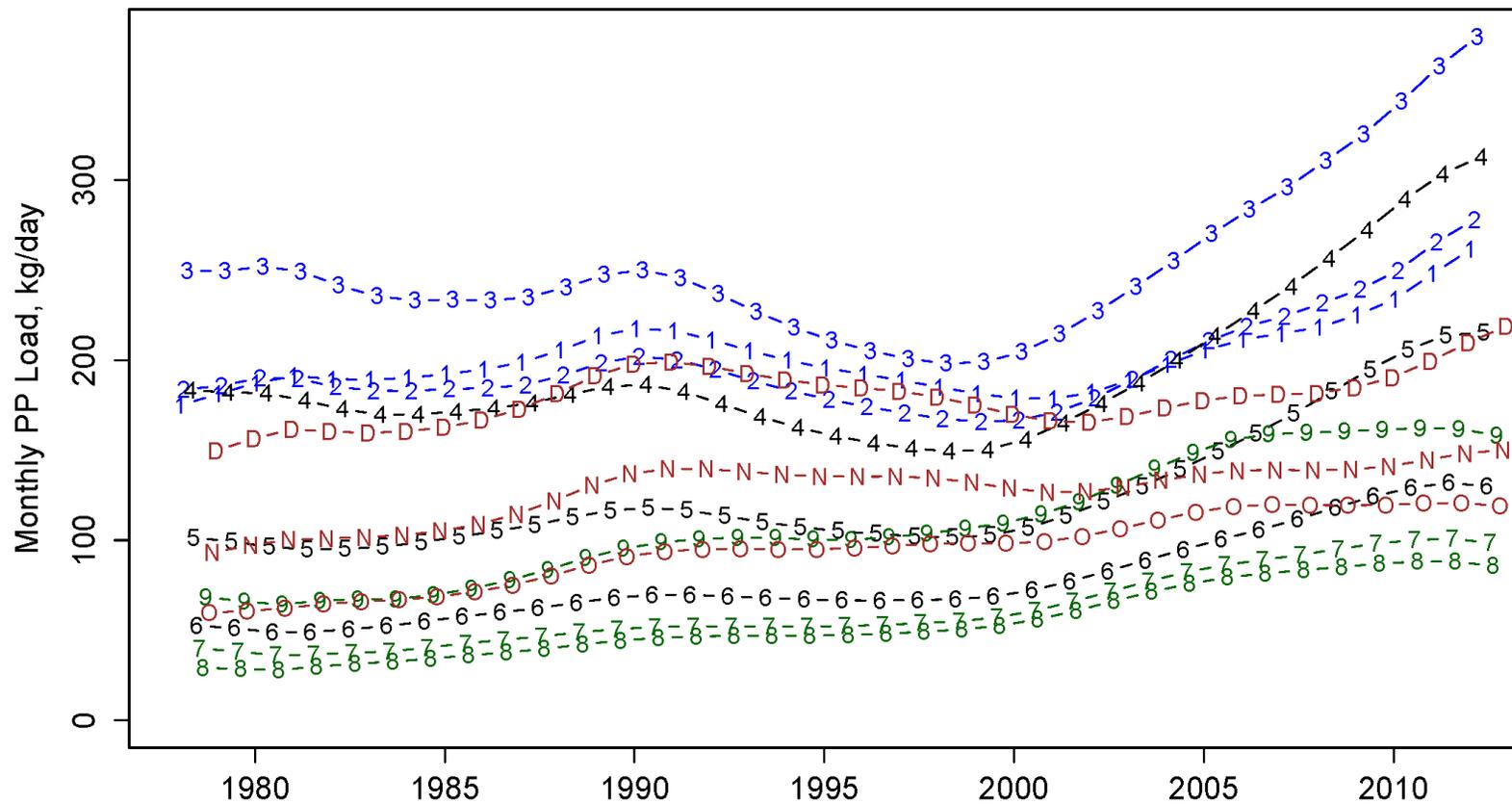


Figure D16. Flow-normalized monthly PP load in the Pamunkey River. (Legend: 1=January; 2=February; 3=March; 4=April; 5=May; 6=June; 7=July; 8=August; 9=September; O=October; N=November; D=December.)

**Appendix D4. Fractional contributions of sub-species loads by the nine Chesapeake tributaries**

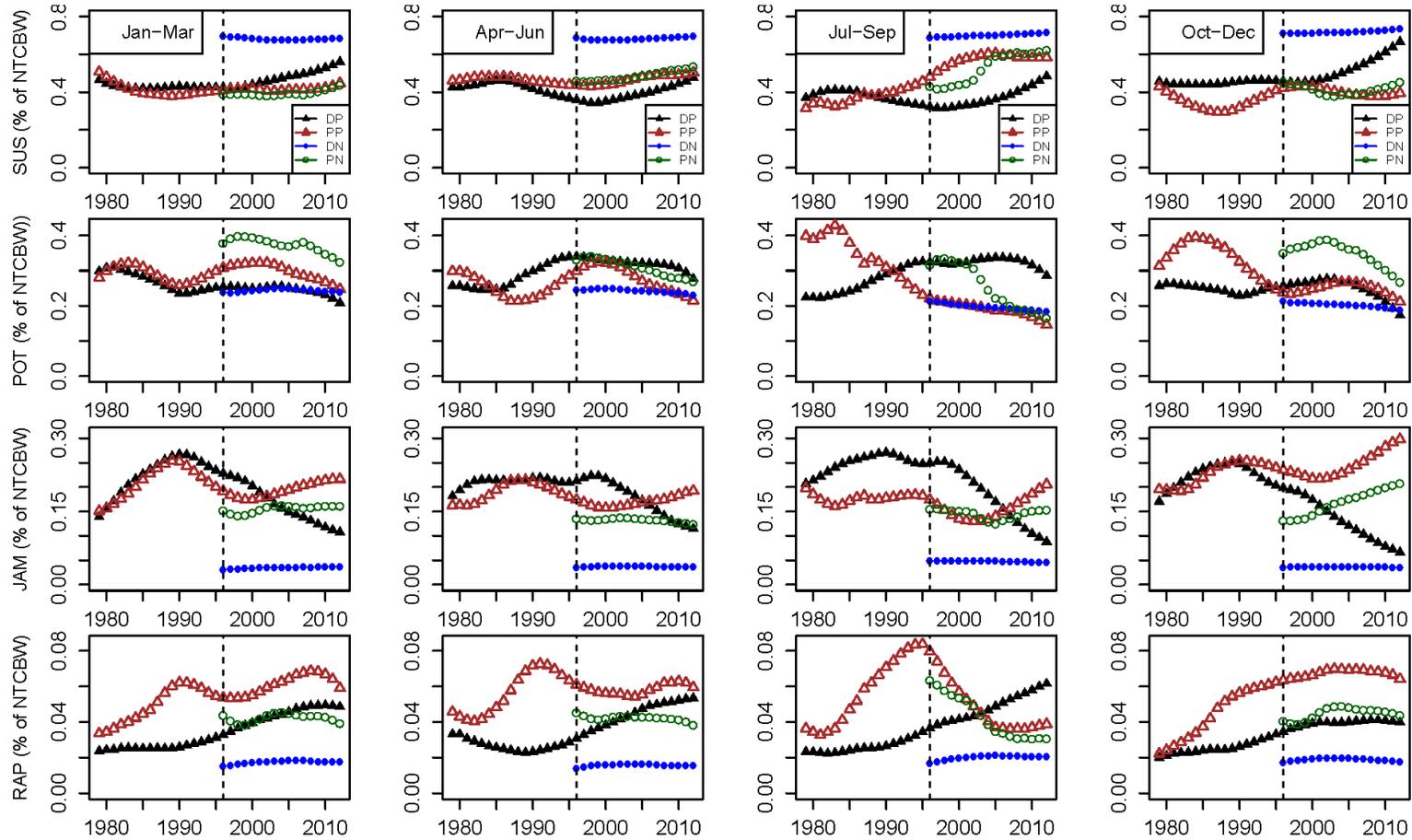


Figure D17. Seasonal fractional contributions of the four largest tributaries to the nontidal Chesapeake Bay watershed (NTCBW) for streamflow discharge and flow-normalized loadings of DP, PP, DN, and PN (SUS: Susquehanna; POT: Potomac; JAM: James; RAP: Rappahannock).

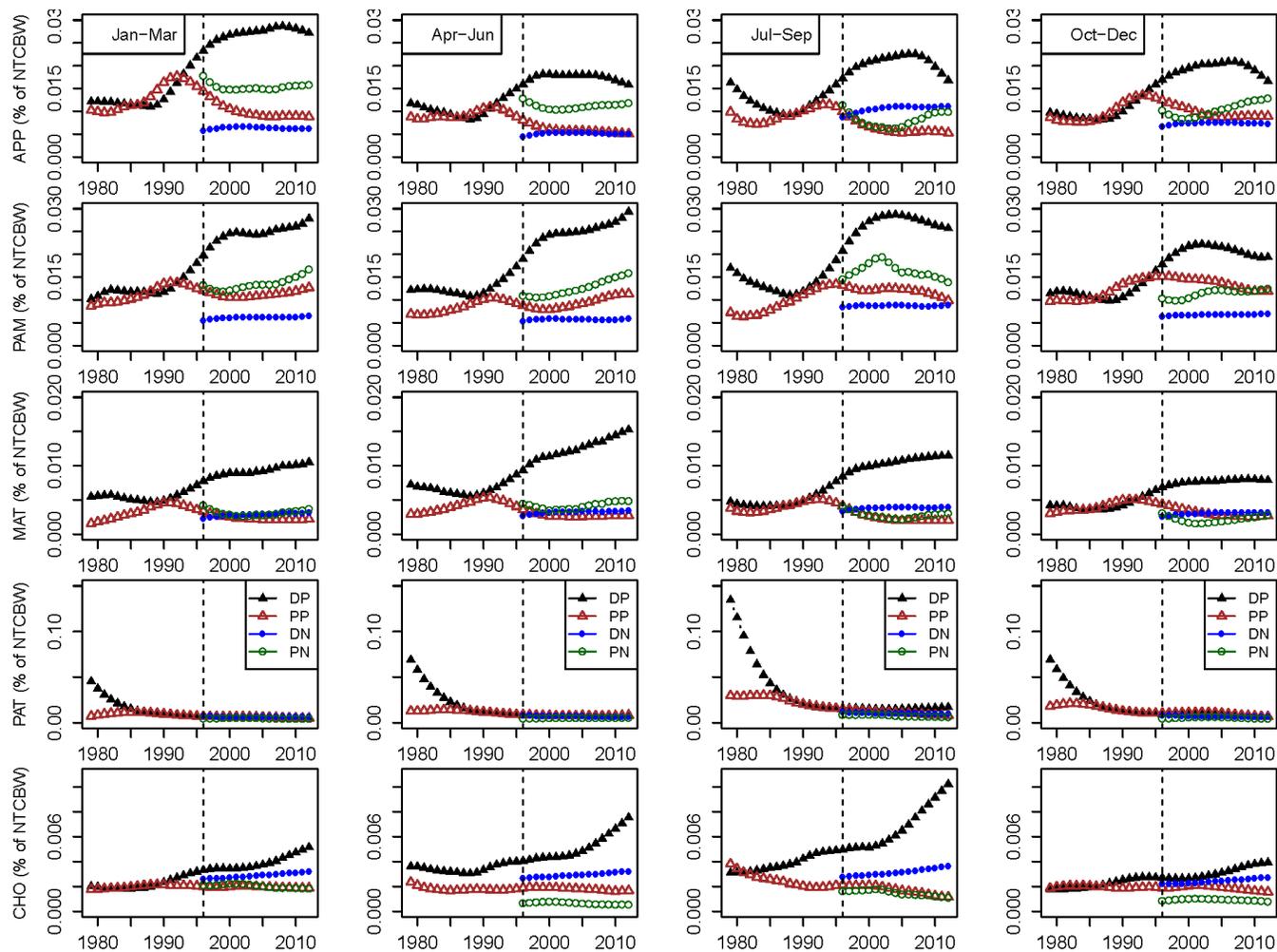


Figure D18. Seasonal fractional contributions of the five smallest tributaries to the nontidal Chesapeake Bay watershed (NTCBW) for streamflow discharge and flow-normalized loadings of DP, PP, DN, and PN (APP: Appomattox; PAM: Pamunkey; MAT: Mattaponi; PAT: Patuxent; CHO: Choptank).

**Appendix D5. Seasonal N:P ratios in the nontidal Chesapeake Bay watershed and the nine tributaries**

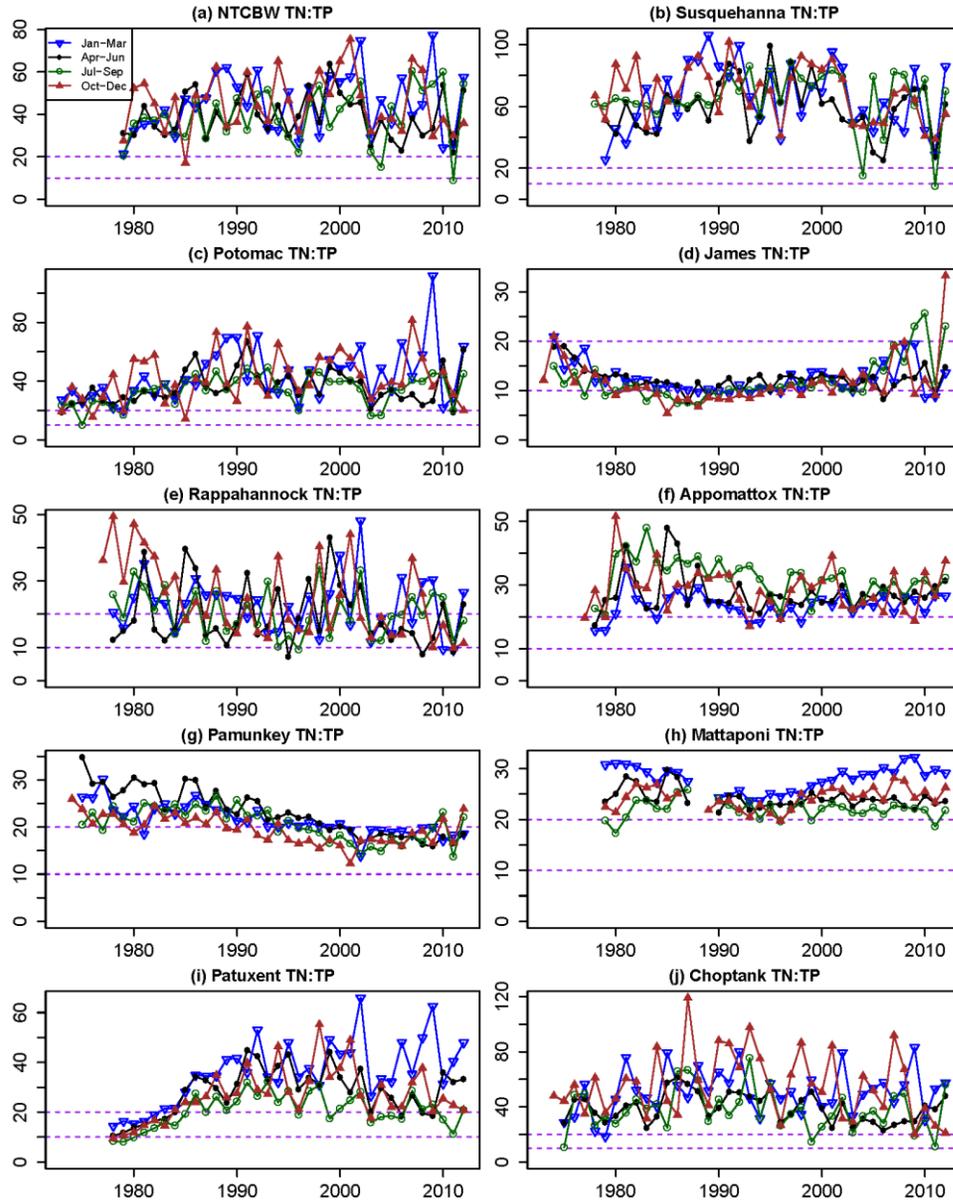


Figure D19. Seasonal TN:TP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW; sum of the nine rivers) and the nine individual tributaries, as calculated based on WRTDS *true-condition* estimates. Y-axis indicates the N:P ratio, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

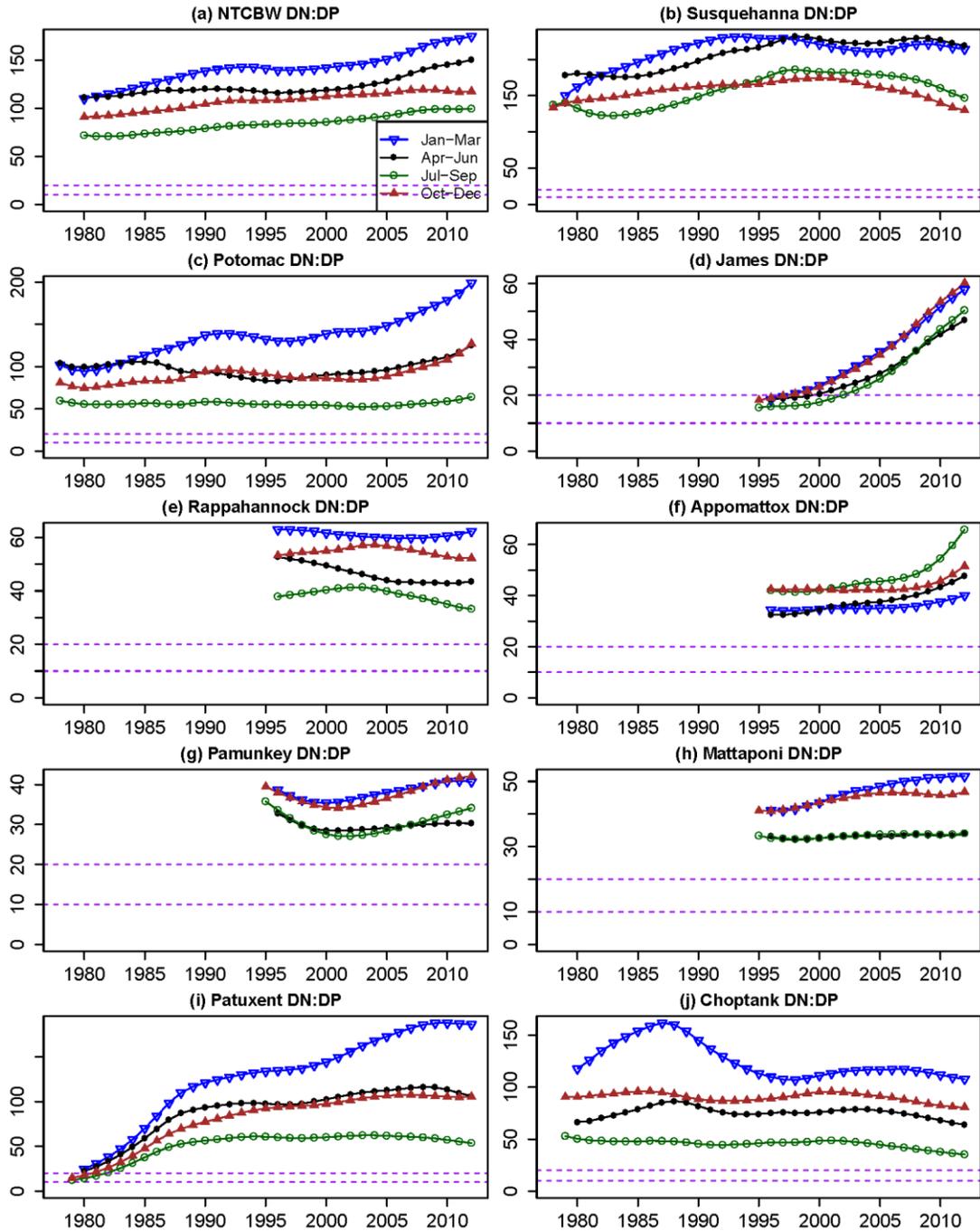


Figure D20. Seasonal DN:DP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW; sum of the nine rivers) and the nine individual tributaries, as calculated based on WRTDS *flow-normalized* estimates. Y-axis indicates the N:P ratio, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

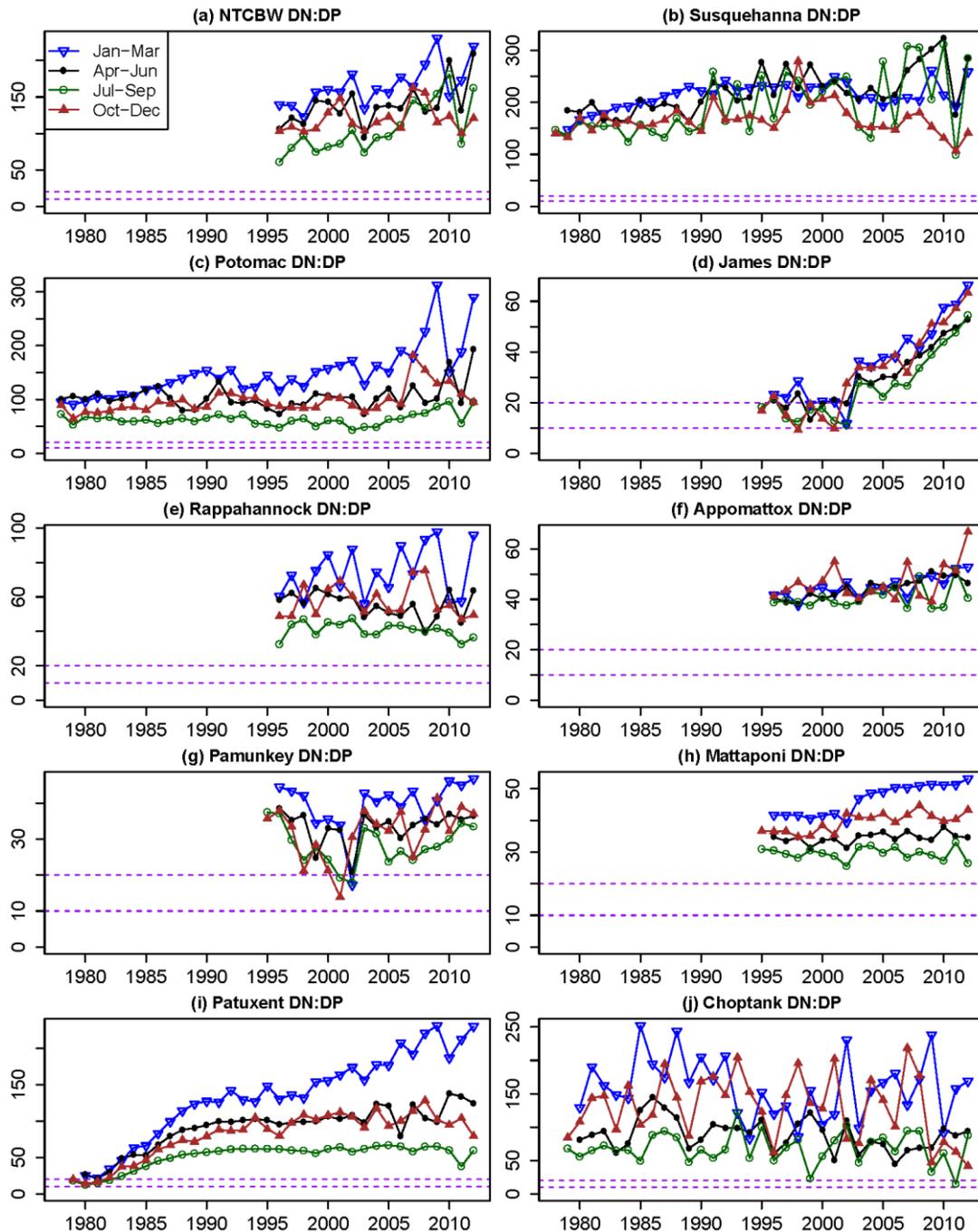


Figure D21. Seasonal DN:DP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW; sum of the nine rivers) and the nine individual tributaries, as calculated based on WRTDS *true-condition* estimates. Y-axis indicates the N:P ratio, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

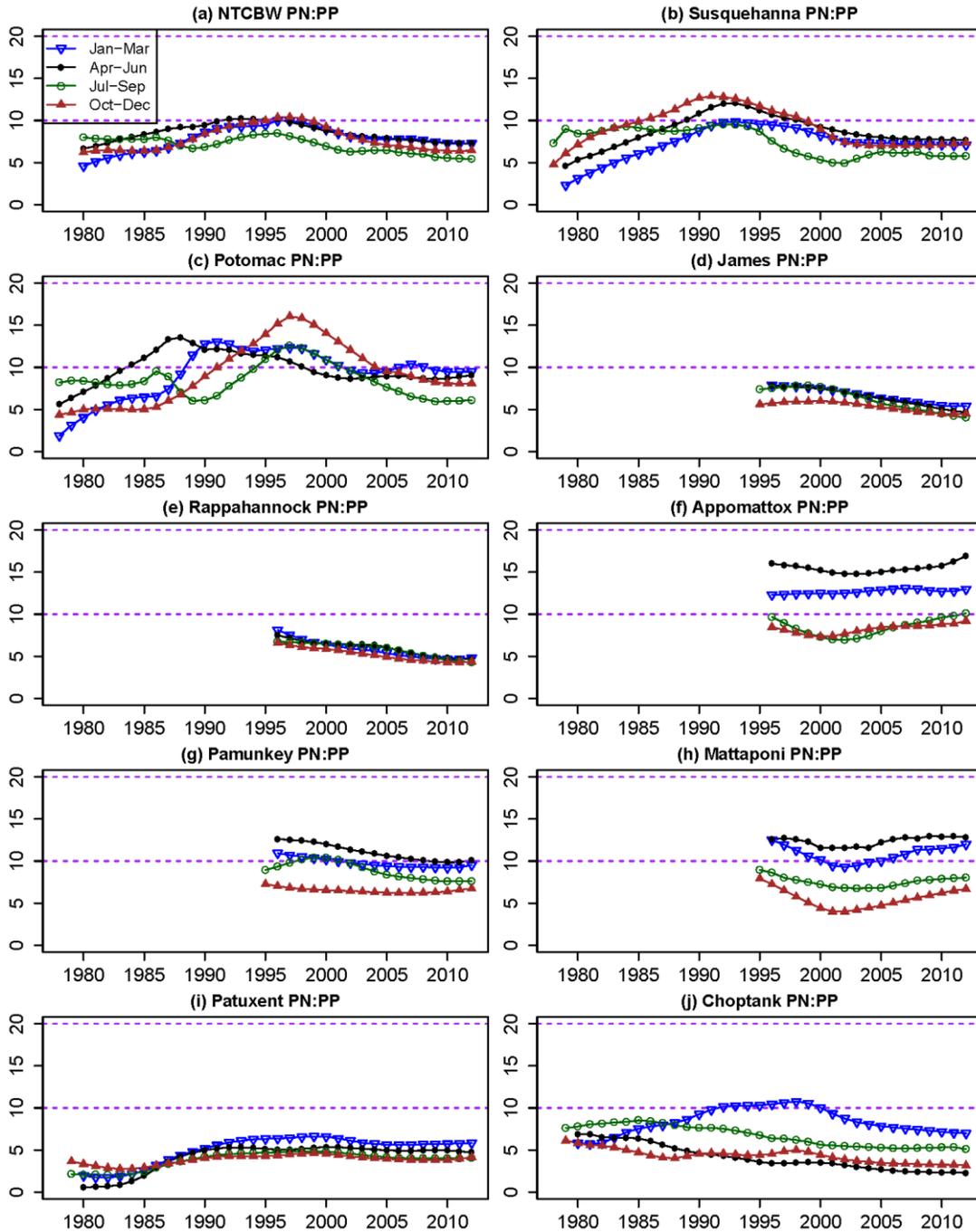


Figure D22. Seasonal PN:PP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW; sum of the nine rivers) and the nine individual tributaries, as calculated based on WRTDS *flow-normalized* estimates. Y-axis indicates the N:P ratio, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

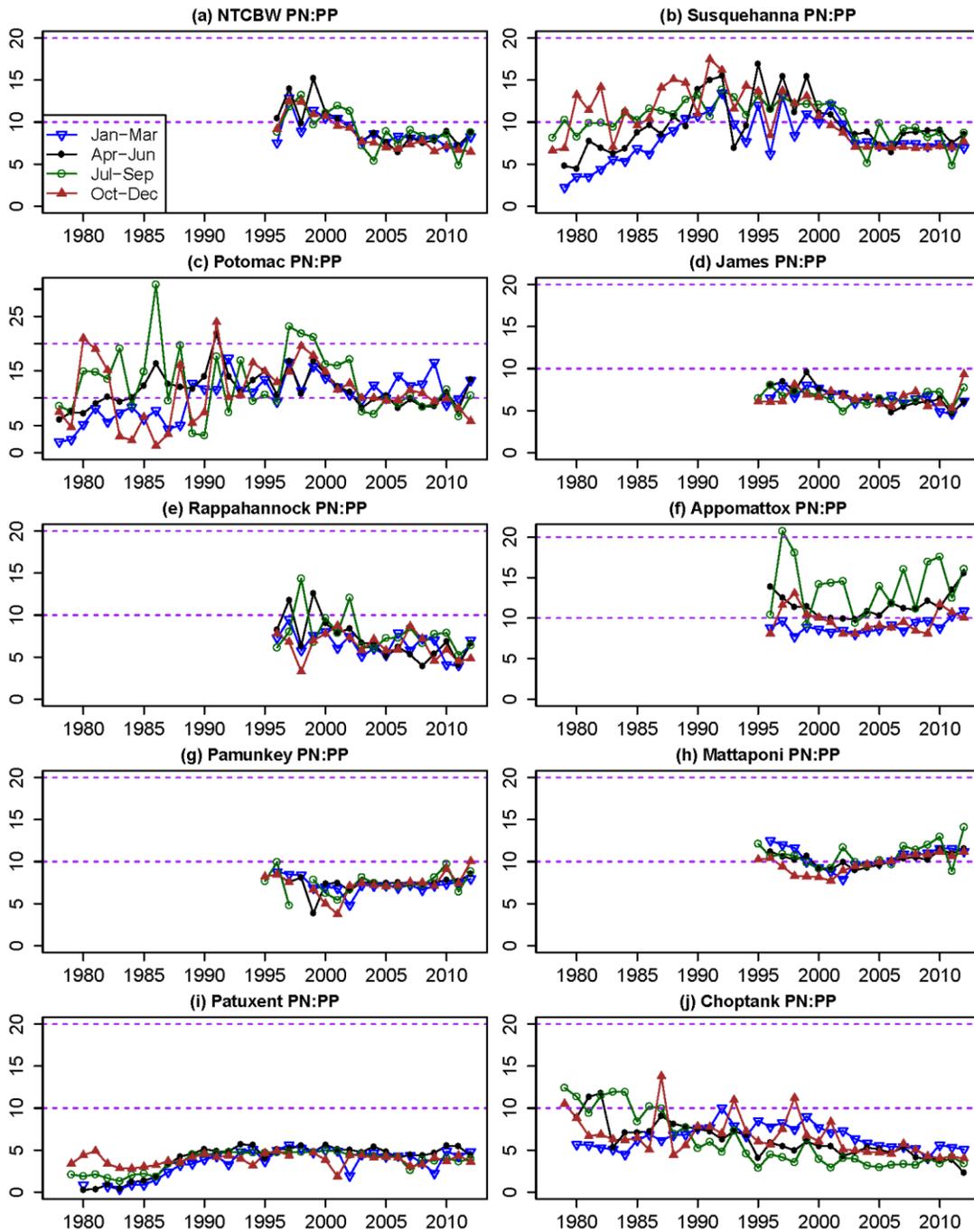


Figure D23. Seasonal PN:PP molar ratios in the nontidal Chesapeake Bay watershed (NTCBW; sum of the nine rivers) and the nine individual tributaries, as calculated based on WRTDS *true-condition* estimates. Y-axis indicates the N:P ratio, with dashed lines separating the three major categories of nutrient limitation: (1) N-limitation (ratio < 10), (2) P-limitation (ratio > 20), and (3) co-limitation by both N and P ( $10 \leq \text{ratio} \leq 20$ ).

## **Appendix D6. WRTDS model residual plots for Susquehanna River at Conowingo**

This appendix encloses the model residual plots for total nitrogen (TN), total phosphorus (TP), and suspended sediment (SS) for the Susquehanna River at Conowingo (Figures D24-D26). These plots are selected as representatives of residual plots for all 45 WRTDS model analyses made for data at the nine USGS River Input Monitoring stations. Each figure contains four graphs, which plot model residuals (observed log concentration minus estimated log concentration) against (a) estimated concentrations, (b) time, (c) discharges, and (d) months, respectively.

Note that on some plots, *e.g.*, Figure D25a, the vertical stripes correspond to left-censored concentration samples. As a general rule of WRTDS, such concentrations are treated as intervals and plotted with vertical lines. As a result, their corresponding estimates and residuals are also indicated with vertical lines graphically. (Readers are referred to the “Censored Data” section in Moyer *et al.* (2012) for details.) Another issue evident in some figures (*e.g.*, Figure D25a) is a clear pattern of diagonal line-up of some data points. This phenomenon relates to rounding issues of USGS sample recording and to the fact that the nature of “round-up” varied over time.

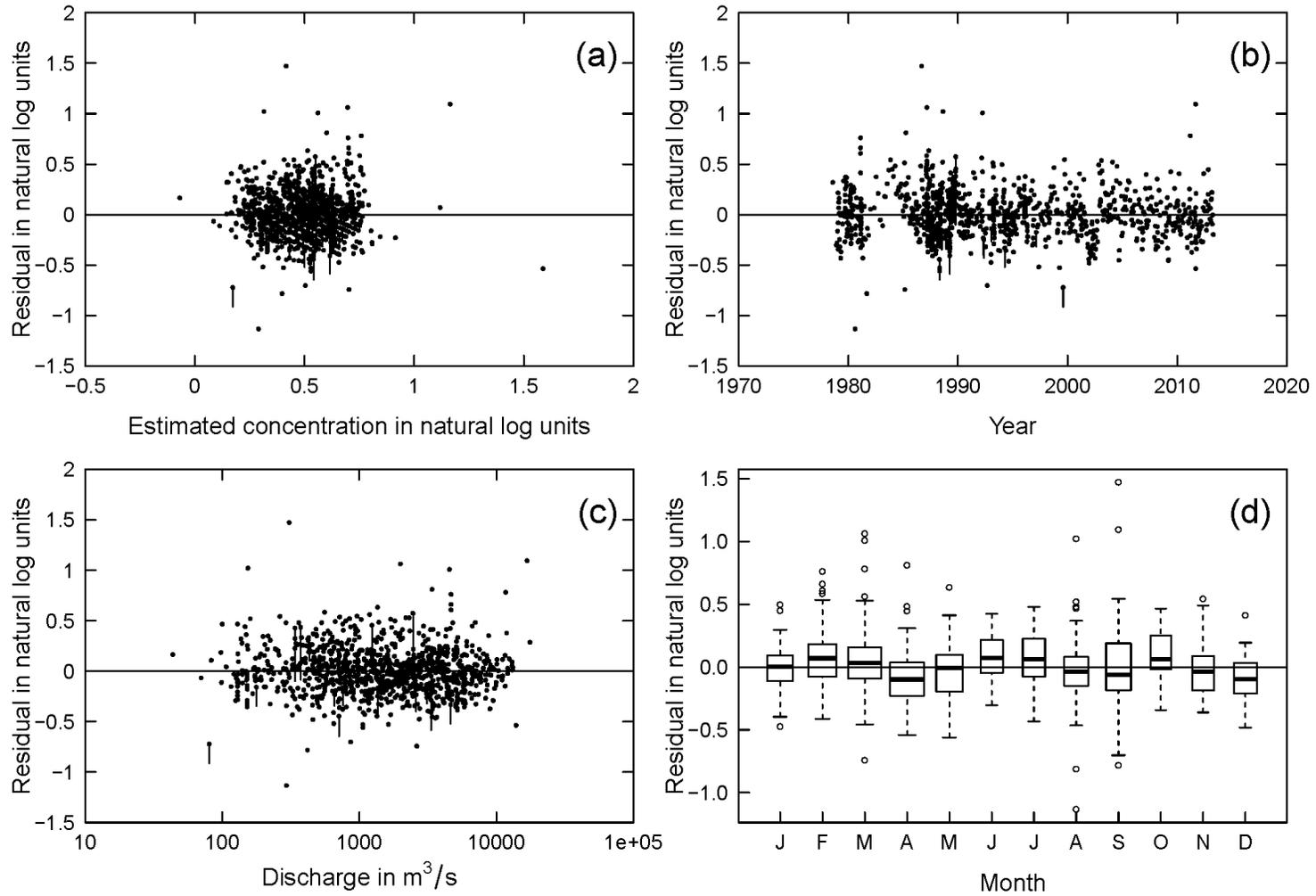


Figure D24. Model residual plot for total nitrogen in the Susquehanna River at Conowingo, MD, showing residual (in natural log units) as a function of estimated concentration (a), year (b), discharge (c) and month (d).

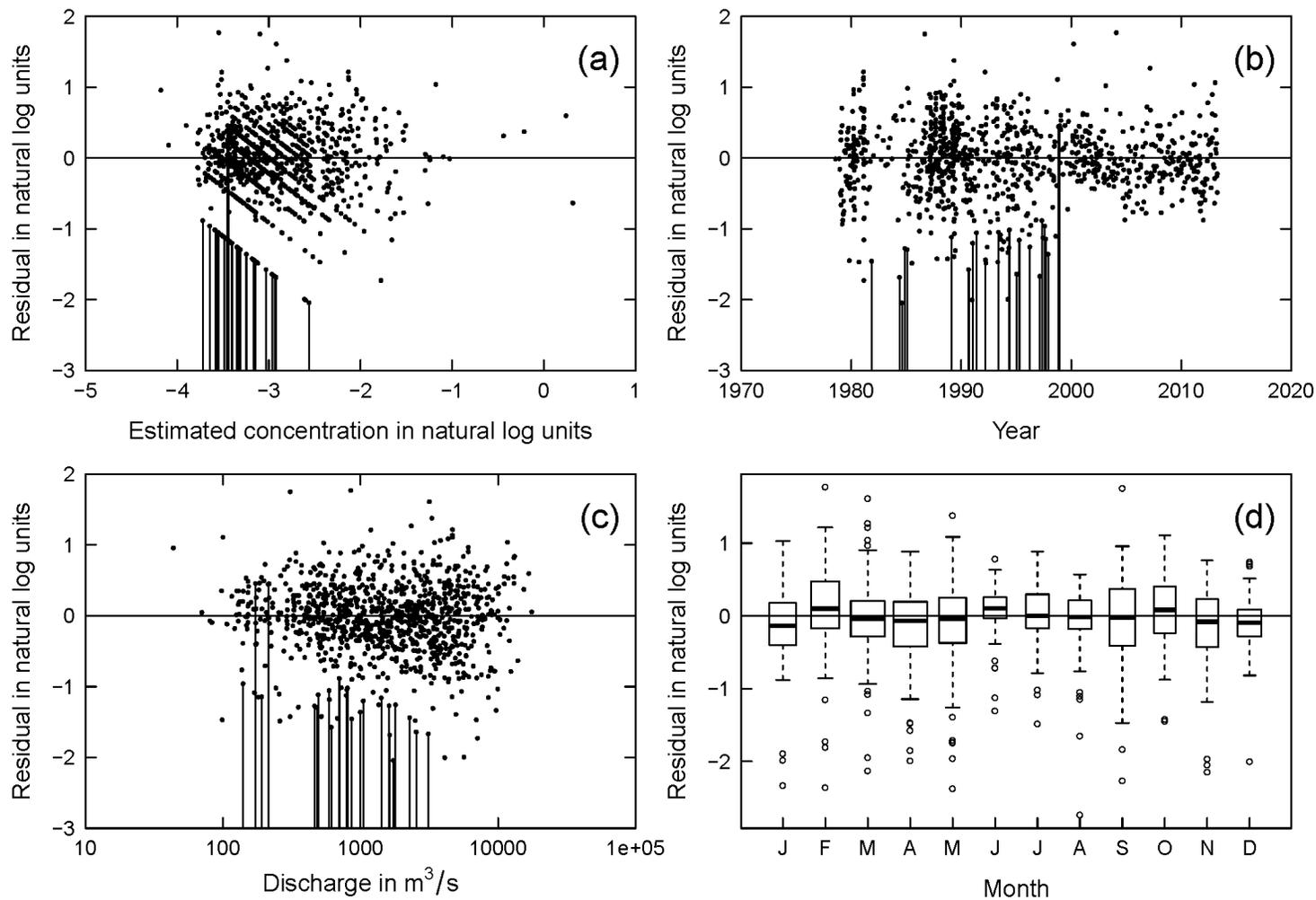


Figure D25. Model residual plot for total phosphorus in the Susquehanna River at Conowingo, MD, showing residual (in natural log units) as a function of estimated concentration (a), year (b), discharge (c) and month (d).

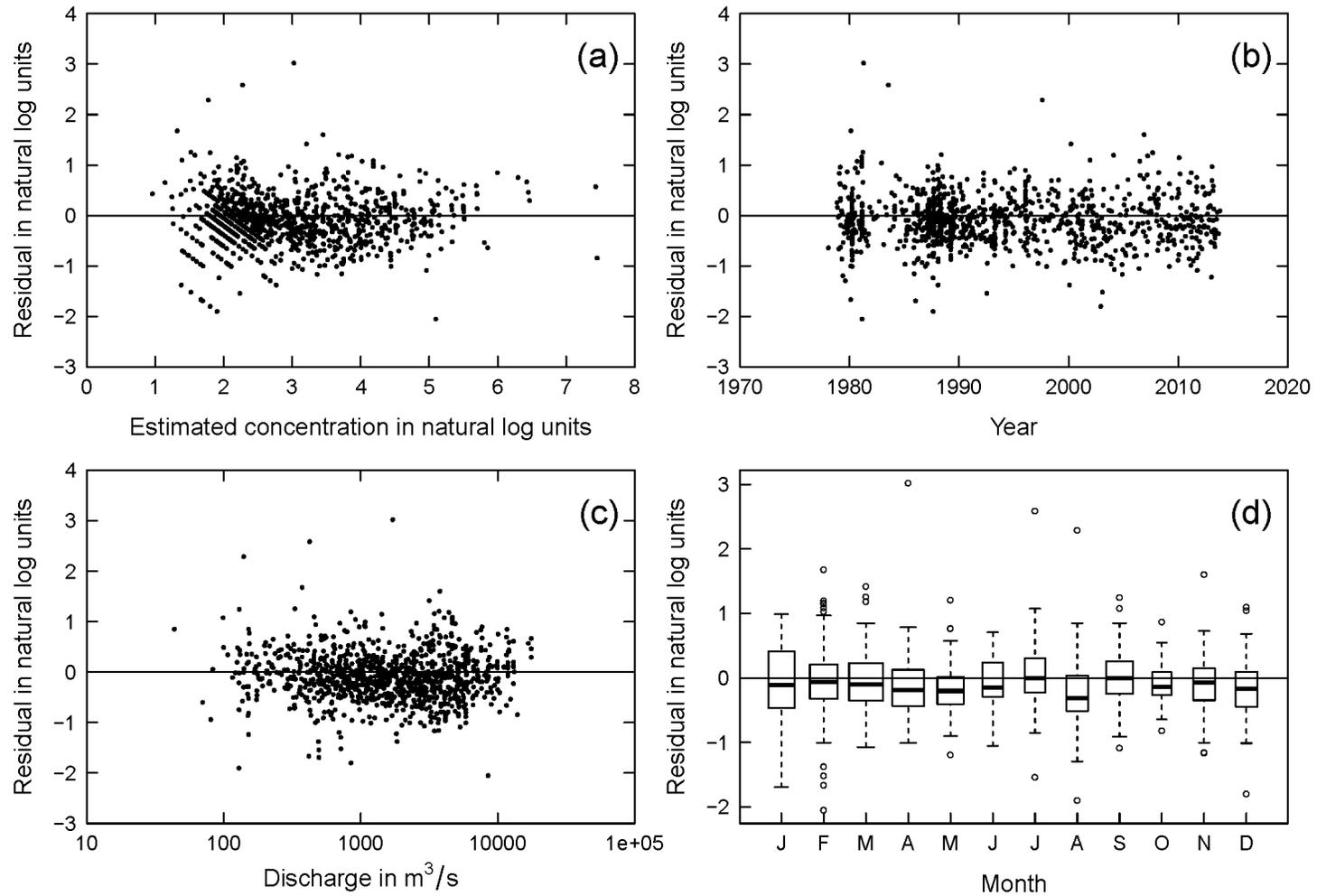


Figure D26. Model residual plot for suspended sediment in the Susquehanna River at Conowingo, MD, showing residual (in natural log units) as a function of estimated concentration (a), year (b), discharge (c) and month (d).

## Literature Cited

- Moyer, D. L., R. M. Hirsch and K. E. Hyer, 2012. Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed. U.S. Geological Survey Scientific Investigations Report 2012-5244, Reston, VA, p. 118. <http://pubs.usgs.gov/sir/2012/5244/>.
- Sprague, L. A., M. J. Langland, S. E. Yochum, R. E. Edwards, J. D. Blomquist, S. W. Phillips, G. W. Shenk and S. D. Preston, 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay Watershed. U.S. Geological Survey Water-Resources Investigations Report 00-4218, Richmond, VA, p. 109. [http://va.water.usgs.gov/online\\_pubs/WRIR/00-4218.htm](http://va.water.usgs.gov/online_pubs/WRIR/00-4218.htm).

**Appendix E. Supporting Information to Chapter 6**

**“What Can We Learn from Limited Data? Statistical Inferences and Uncertainties of Riverine Fluxes and Trends with Limited Sampling of Extreme-Flow Events”**

Table E1. Artificial concentration samples for Marietta and Conowingo during six major storm events between 1986 and 2014.

Site	Date	Daily flow (m <sup>3</sup> /s)	SS (mg/L)	TP (mg/L)	TN (mg/L)
Marietta <sup>a</sup>	1996/1/20	12,431	194	0.09	2.3
	1996/1/21	17,613	1200	0.29	2.8
	1996/1/21	17,613	1000	0.29	2.8
	1996/1/21	17,613	863	-	-
	2004/9/20	15,433	3680	1.17	20.0
	2011/9/8	16,764	2980	2.31	10.0
	2011/9/11	16,764	2980	2.31	10.0
Conowingo <sup>b</sup>	2006/6/29	10,081	1150	0.88	3.3

<sup>a</sup> assumed to have been measured at Marietta and have the same concentrations observed at Conowingo.

<sup>b</sup> assumed to have been measured at Conowingo and have the same concentrations observed at Marietta.

Table E2. Comparison of net scour estimates during six storm events between 1986 and 2014 under three sampling scenarios.

Storm	Date	Daily flow (m <sup>3</sup> /s)		Sampled (Y/N)?		SS fractional difference		TP fractional difference		TN fractional difference	
		Mar	Cono	Mar	Cono	(S <sub>2</sub> -S <sub>1</sub> )/S <sub>1</sub>	(S <sub>3</sub> -S <sub>1</sub> )/S <sub>1</sub>	(S <sub>2</sub> -S <sub>1</sub> )/S <sub>1</sub>	(S <sub>3</sub> -S <sub>1</sub> )/S <sub>1</sub>	(S <sub>2</sub> -S <sub>1</sub> )/S <sub>1</sub>	(S <sub>3</sub> -S <sub>1</sub> )/S <sub>1</sub>
						Sub-scour only <sup>a</sup>	Artificial data <sup>b</sup>	Sub-scour only <sup>a</sup>	Artificial data <sup>b</sup>	Sub-scour only <sup>a</sup>	Artificial data <sup>b</sup>
Storm 1	1993/4/1	10,760	11,610	Y	Y	39%	46%	28%	-14%	61%	15%
	1993/4/2	12,176	13,224	Y	Y	126%	95%	45%	-22%	76%	16%
	1993/4/3	12,205	13,054	Y	Y	116%	102%	46%	-25%	92%	20%
Storm 2 (Ice Jam)	1996/1/20	12,205	12,431	N	Y	92%	49%	-24%	-30%	50%	11%
	1996/1/21	15,744	17,613	N	Y	-207%	-44%	-84%	-100%	1412%	92%
	1996/1/22	10,987	12,120	Y	Y	-1124%	-450%	-48%	-54%	1170%	300%
Storm 3 (Ivan)	2004/9/20	14,073	15,433	N	Y	-128%	-77%	-155%	-95%	-96%	-90%
Storm 4	2006/6/29	11,412	11,412	Y	N	-190%	-74%	-214%	-105%	-96%	-82%
Storm 5	2011/3/12	11,978	11,723	Y	Y	-71%	0%	-64%	0%	-40%	-1%
	2011/9/8	14,866	16,764	N	Y	-108%	-44%	-96%	-53%	-92%	-61%
Storm 6 (TS Lee)	2011/9/9	17,443	20,077	N	N	-102%	-41%	-84%	-48%	-91%	-57%
	2011/9/10	13,224	13,960	Y	Y	-122%	-60%	-108%	-63%	-89%	-71%

<sup>a</sup> For the sub-scour data scenario, all samples from flows above 11,300 m<sup>3</sup>/s were removed at both locations.

<sup>b</sup> For the artificial data scenario, artificial samples were inserted at the non-sampled or “N” location of highlighted dates.

Notations:

Mar (M): Marietta; Cono (C): Conowingo; green cell: fractional change < -30%; red cell: fractional change > +30%.

S<sub>1</sub>: Net scour estimates, as estimated using original samples.

S<sub>2</sub>: Net scour estimates, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

S<sub>3</sub>: Net scour estimates, as estimated using original samples plus artificial samples on five dates (1996/1/20, 1996/1/21, 2004/9/20, 2011/9/8, 2011/09/11) for Marietta and one date (2006/6/29) for Conowingo; see Table E1 for details.

Table E3. Comparison of SS load estimates during six storm events in 1986-2014 under three sampling scenarios.

Storm	Date	Daily flow (m <sup>3</sup> /s)		Sampled (Y/N)?		SS fractional difference			
		Mar	Cono	Mar	Cono	(M <sub>2</sub> -M <sub>1</sub> )/M <sub>1</sub>	(M <sub>3</sub> -M <sub>1</sub> )/M <sub>1</sub>	(C <sub>2</sub> -C <sub>1</sub> )/C <sub>1</sub>	(C <sub>3</sub> -C <sub>1</sub> )/C <sub>1</sub>
						Sub-scour only (-47 samples) <sup>a</sup>	Artificial data (+7 samples) <sup>b</sup>	Sub-scour only (-60 samples) <sup>a</sup>	Artificial data (+1 sample) <sup>b</sup>
Storm 1	1993/4/1	10,760	11,610	Y	Y	-13%	12%	-28%	0%
	1993/4/2	12,176	13,224	Y	Y	-15%	13%	-33%	0%
	1993/4/3	12,205	13,054	Y	Y	-15%	13%	-32%	0%
Storm 2 ( <i>Ice Jam</i> )	1996/1/20	12,205	12,431	N	Y	-10%	15%	-46%	0%
	1996/1/21	15,744	17,613	N	Y	-10%	17%	-63%	0%
	1996/1/22	10,987	12,120	Y	Y	-10%	15%	-45%	0%
Storm 3 ( <i>Ivan</i> )	2004/9/20	14,073	15,433	N	Y	54%	147%	-75%	-12%
Storm 4	2006/6/29	11,412	11,412	Y	N	-3%	21%	-69%	-13%
Storm 5	2011/3/12	11,978	11,723	Y	Y	13%	0%	-12%	0%
	2011/9/8	14,866	16,764	N	Y	16%	142%	-80%	-4%
	2011/9/9	17,443	20,077	N	N	11%	164%	-81%	-4%
Storm 6 ( <i>TS Lee</i> )	2011/9/10	13,224	13,960	Y	Y	22%	120%	-77%	-3%

<sup>a</sup> For the flow-censored data scenario, all samples from flows above 11,300 m<sup>3</sup>/s and all samples collected during the entire hydrographs of the six storm events were removed at both locations.

<sup>b</sup> For the artificial data scenario, artificial samples were inserted at the non-sampled or “N” location of highlighted dates.

**Notations:**

Mar (M): Marietta; Cono (C): Conowingo; green cell: fractional change < -30%; red cell: fractional change > +30%.

C<sub>1</sub>: WRTDS daily loading at Conowingo, as estimated using original samples.

C<sub>2</sub>: WRTDS daily loading at Conowingo, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

C<sub>3</sub>: WRTDS daily loading at Conowingo, as estimated using original samples plus artificial sample on one date (2006/6/29).

M<sub>1</sub>: WRTDS daily loading at Marietta, as estimated using original samples.

M<sub>2</sub>: WRTDS daily loading at Marietta, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

M<sub>3</sub>: WRTDS daily loading at Marietta, as estimated using original samples plus artificial samples on five dates (1996/1/20, 1996/1/21, 2004/9/20, 2011/9/8, 2011/09/11); see Table E1.

Table E4. Comparison of TP load estimates during six storm events in 1986-2014 under three sampling scenarios.

Storm	Date	Daily flow (m <sup>3</sup> /s)		Sampled (Y/N)?		TP fractional difference			
		Mar	Cono	Mar	Cono	(M <sub>2</sub> -M <sub>1</sub> )/M <sub>1</sub>	(M <sub>3</sub> -M <sub>1</sub> )/M <sub>1</sub>	(C <sub>2</sub> -C <sub>1</sub> )/C <sub>1</sub>	(C <sub>3</sub> -C <sub>1</sub> )/C <sub>1</sub>
						Sub-scour only (-47 samples) <sup>a</sup>	Artificial data (+6 samples) <sup>b</sup>	Sub-scour only (-46 samples) <sup>a</sup>	Artificial data (+1 sample) <sup>b</sup>
Storm 1	1993/4/1	10,760	11,610	Y	Y	-14%	-5%	-31%	0%
	1993/4/2	12,176	13,224	Y	Y	-16%	-6%	-35%	0%
	1993/4/3	12,205	13,054	Y	Y	-16%	-6%	-34%	0%
Storm 2 ( <i>Ice Jam</i> )	1996/1/20	12,205	12,431	N	Y	-19%	-20%	-4%	0%
	1996/1/21	15,744	17,613	N	Y	-25%	-33%	4%	0%
	1996/1/22	10,987	12,120	Y	Y	-18%	-16%	-5%	0%
Storm 3 ( <i>Ivan</i> )	2004/9/20	14,073	15,433	N	Y	21%	25%	-23%	-5%
Storm 4	2006/6/29	11,412	11,412	Y	N	-19%	12%	-45%	-4%
Storm 5	2011/3/12	11,978	11,723	Y	Y	6%	0%	-17%	0%
Storm 6 ( <i>TS Lee</i> )	2011/9/8	14,866	16,764	N	Y	26%	103%	-53%	-1%
	2011/9/9	17,443	20,077	N	N	21%	120%	-54%	-2%
	2011/9/10	13,224	13,960	Y	Y	32%	87%	-50%	-1%

<sup>a</sup> For the flow-censored data scenario, all samples from flows above 11,300 m<sup>3</sup>/s and all samples collected during the entire hydrographs of the six storm events were removed at both locations.

<sup>b</sup> For the artificial data scenario, artificial samples were inserted at the non-sampled or “N” location of highlighted dates.

Notations:

Mar (M): Marietta; Cono (C): Conowingo; green cell: fractional change < -30%; red cell: fractional change > +30%.

C<sub>1</sub>: WRTDS daily loading at Conowingo, as estimated using original samples.

C<sub>2</sub>: WRTDS daily loading at Conowingo, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

C<sub>3</sub>: WRTDS daily loading at Conowingo, as estimated using original samples plus artificial sample on one date (2006/6/29).

M<sub>1</sub>: WRTDS daily loading at Marietta, as estimated using original samples.

M<sub>2</sub>: WRTDS daily loading at Marietta, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

M<sub>3</sub>: WRTDS daily loading at Marietta, as estimated using original samples plus artificial samples on five dates (1996/1/20, 1996/1/21, 2004/9/20, 2011/9/8, 2011/09/11); see Table E1.

Table E5. Comparison of TN load estimates during six storm events in 1986-2014 under three sampling scenarios.

Storm	Date	Daily flow (m <sup>3</sup> /s)		Sampled (Y/N)?		TN fractional difference			
		Mar	Cono	Mar	Cono	(M <sub>2</sub> -M <sub>1</sub> )/M <sub>1</sub>	(M <sub>3</sub> -M <sub>1</sub> )/M <sub>1</sub>	(C <sub>2</sub> -C <sub>1</sub> )/C <sub>1</sub>	(C <sub>3</sub> -C <sub>1</sub> )/C <sub>1</sub>
						Sub-scour only (-47 samples) <sup>a</sup>	Artificial data (+6 samples) <sup>b</sup>	Sub-scour only (-46 samples) <sup>a</sup>	Artificial data (+1 sample) <sup>b</sup>
Storm 1	1993/4/1	10,760	11,610	Y	Y	0%	2%	-9%	0%
	1993/4/2	12,176	13,224	Y	Y	2%	3%	-11%	0%
	1993/4/3	12,205	13,054	Y	Y	2%	3%	-10%	0%
Storm 2 ( <i>Ice Jam</i> )	1996/1/20	12,205	12,431	N	Y	7%	3%	-7%	0%
	1996/1/21	15,744	17,613	N	Y	17%	2%	-11%	0%
	1996/1/22	10,987	12,120	Y	Y	4%	3%	-7%	0%
Storm 3 ( <i>Ivan</i> )	2004/9/20	14,073	15,433	N	Y	-28%	354%	-83%	-10%
Storm 4	2006/6/29	11,412	11,412	Y	N	-23%	49%	-54%	-12%
Storm 5	2011/3/12	11,978	11,723	Y	Y	0%	0%	-14%	0%
Storm 6 ( <i>TS Lee</i> )	2011/9/8	14,866	16,764	N	Y	-11%	179%	-70%	-4%
	2011/9/9	17,443	20,077	N	N	-12%	226%	-75%	-4%
	2011/9/10	13,224	13,960	Y	Y	-8%	146%	-63%	-3%

<sup>a</sup> For the flow-censored data scenario, all samples from flows above 11,300 m<sup>3</sup>/s and all samples collected during the entire hydrographs of the six storm events were removed at both locations.

<sup>b</sup> For the artificial data scenario, artificial samples were inserted at the non-sampled or “N” location of highlighted dates.

**Notations:**

Mar (M): Marietta; Cono (C): Conowingo; green cell: fractional change < -30%; red cell: fractional change > +30%.

C<sub>1</sub>: WRTDS daily loading at Conowingo, as estimated using original samples.

C<sub>2</sub>: WRTDS daily loading at Conowingo, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

C<sub>3</sub>: WRTDS daily loading at Conowingo, as estimated using original samples plus artificial sample on one date (2006/6/29).

M<sub>1</sub>: WRTDS daily loading at Marietta, as estimated using original samples.

M<sub>2</sub>: WRTDS daily loading at Marietta, as estimated using flow-censored samples that exclude all scour-flow samples and all samples collected during the entire hydrographs of the six storm events.

M<sub>3</sub>: WRTDS daily loading at Marietta, as estimated using original samples plus artificial samples on five dates (1996/1/20, 1996/1/21, 2004/9/20, 2011/9/8, 2011/09/11); see Table E1.

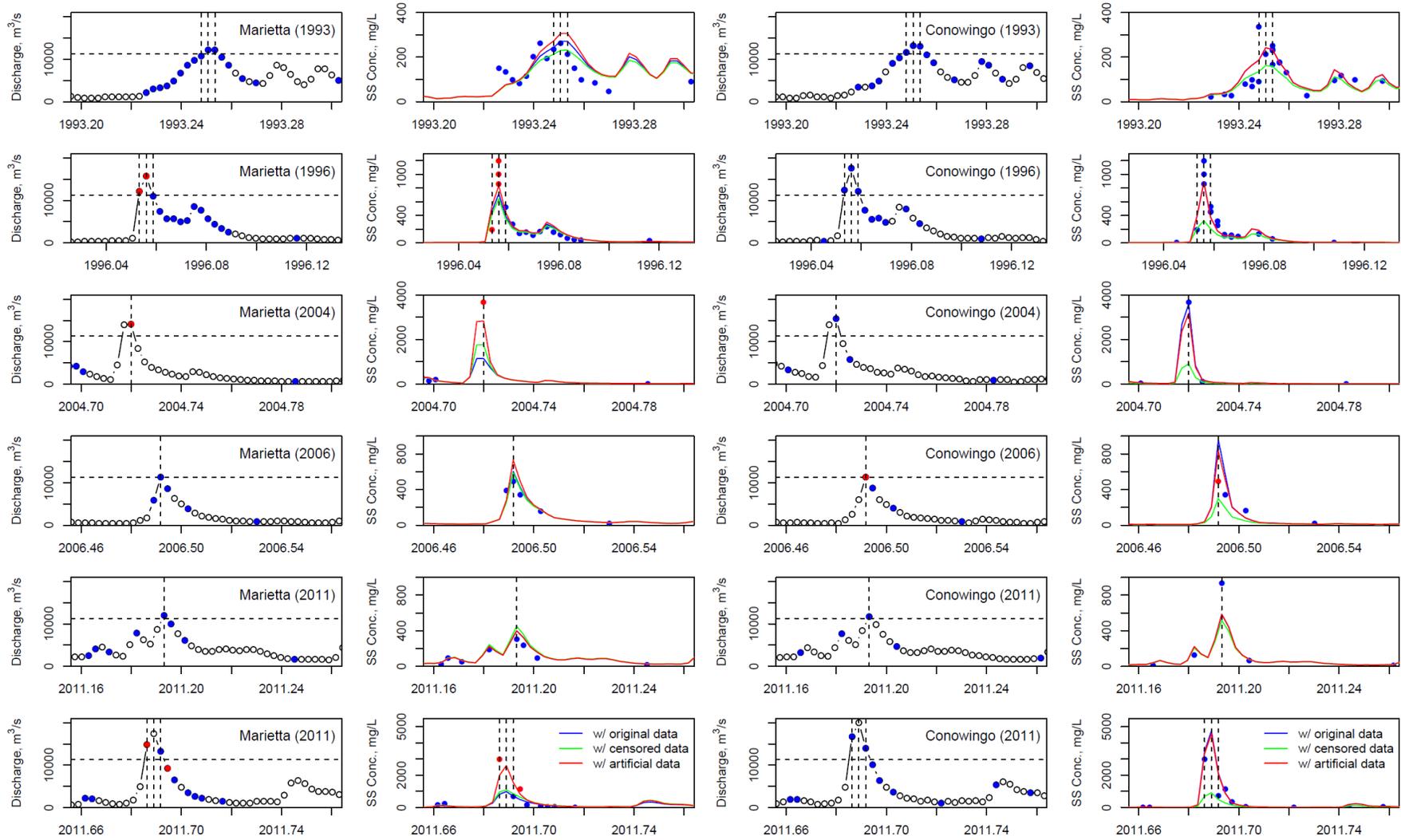


Figure E1. Comparison of hydrograph and suspended sediment (SS) chemograph for Marietta and Conowingo during six major storm events in 1986-2014 and WRTDS concentration estimates obtained based on three different sampling scenarios. Horizontal dashed lines indicate the scour threshold of 11,300 m<sup>3</sup>/s. Vertical dashed lines indicate the key dates during each event, for which the differences of WRTDS estimates among the three sampling scenarios are quantified and summarized in Table E2 and Table E3.

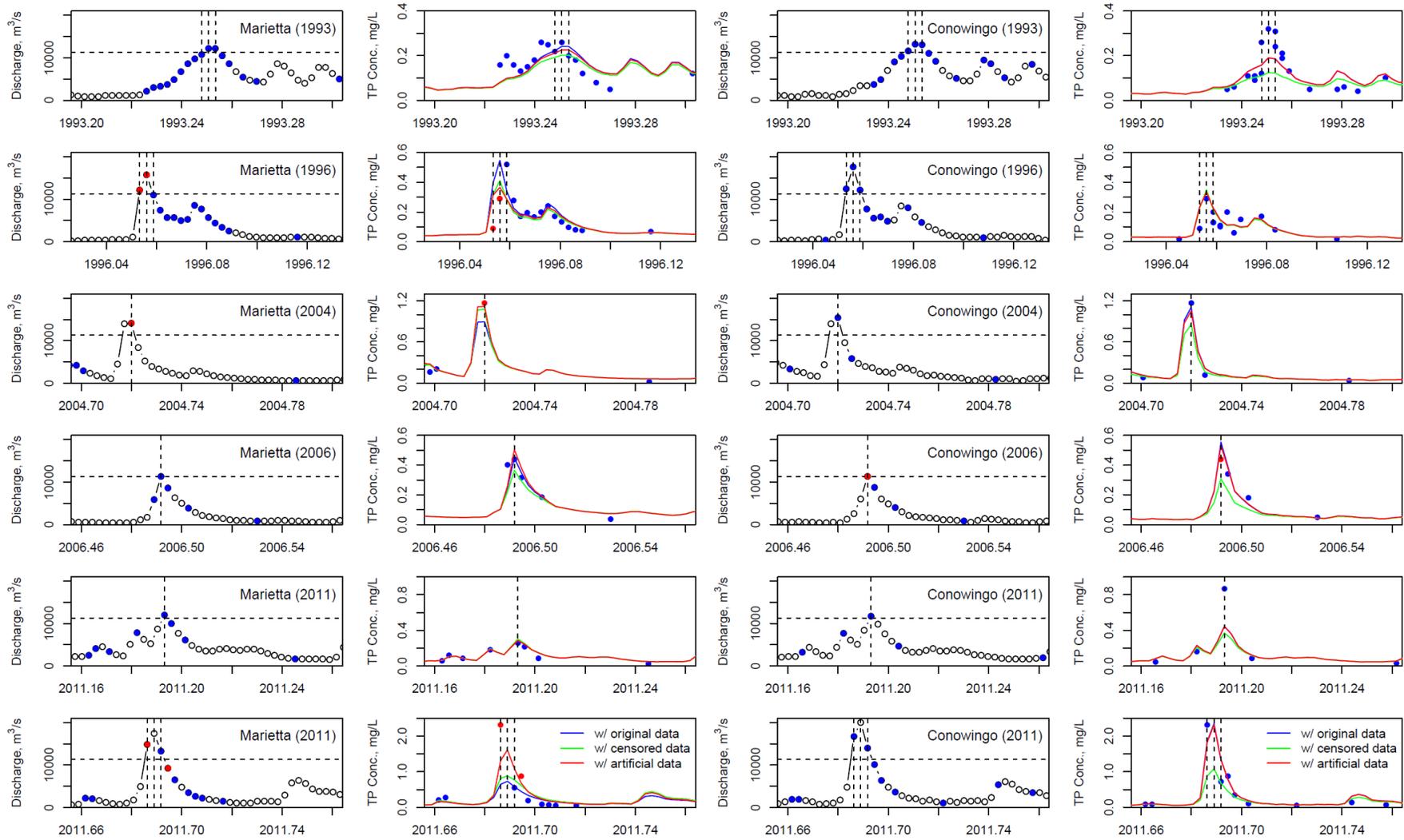


Figure E2. Comparison of the hydrograph and total phosphorus (TP) chemograph for Marietta and Conowingo during six major storm events in 1986-2014 and WRTDS concentration estimates obtained based on three different sampling scenarios. Horizontal dashed lines indicate the scour threshold of 11,300 m<sup>3</sup>/s. Vertical dashed lines indicate the key dates during each event, for which the differences of WRTDS estimates among the three sampling scenarios are quantified and summarized in Table E2 and Table E4.

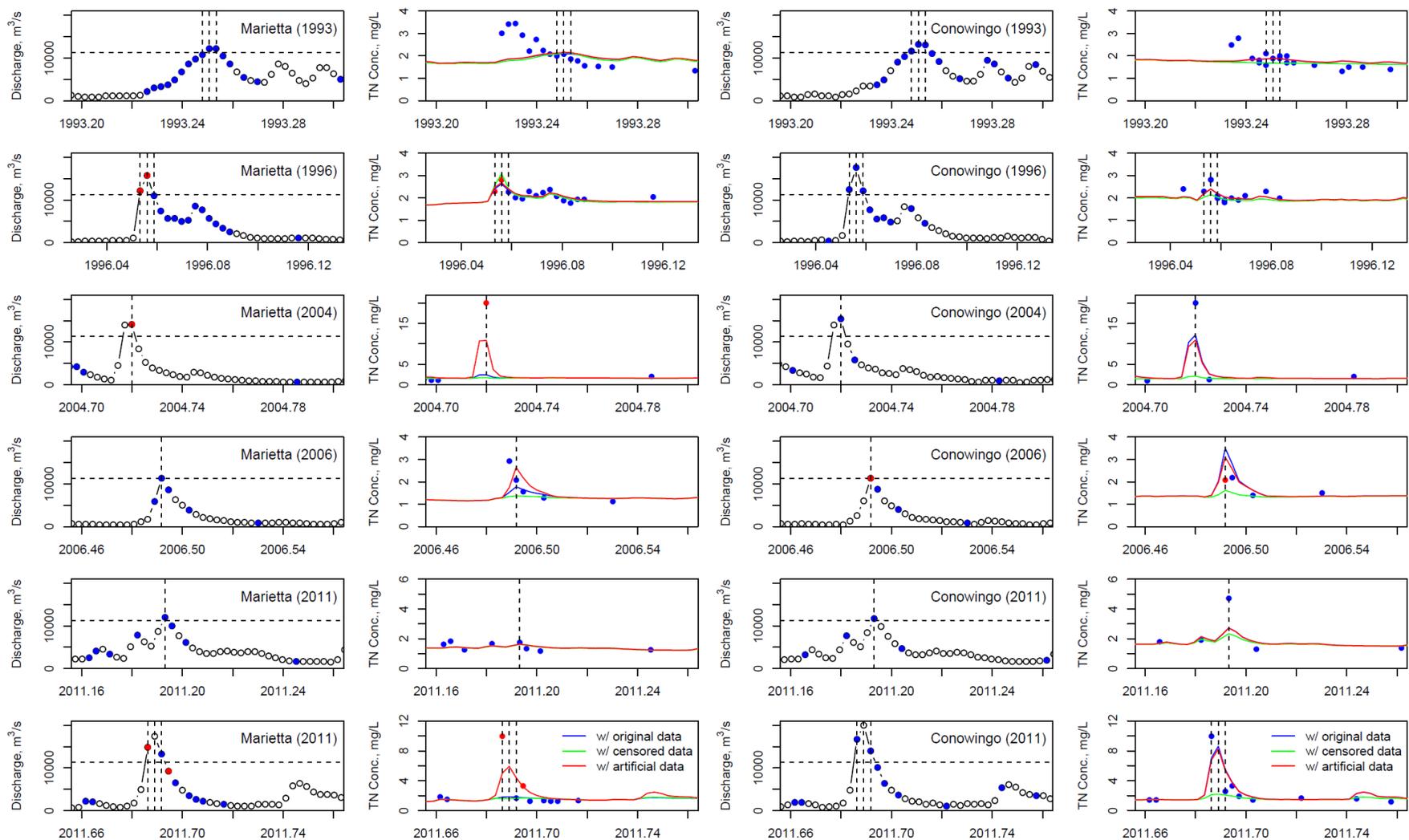


Figure E3. Comparison of the hydrograph and total nitrogen (TN) chemograph for Marietta and Conowingo during six major storm events in 1986-2014 and WRTDS concentration estimates obtained based on three different sampling scenarios. Horizontal dashed lines indicate the scour threshold of 11,300 m<sup>3</sup>/s. Vertical dashed lines indicate the key dates during each event, for which the differences of WRTDS estimates among the three sampling scenarios are quantified and summarized in Table E2 and Table E5.

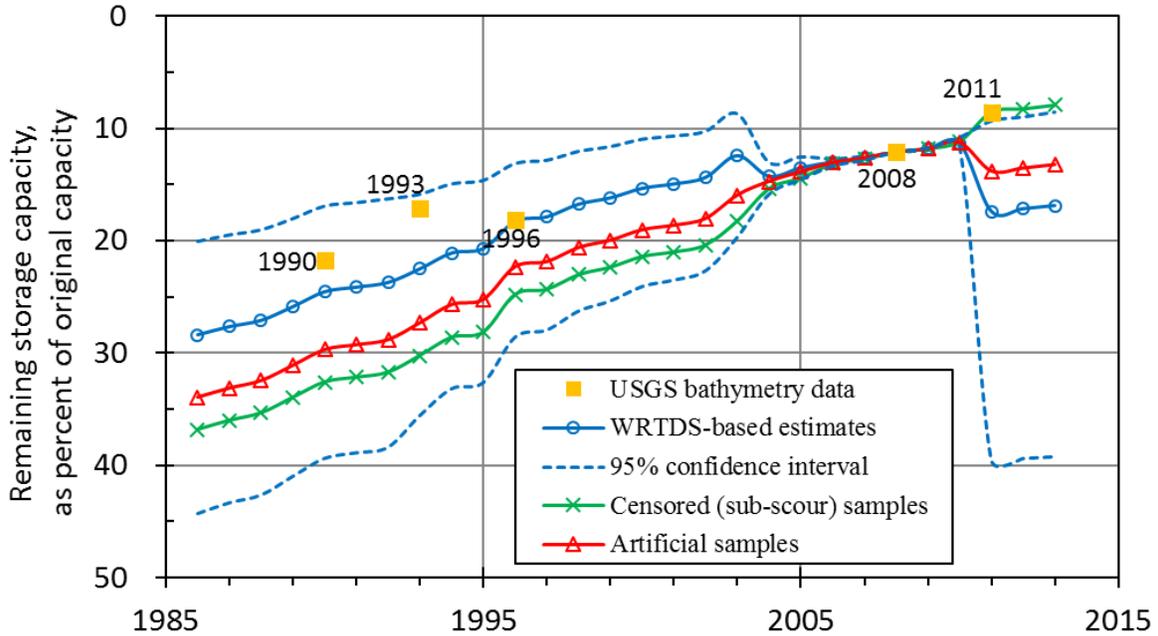
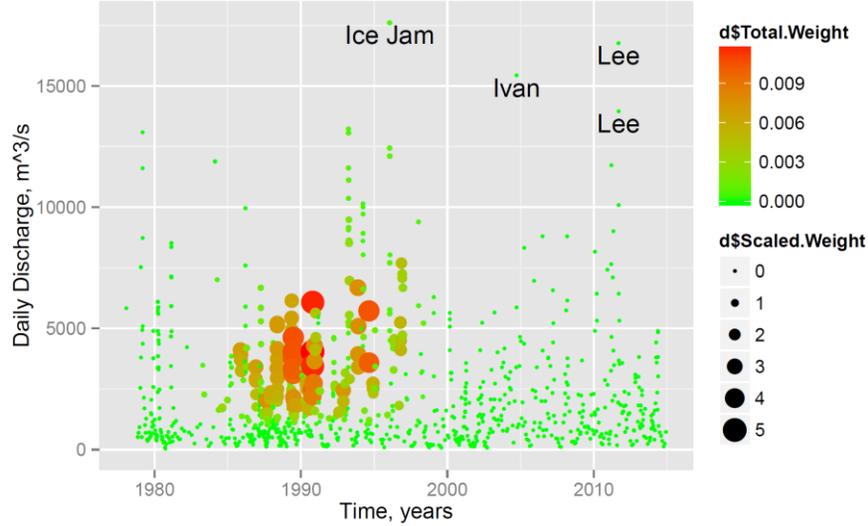
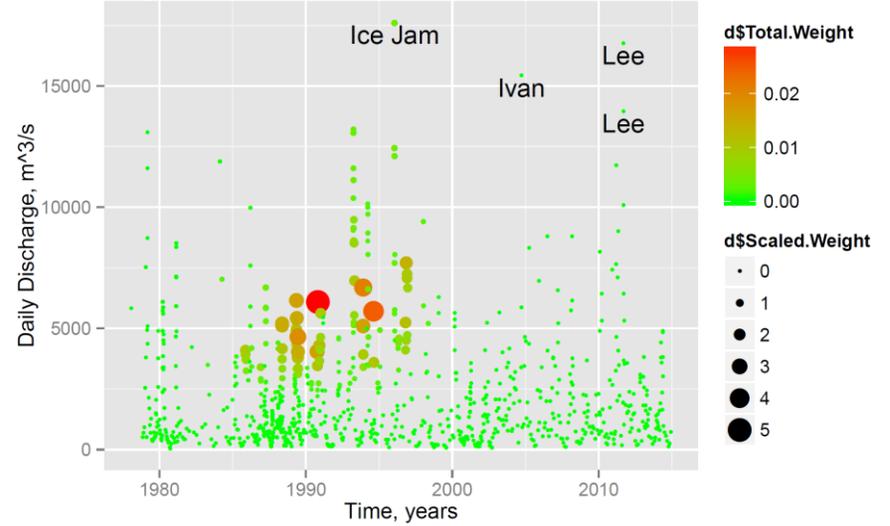


Figure E4. Estimated remaining sediment storage capacity in Conowingo Reservoir for the period between 1987 and 2013, expressed as percentage of the original storage capacity in 1929 (modified after Zhang *et al.* (2016)). The curves were reconstructed by assuming that the 2008 bathymetry-based estimate of sediment capacity (Langland, 2015) is correct and by estimating pre-2008 and post-2008 storages using different scenarios of WRTDS annual mass-balance estimates. Specifically, solid blue line is based on WRTDS estimates using the original concentration record. Dashed blue lines indicate 95% confidence intervals, as obtained from uncertainty analyses of 100 WRTDS model runs on bootstrap samples. Solid green line is based on WRTDS estimates using flow-censored concentration records at both sites, which eliminates all concentration samples with discharge above the scour threshold (*i.e.*, 11,300 m<sup>3</sup>/s) and all samples collected during the entire hydrographs of the six storm events. Solid red line is based on WRTDS estimates using concentration records with added artificial samples at both sites. (See Figure E2 and Table E3 for details.)

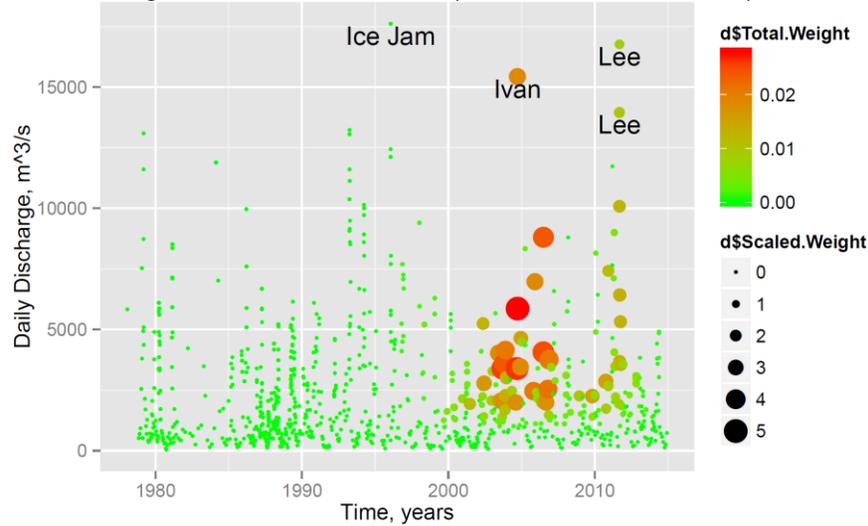
Total Weight Relative to 1990/09/01 (5,663 m<sup>3</sup>/s; 200,000 cfs)



Total Weight Relative to 1990/09/01 (14,158 m<sup>3</sup>/s; 500,000 cfs)



Total Weight Relative to 2005/09/01 (5,663 m<sup>3</sup>/s; 200,000 cfs)



Total Weight Relative to 2005/09/01 (14,158 m<sup>3</sup>/s; 500,000 cfs)

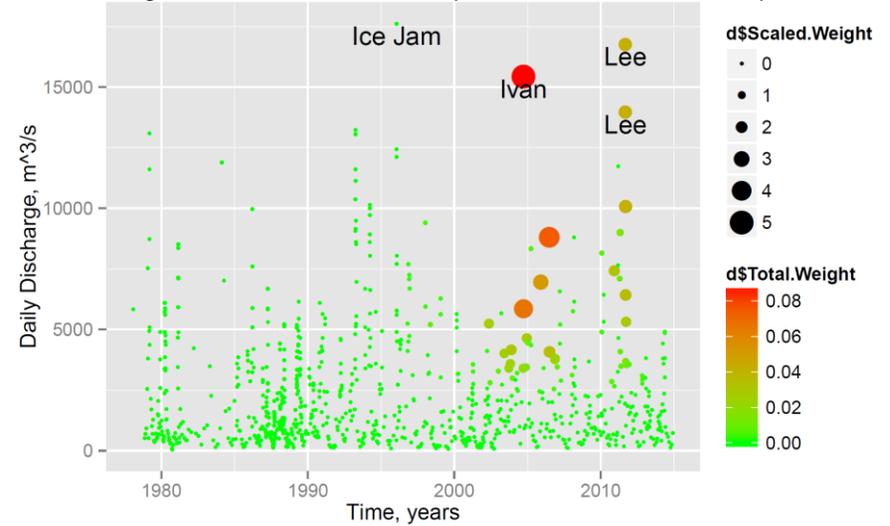


Figure E5. WRTDS-assigned weights to each available sediment concentration sample at Conowingo in models developed for four hypothetical cases of time and discharge. To aid comparison among points, weight values are represented by both the size and the color.

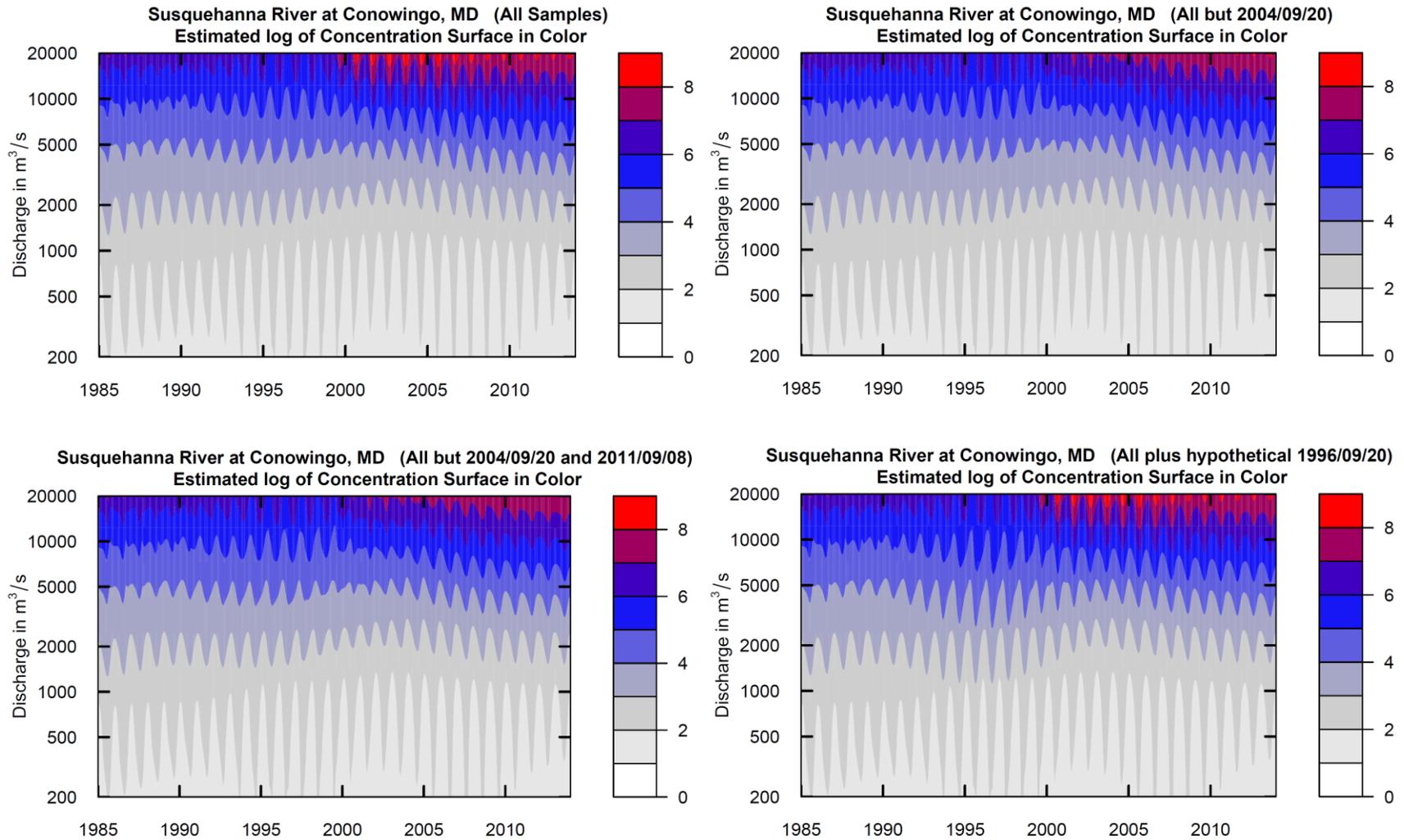


Figure E6. WRTDS-regression surface for suspended sediment at Conowingo, showing estimated log(concentration) as a function of time and discharge. These WRTDS surfaces were obtained based on four different scenarios of concentration sampling.

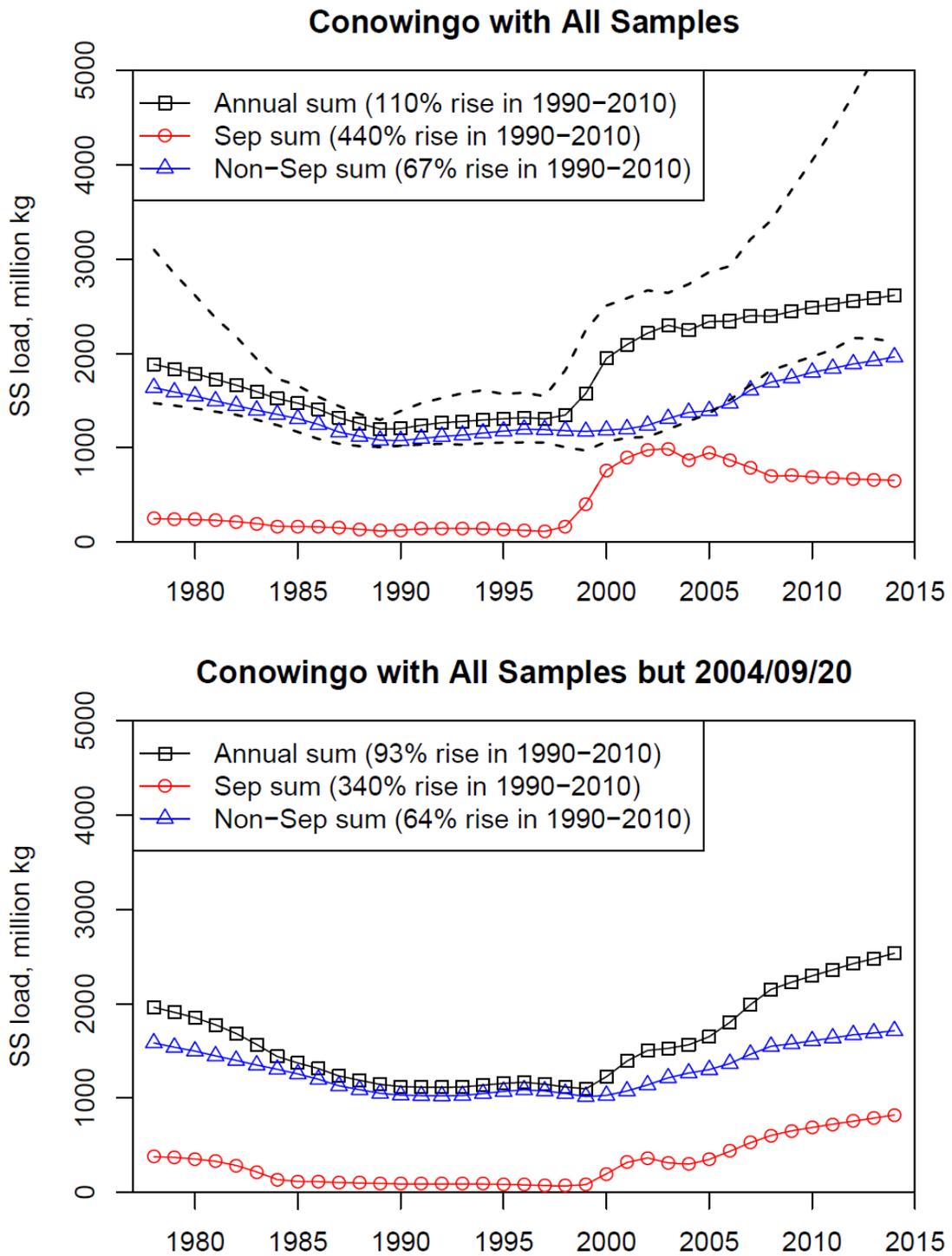


Figure E7. WRTDS flow-normalized trends of suspended sediment load at Conowingo, as obtained based on two different scenarios of concentration sampling, *i.e.*, one with the 2004/09/20 Hurricane Ivan sample (top panel) and one without it (bottom panel).

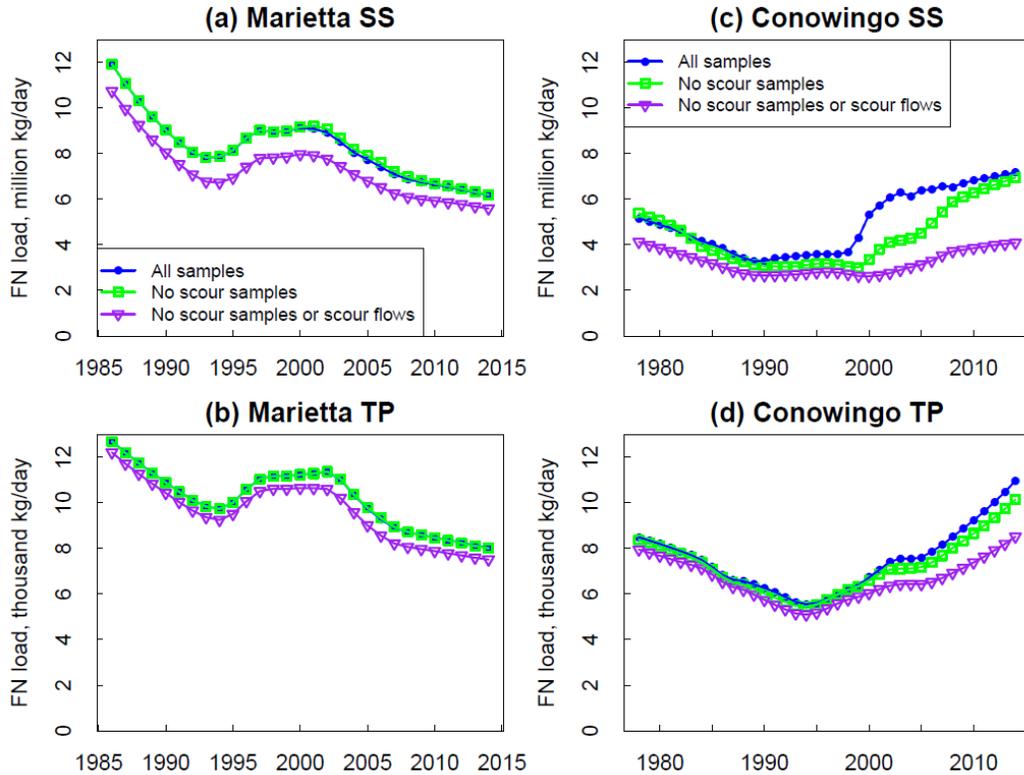


Figure E8. Flow-normalized annual trends of suspended sediment (SS) and total phosphorus (TP) loadings at Marietta (reservoir inlet) and Conowingo (reservoir outlet), as obtained based on three different scenarios of sampling data. Solid blue line is based on the original concentration record. Solid green line is based on a flow-censored concentration record, which eliminates all concentration samples with discharge  $> 11,300 \text{ m}^3/\text{s}$ . Solid purple line is based on censored records for both concentration and discharge, which eliminates all concentration samples with discharge  $> 11,300 \text{ m}^3/\text{s}$  and replaces all discharges  $> 11,300 \text{ m}^3/\text{s}$  with a nominal but small value, *i.e.*,  $20 \text{ m}^3/\text{s}$ .

### Literature Cited

Langland, M. J., 2015. Sediment transport and capacity change in three reservoirs, Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey Open-File Report 2014-1235, Reston, VA, p. 18.  
<http://dx.doi.org/10.3133/ofr20141235>.

Zhang, Q., R. M. Hirsch and W. P. Ball, 2016. Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Sedimentation. *Environ. Sci. Technol.* 50:1877-1886, DOI: 10.1021/acs.est.5b04073.

**Appendix F. Supporting Information to Chapter 8**

**“Non-stationary Concentration-Discharge Relationships: A Synthesis of Nutrient  
and Sediment Patterns in the Major Tributaries to Chesapeake Bay”**

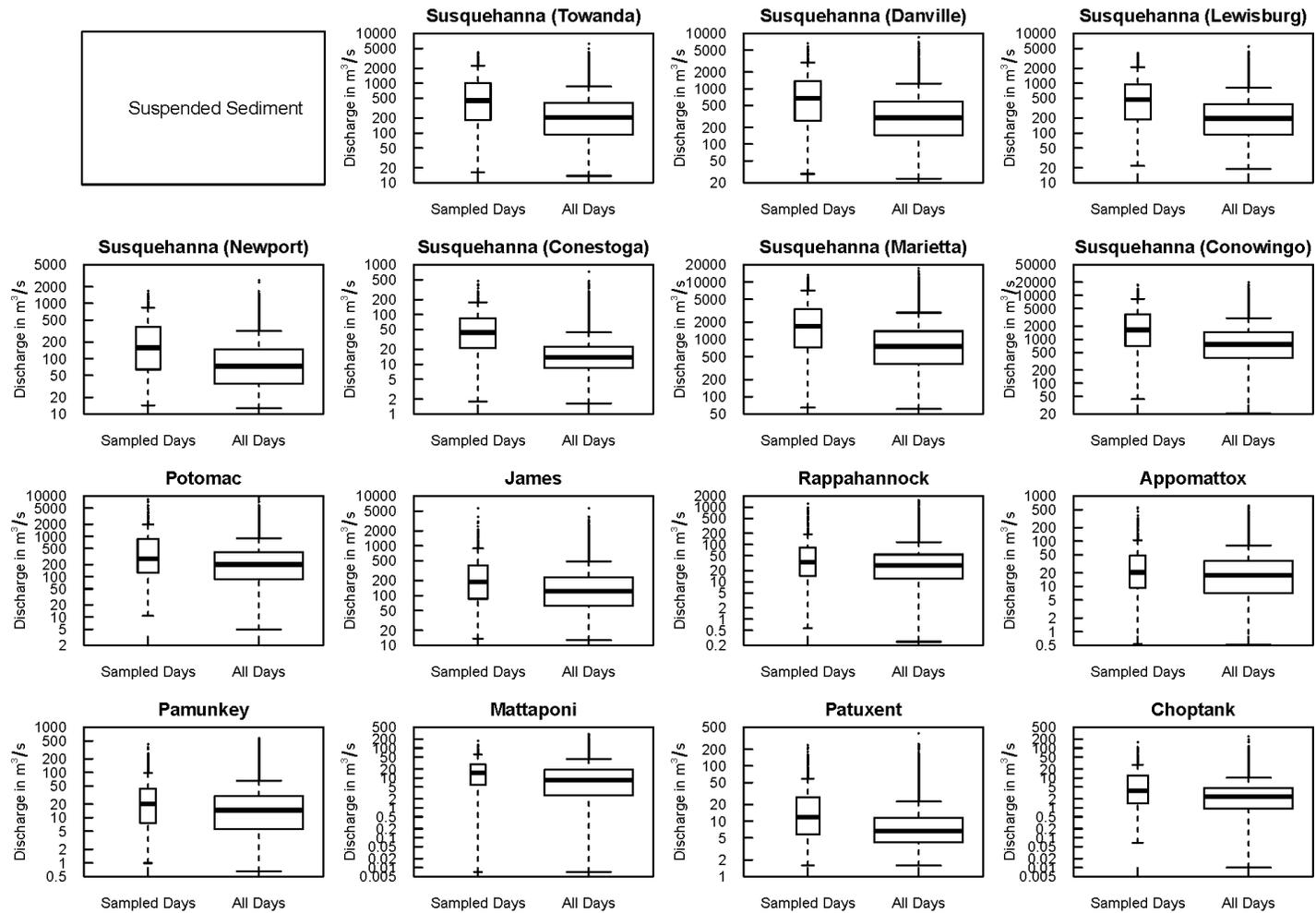


Figure F1. Distributions of discharge on days with *suspended sediment* (SS) samples and discharge on all days at the 15 Chesapeake sites in the period of record.

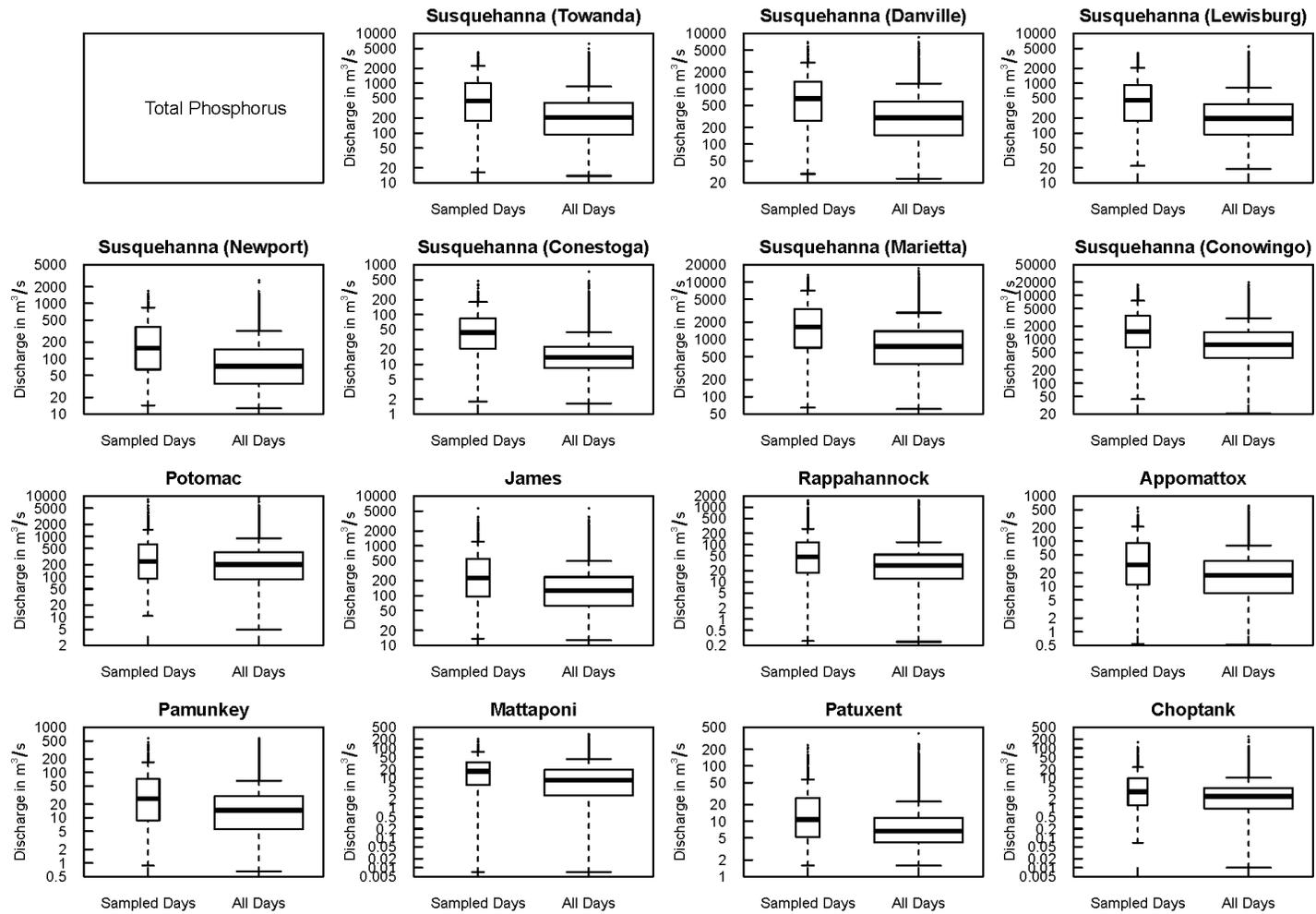


Figure F2. Distributions of discharge on days with *total phosphorus* (TP) samples and discharge on all days at the 15 Chesapeake sites in the period of record.

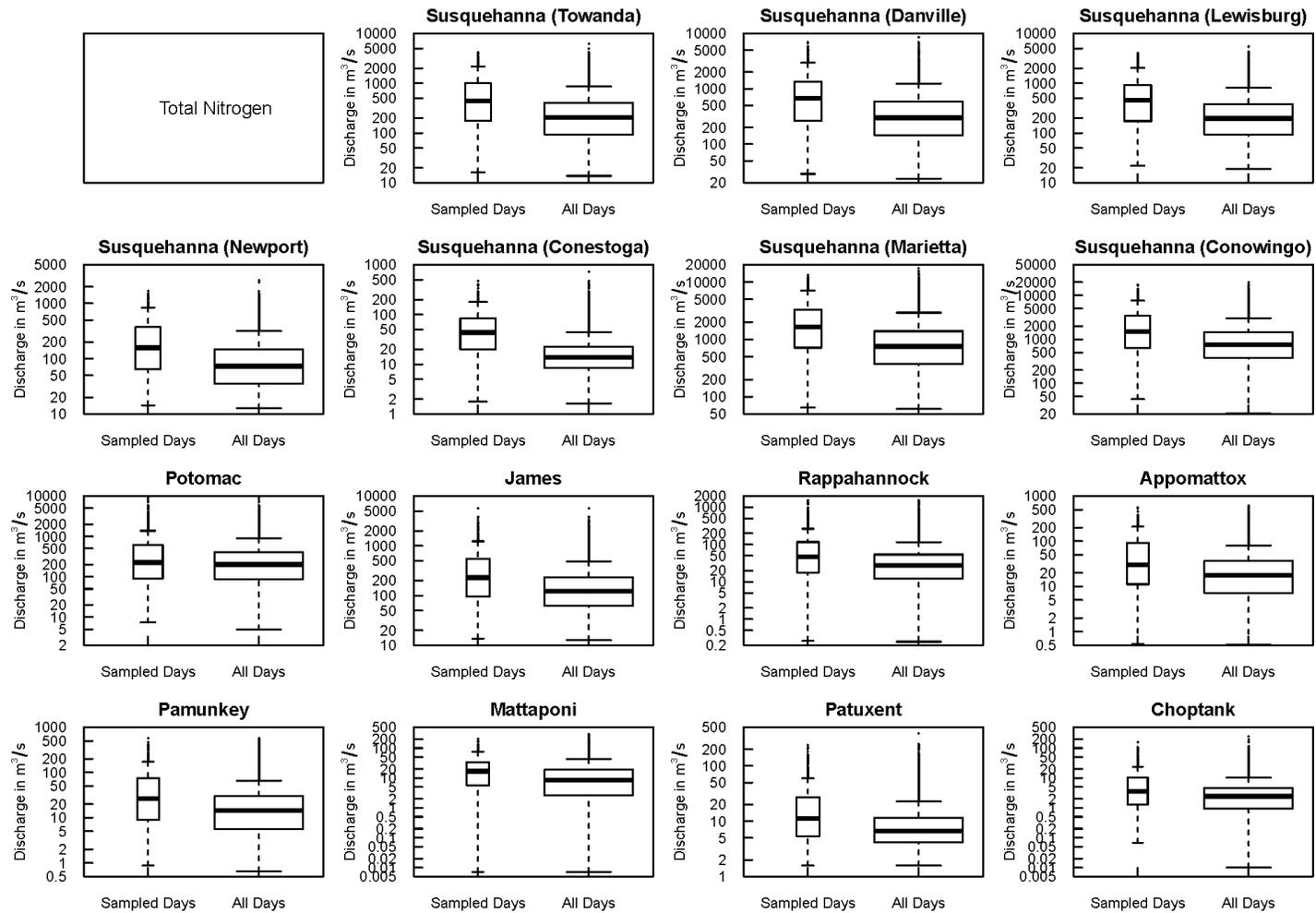


Figure F3. Distributions of discharge on days with *total nitrogen* (TN) samples and discharge on all days at the 15 Chesapeake sites in the period of record.

**Appendix G. Supporting Information to Chapter 9**

**“Improving Riverine Constituent Concentration and Flux Estimation by  
Accounting for Antecedent Discharge Conditions”**

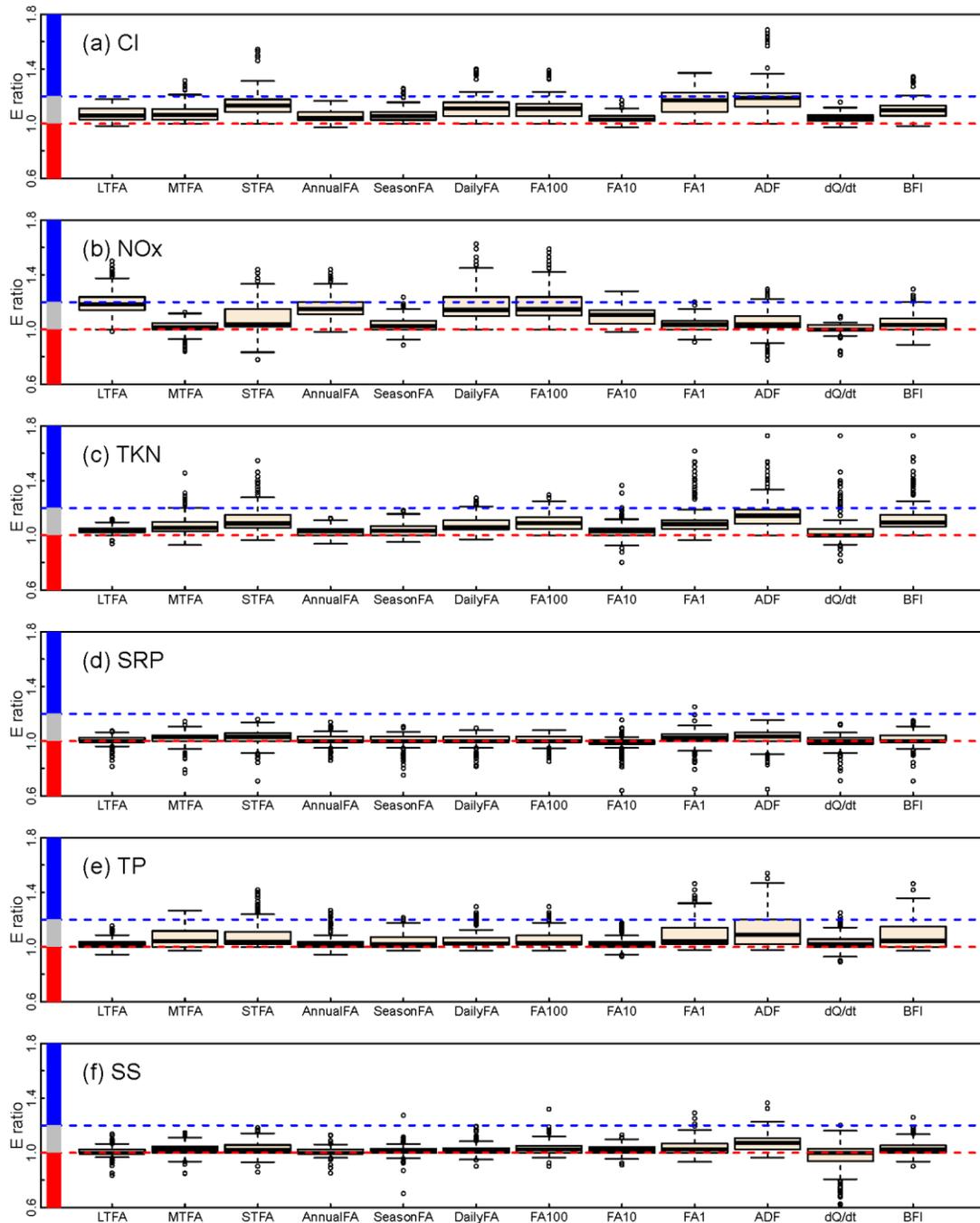


Figure G1. Boxplots of ratio between the modified Nash-Sutcliffe efficiency ( $E$ ) of each modified model and  $E$  of the original model for the nine sites for (a) Cl, (b) NO<sub>x</sub>, (c) TKN, (d) SRP, (e) TP, and (f) SS. The ratios are divided into three regions: (1) major improvement (ratio > 1.2; blue bars), (2) minor improvement (1 < ratio < 1.2; grey bars), and (3) inferior performance (ratio < 1.0; red bars). Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy B*.

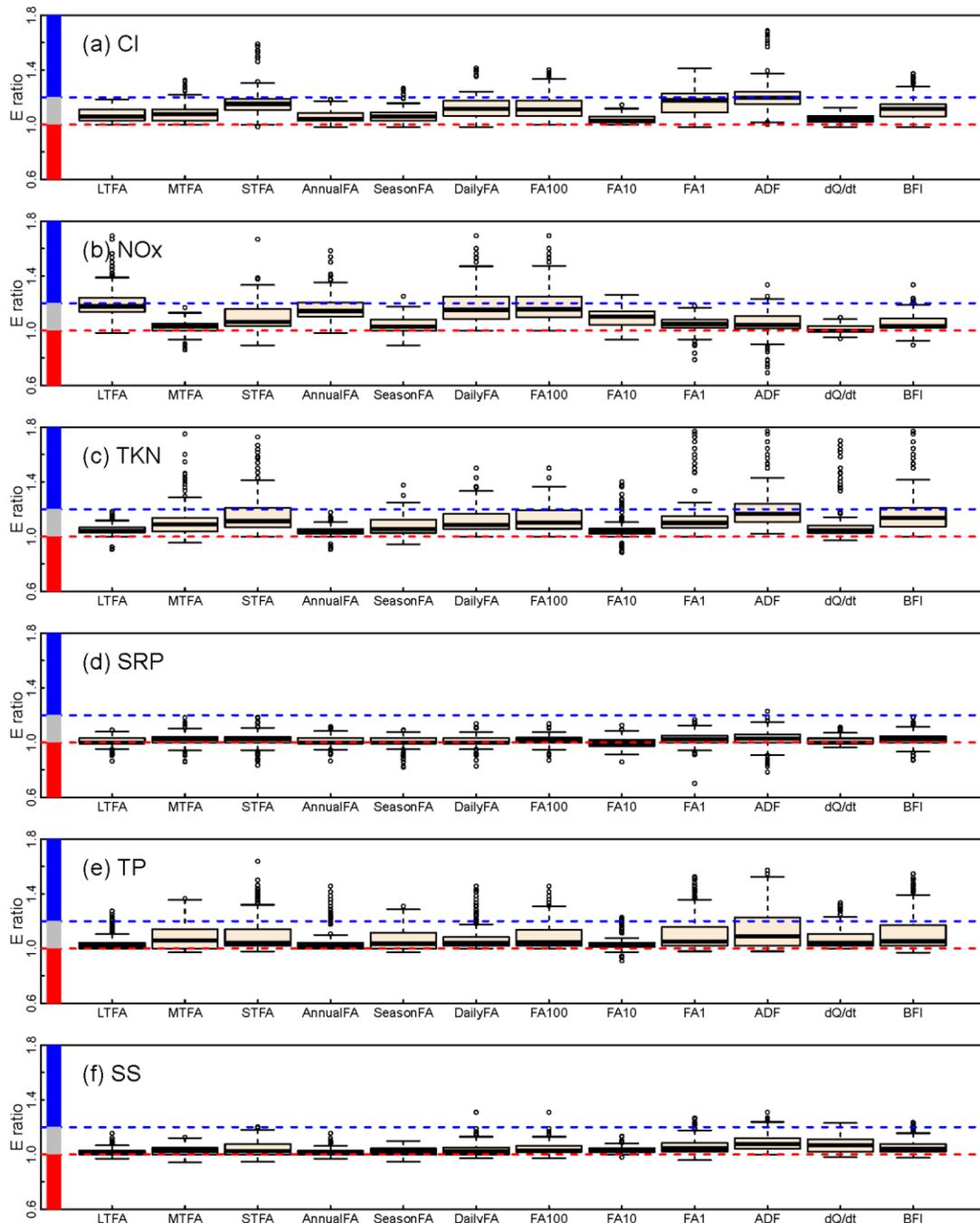


Figure G2. Boxplots of ratio between the modified Nash-Sutcliffe efficiency ( $E$ ) of each modified model and  $E$  of the original model for the nine sites for (a) Cl, (b)  $\text{NO}_x$ , (c) TKN, (d) SRP, (e) TP, and (f) SS. The ratios are divided into three regions: (1) major improvement (ratio  $> 1.2$ ; blue bars), (2) minor improvement ( $1 < \text{ratio} < 1.2$ ; grey bars), and (3) inferior performance (ratio  $< 1.0$ ; red bars). Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy C*.

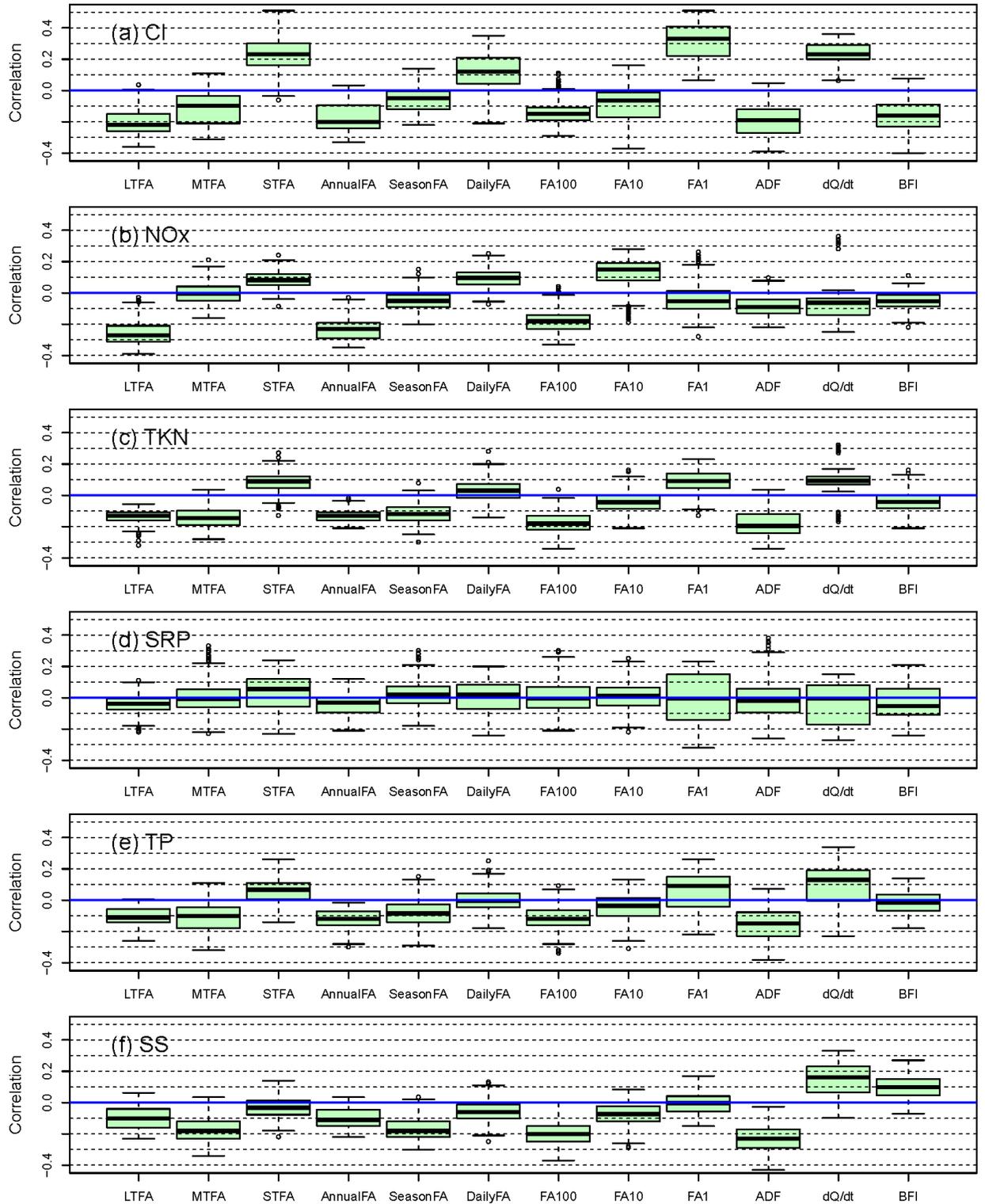


Figure G3. Boxplots of Spearman's correlation coefficient ( $\gamma$ ) between residuals from the original WRTDS model and each of the twelve proposed flow variables. Each boxplot represents 30 replicates at all 9 sites (*i.e.*, 270 cases) under *sampling strategy A*.

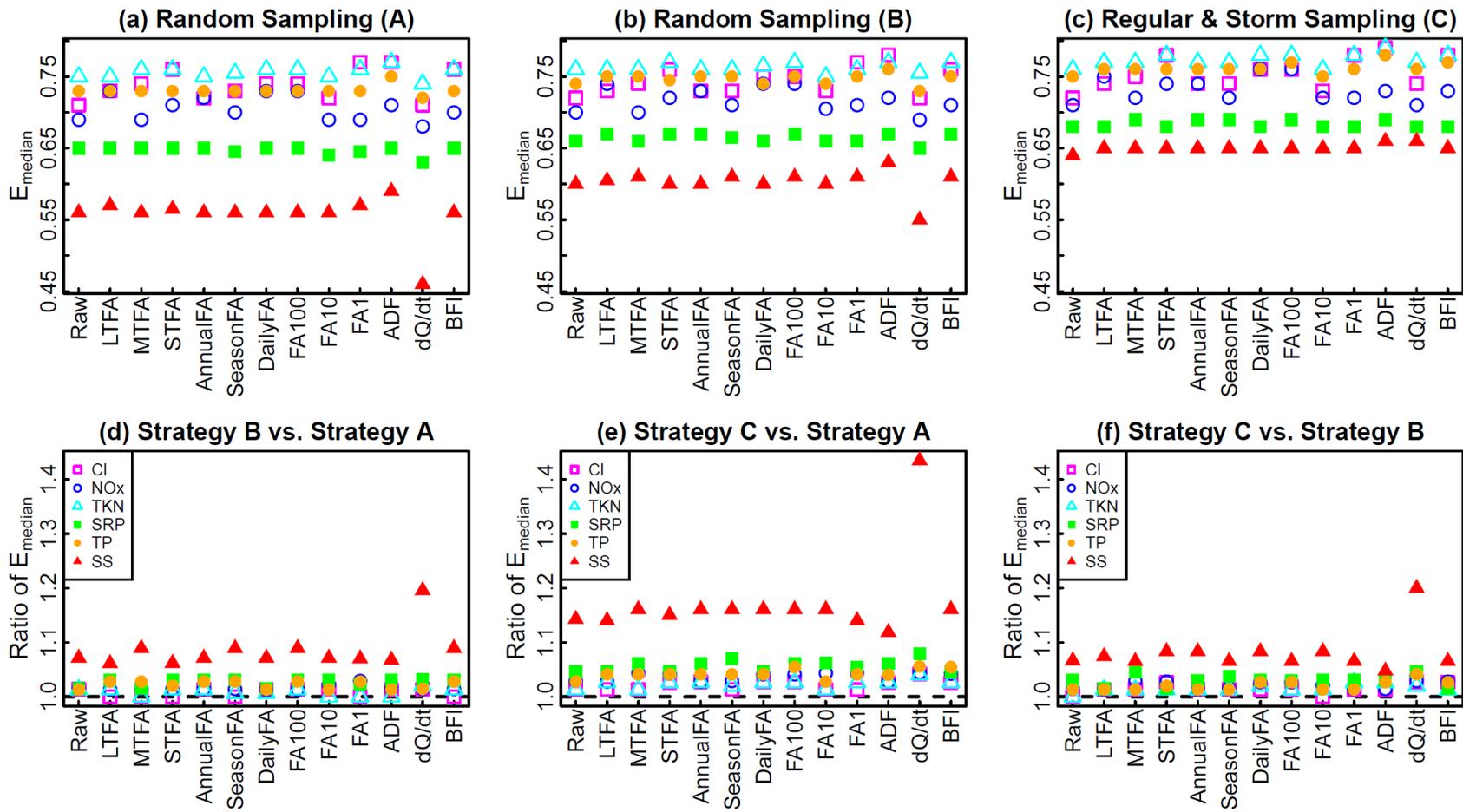


Figure G4. Performance of the original and modified models under the three sampling strategies. Plots (a)-(c) summarizes the performance with medians of modified Nash-Sutcliffe efficiency ( $E$ ) for *flux* based on 30 replicates at all 9 sites (*i.e.*, 270 cases) under each sampling strategy. Plots (d)-(f) compares the three sampling strategies directly using the ratios of  $E$  median.

*Page intentionally left blank*

**Appendix H. Supporting Information to Chapter 10**

**“Evaluation of Methods for Estimating Long-Range Dependence in Water Quality  
Time Series with Irregular Sampling”**

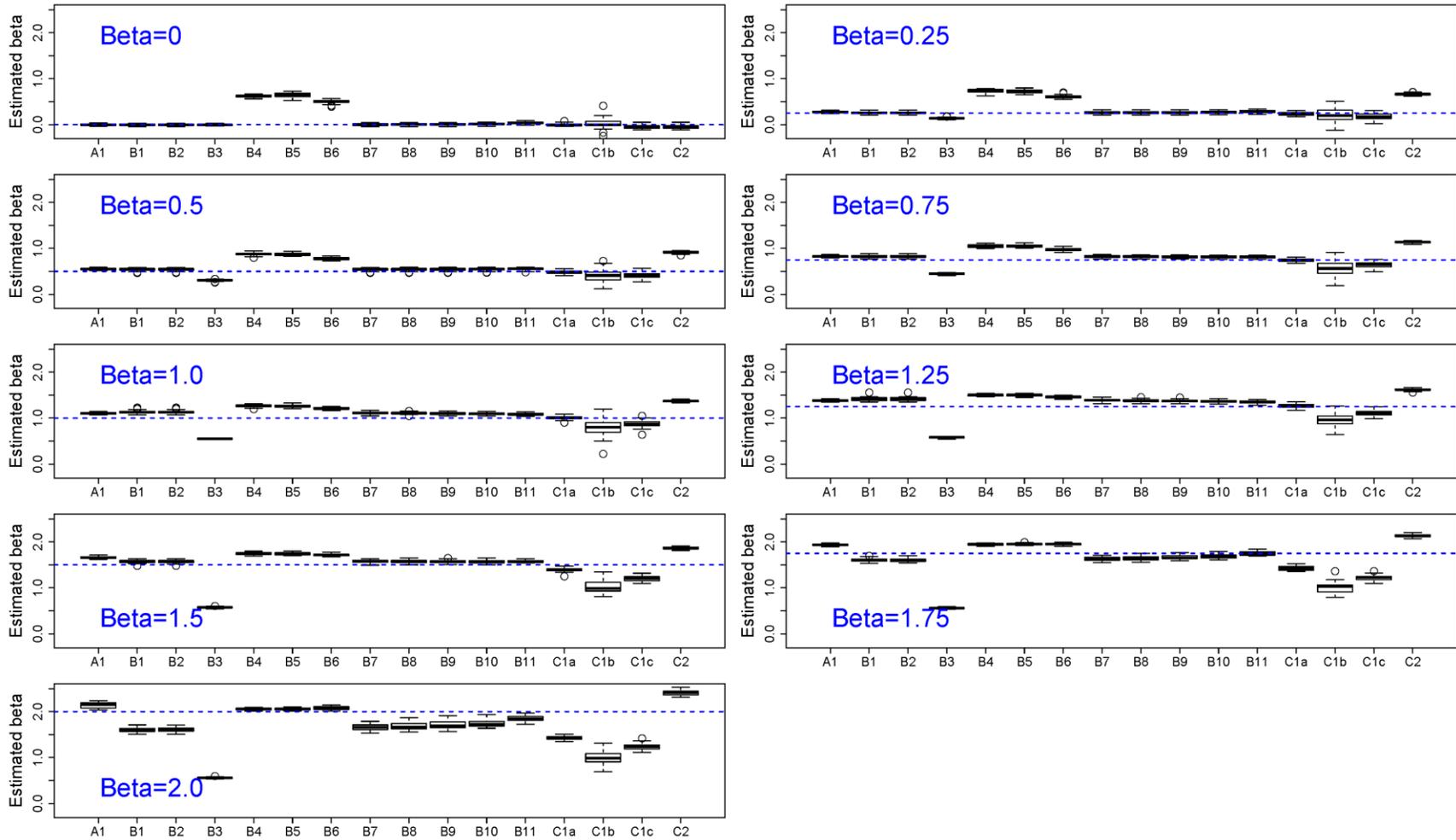


Figure H1. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 0.01, \mu = 1$ ) distributed gap intervals.

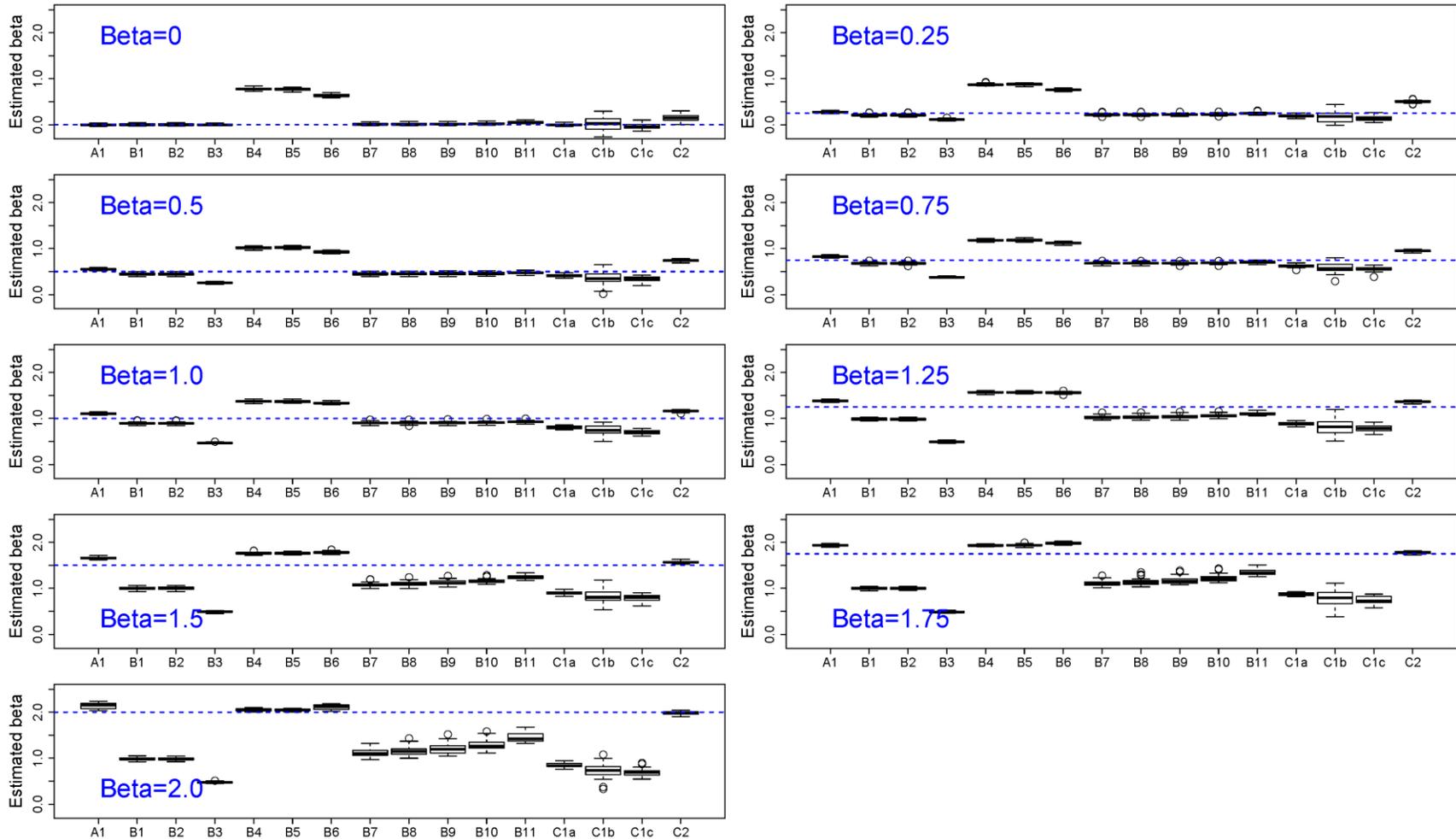


Figure H2. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 0.1, \mu = 1$ ) distributed gap intervals.

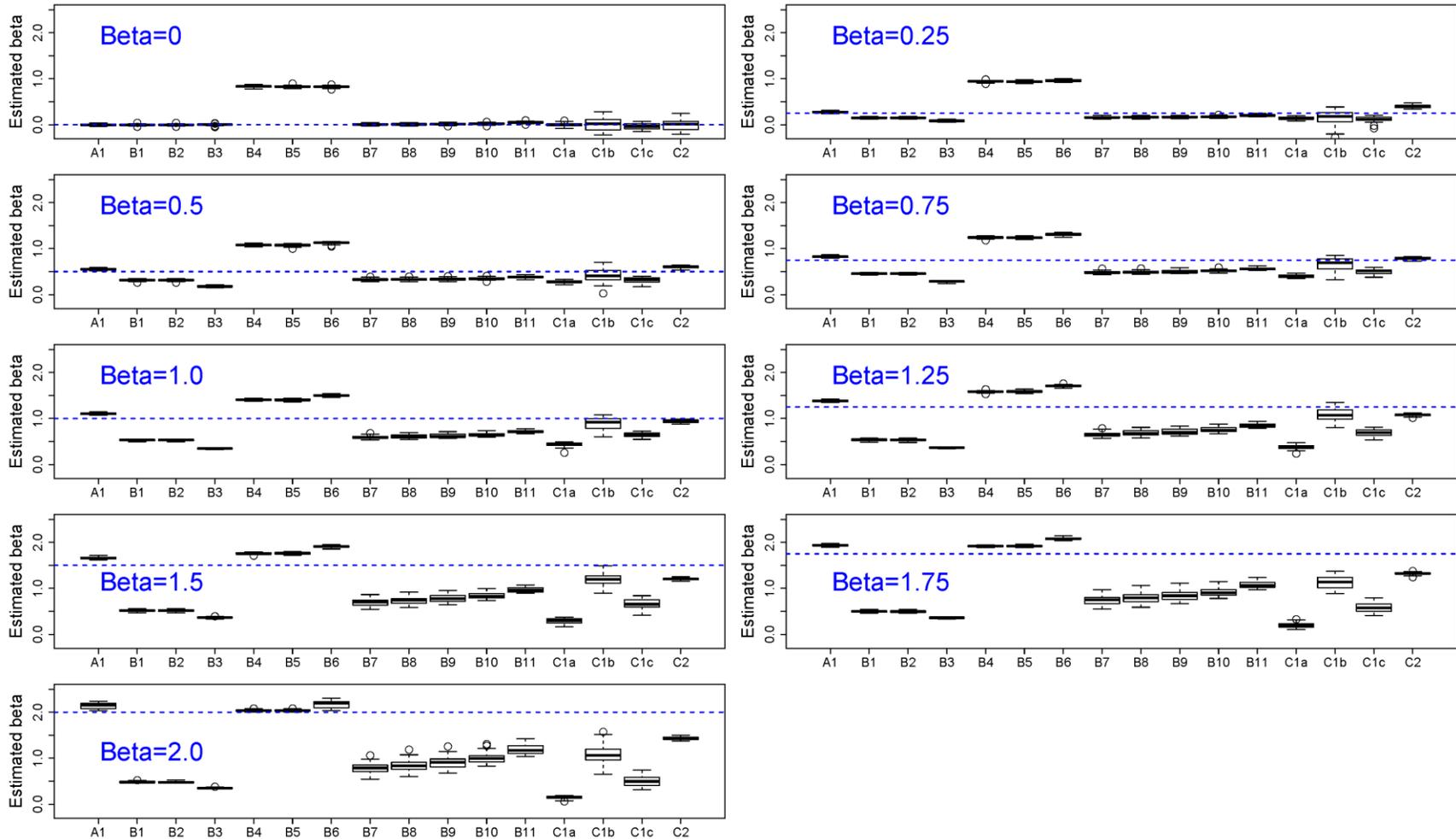


Figure H3. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 1, \mu = 1$ ) distributed gap intervals.

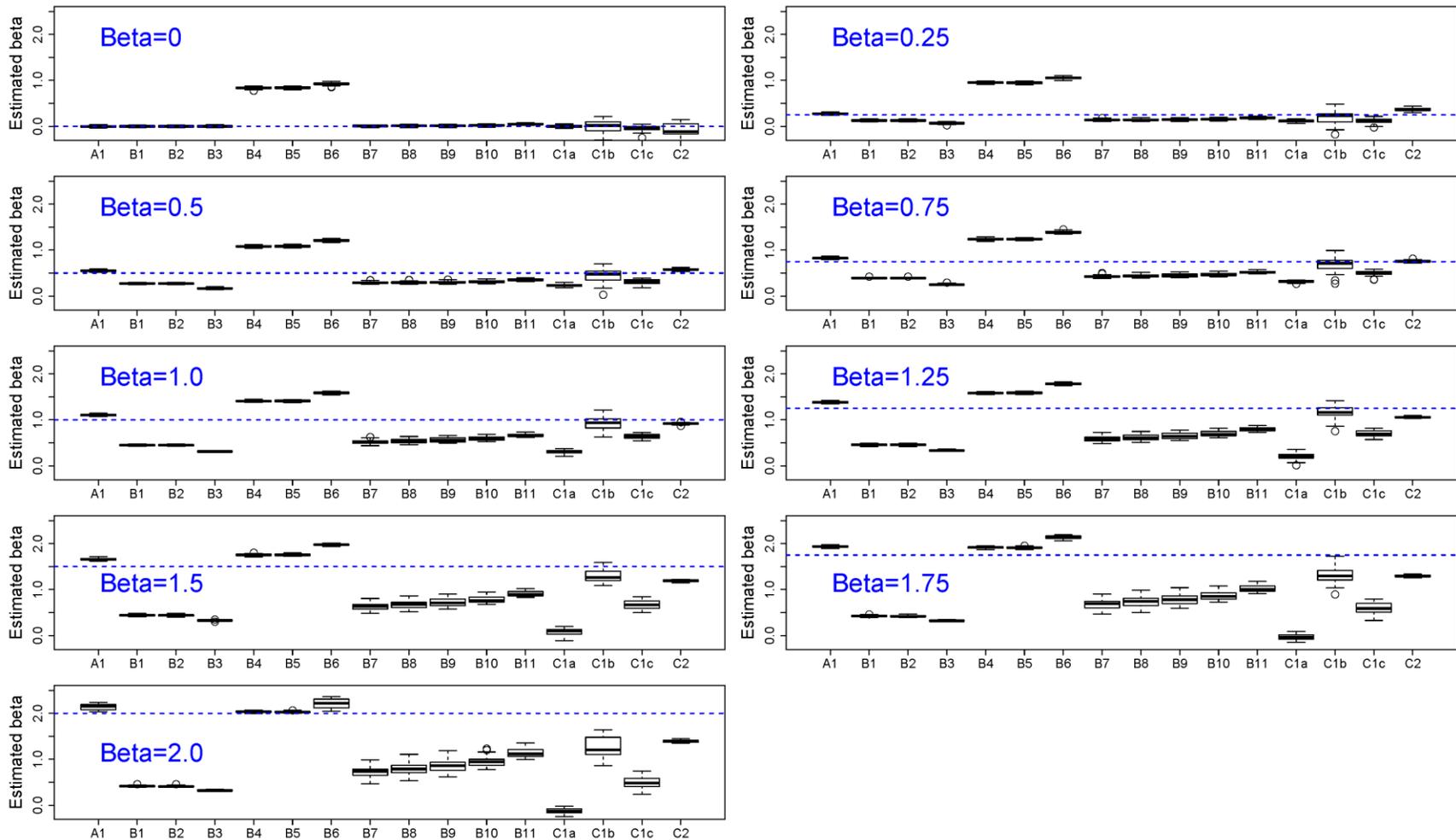


Figure H4. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 10, \mu = 1$ ) distributed gap intervals.

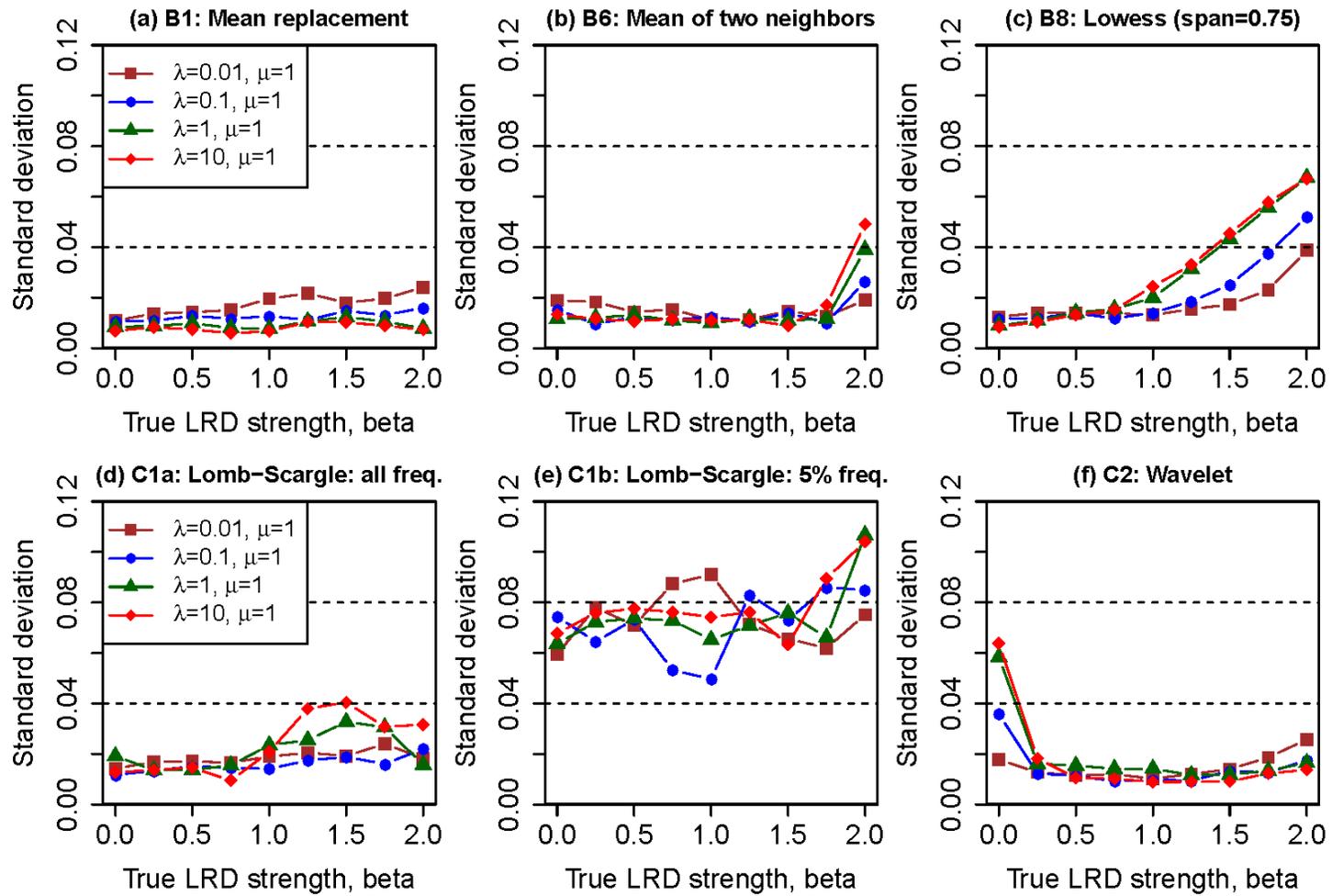


Figure H5. Comparison of standard deviation in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 2.

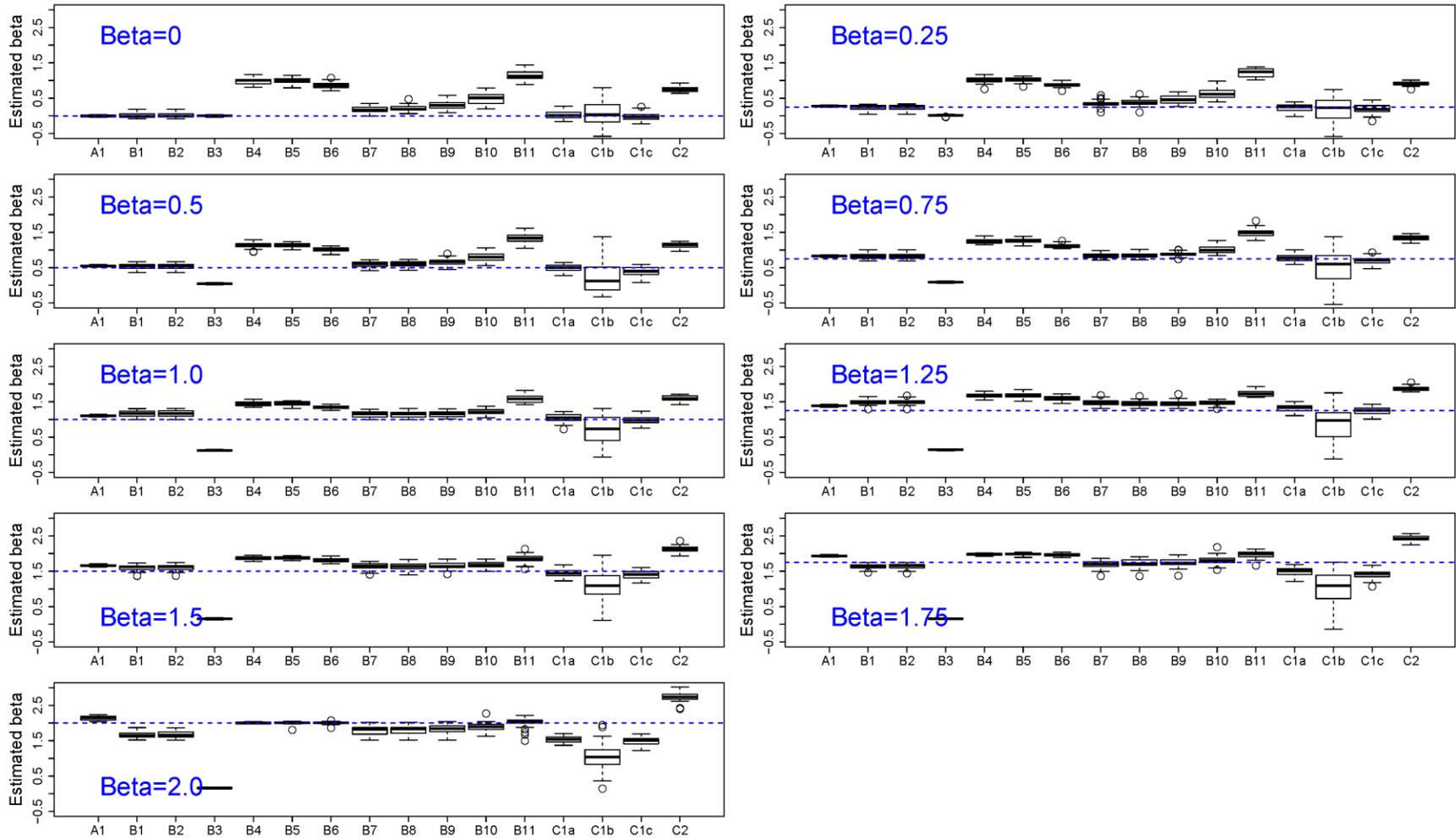


Figure H6. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 0.01, \mu = 14$ ) distributed gap intervals.

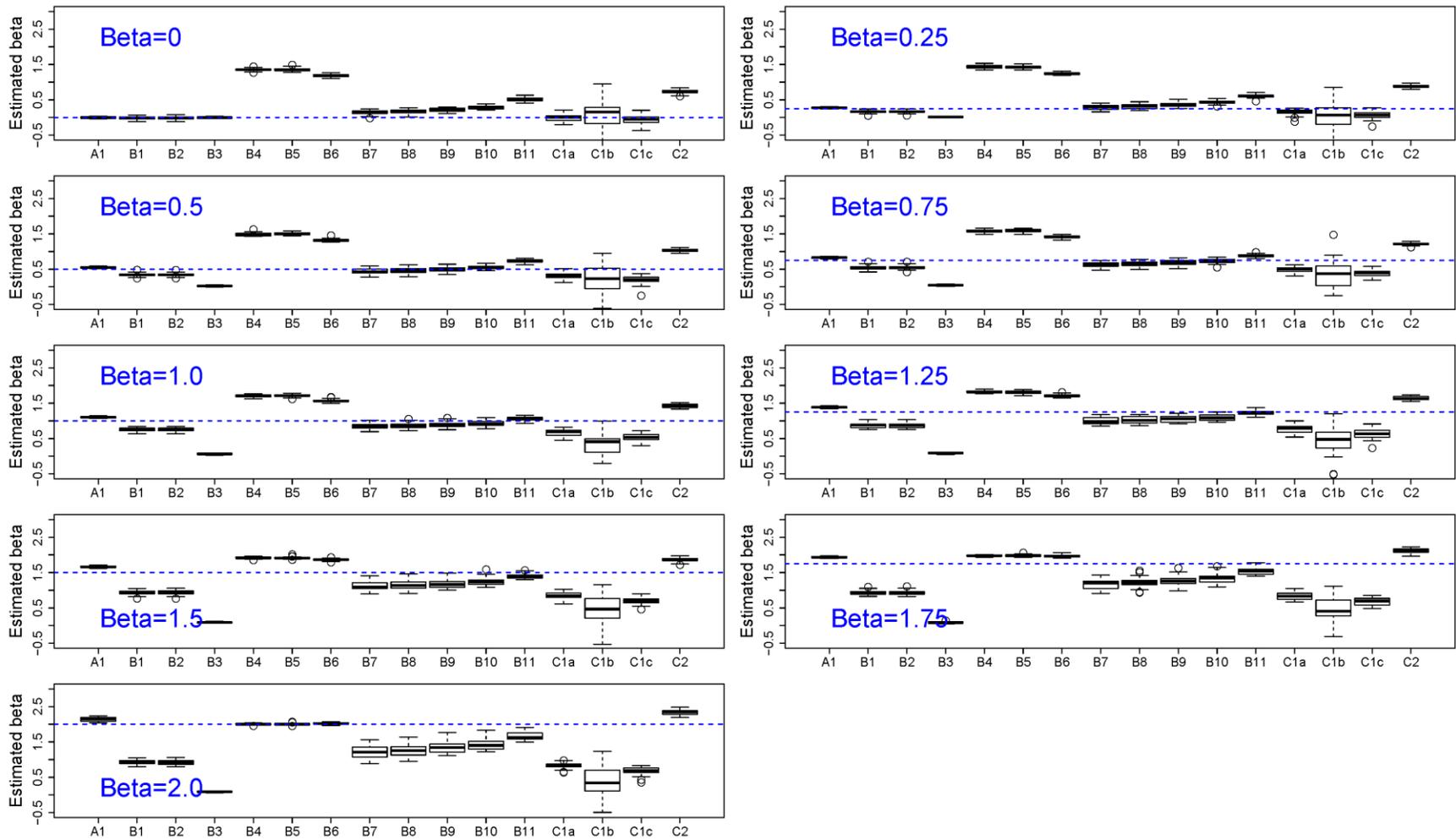


Figure H7. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 0.1, \mu = 14$ ) distributed gap intervals.

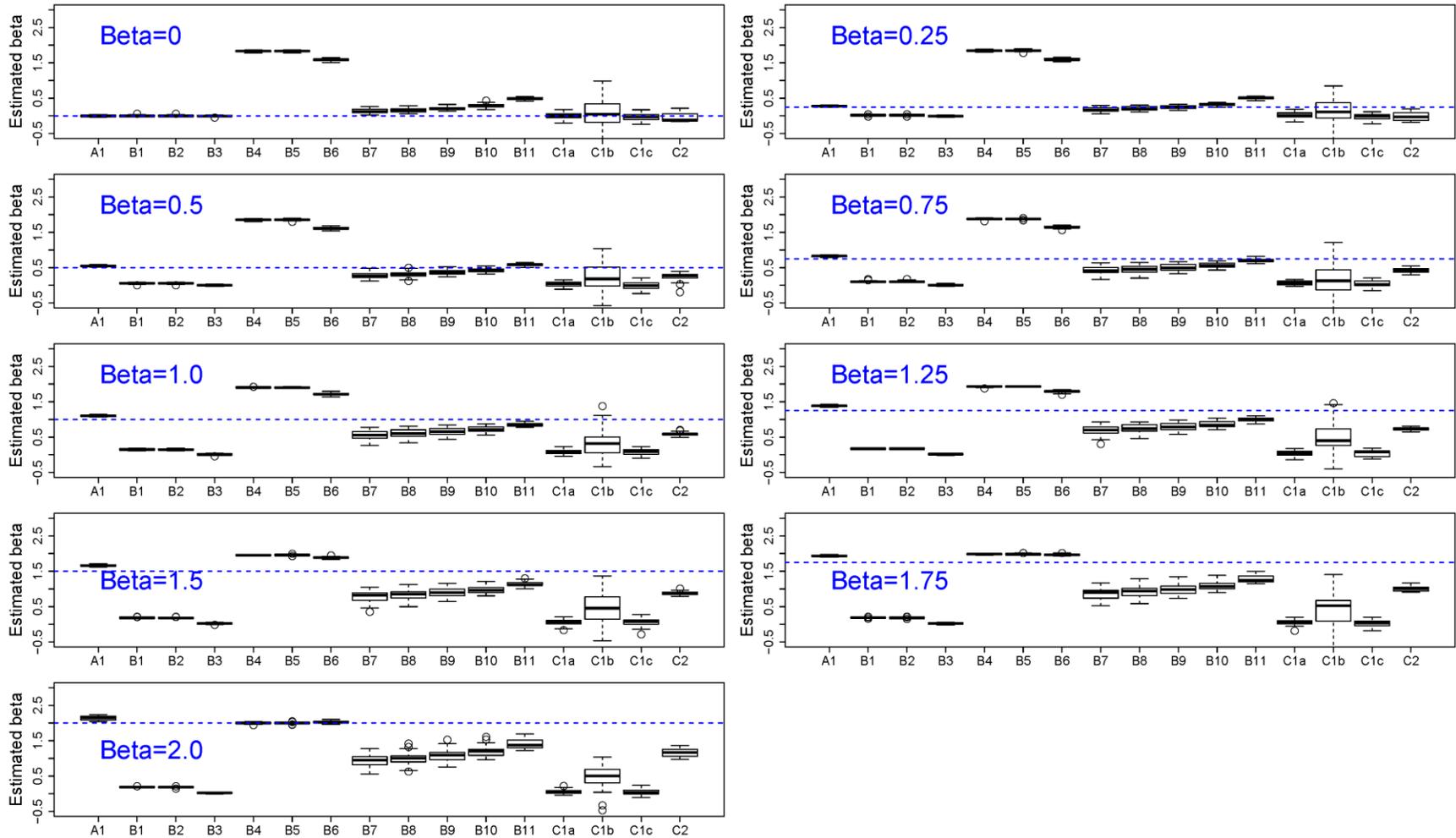


Figure H8. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 1, \mu = 14$ ) distributed gap intervals.

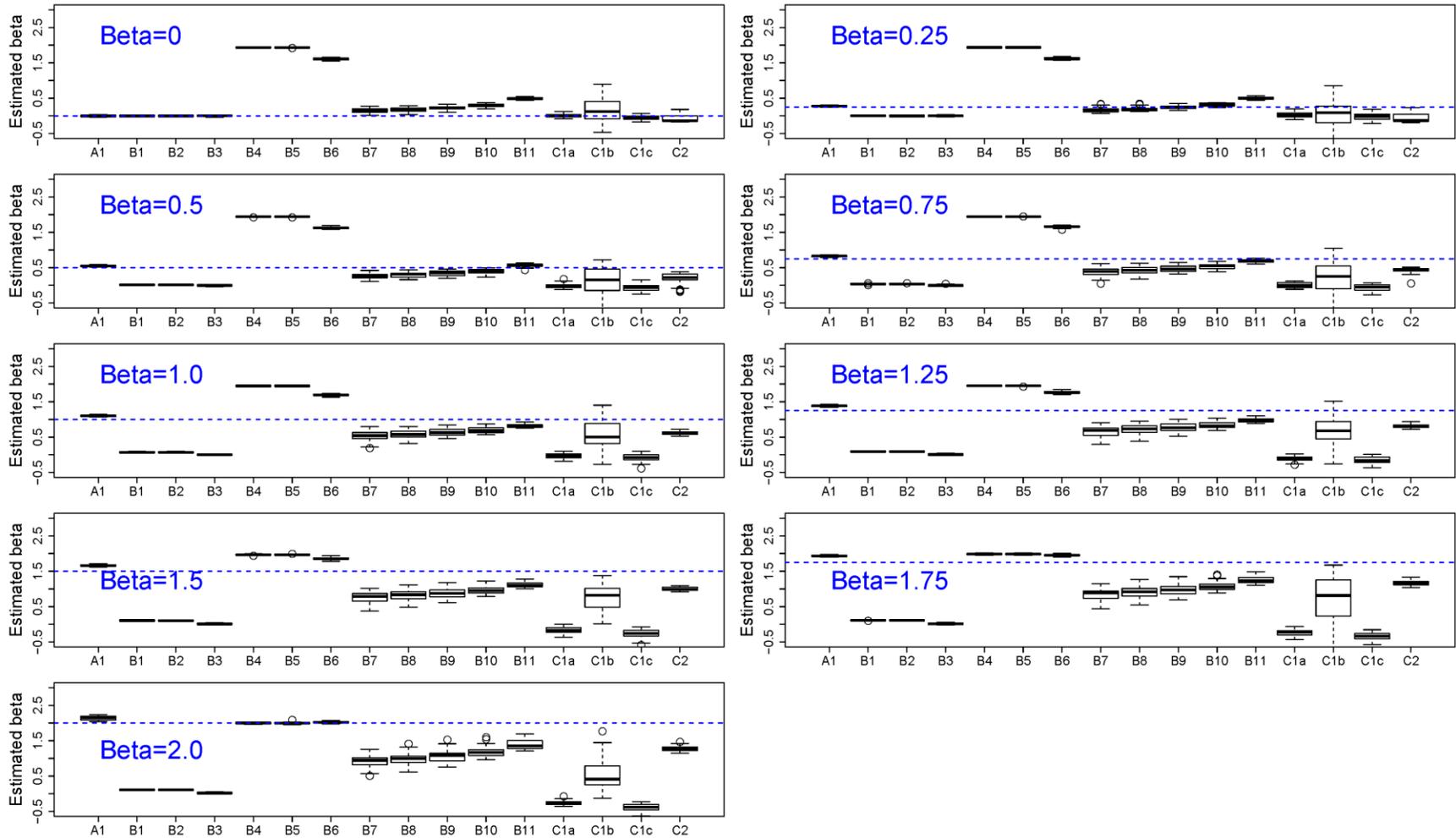


Figure H9. Comparison of methods for estimating long-range dependence (LRD) in irregular data (30 replicates) that are simulated with varying LRD, series length of 9125, and NB ( $\lambda = 10$ ,  $\mu = 14$ ) distributed gap intervals.

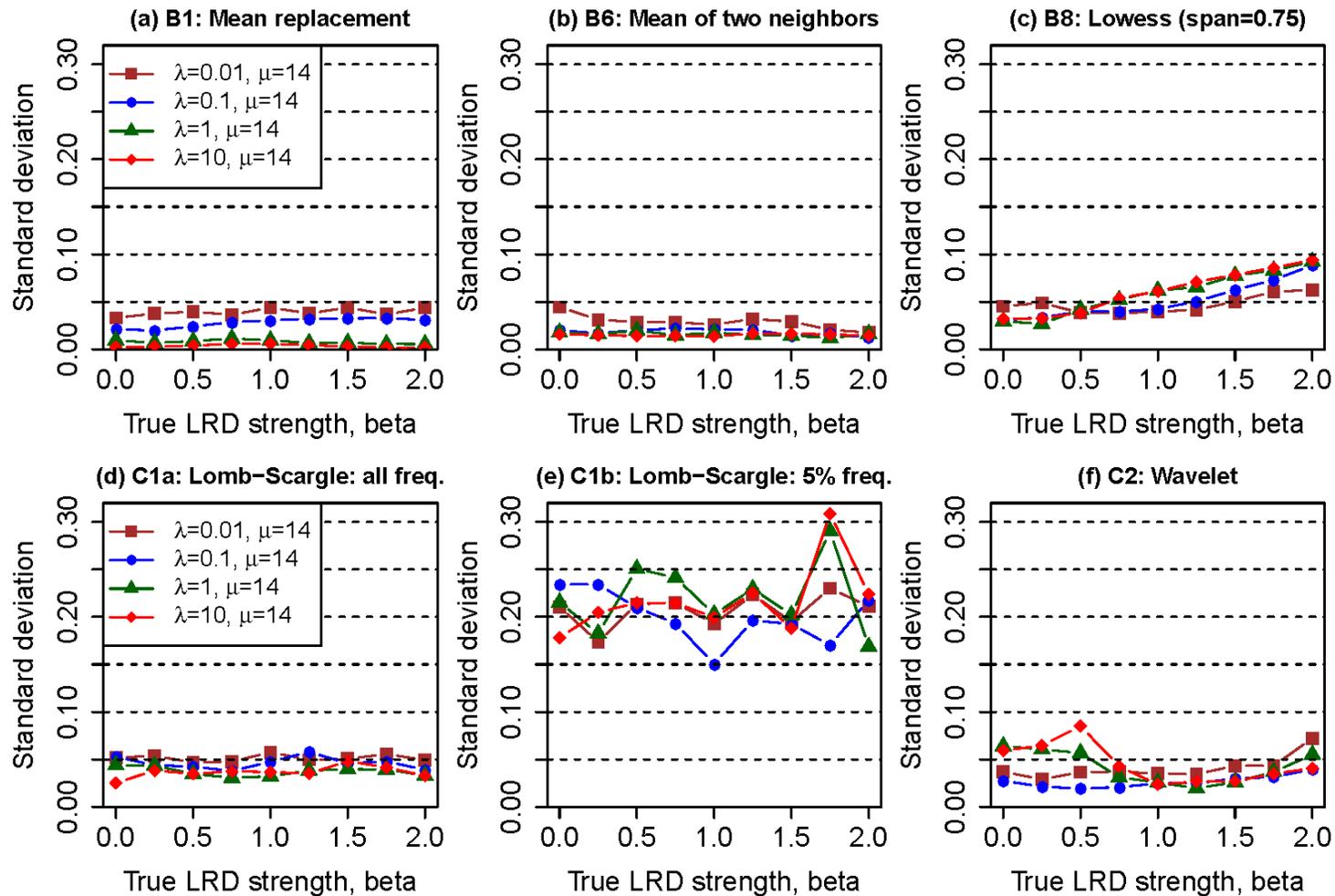


Figure H10. Comparison of standard deviation in estimated LRD strengths in irregular data that are simulated with varying LRDs (30 replicates), series length of 9125, and mean gap interval of 15.

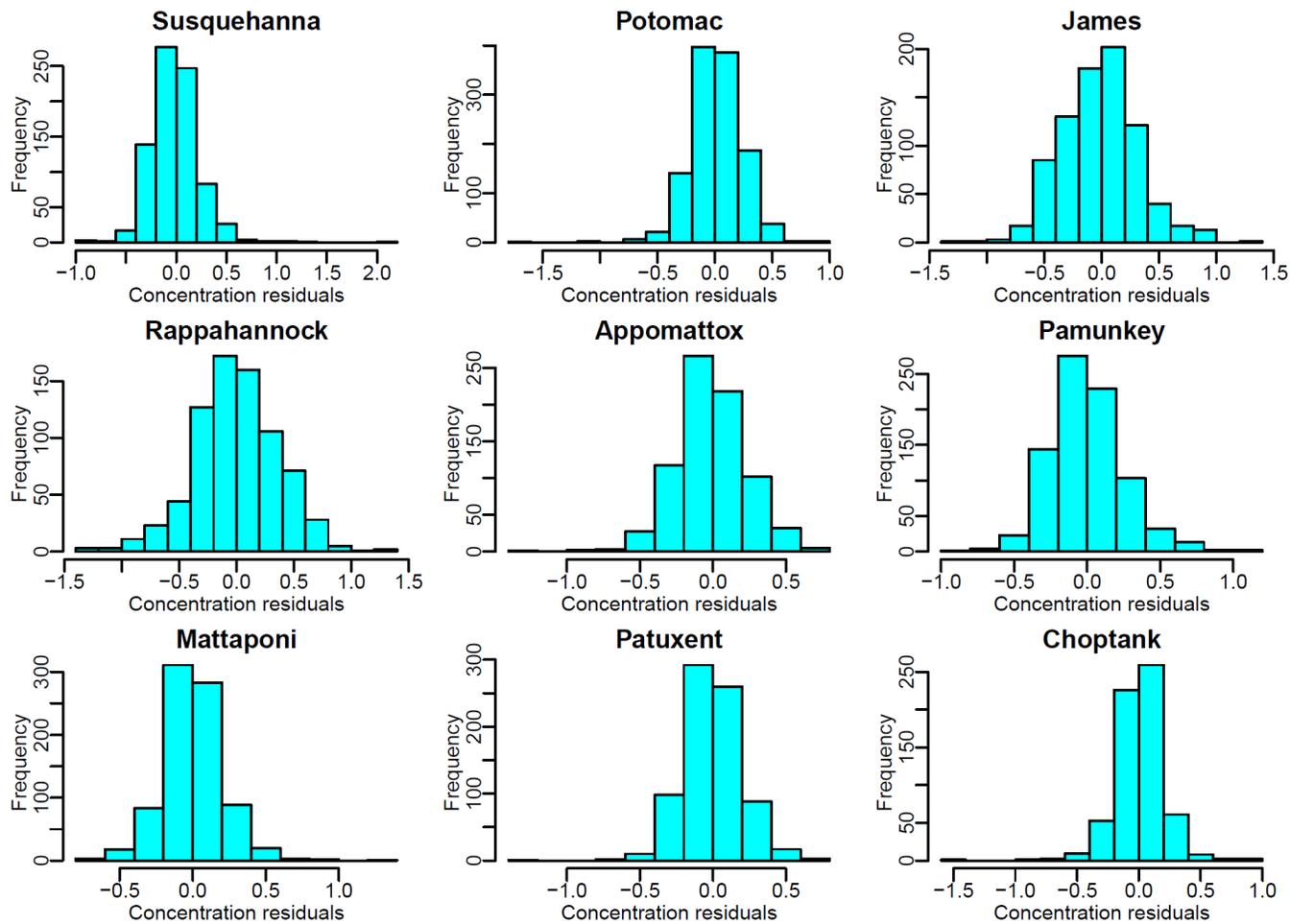


Figure H11. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for total nitrogen (TN) at the nine Chesapeake Bay monitoring sites. See Table 10.1 for site and data details.

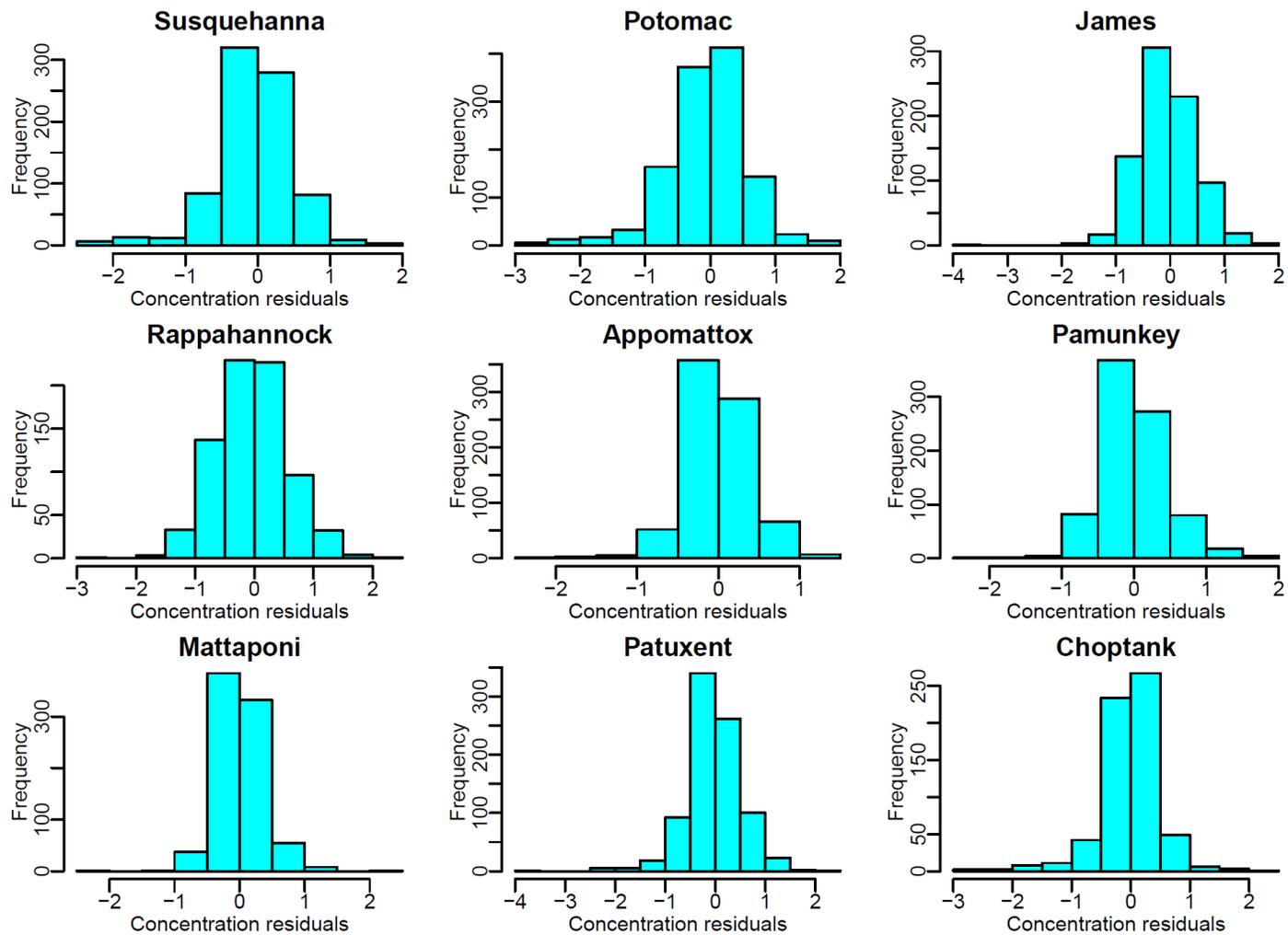


Figure H12. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for total phosphorus (TP) at the nine Chesapeake Bay monitoring sites. See Table 10.1 for site and data details.

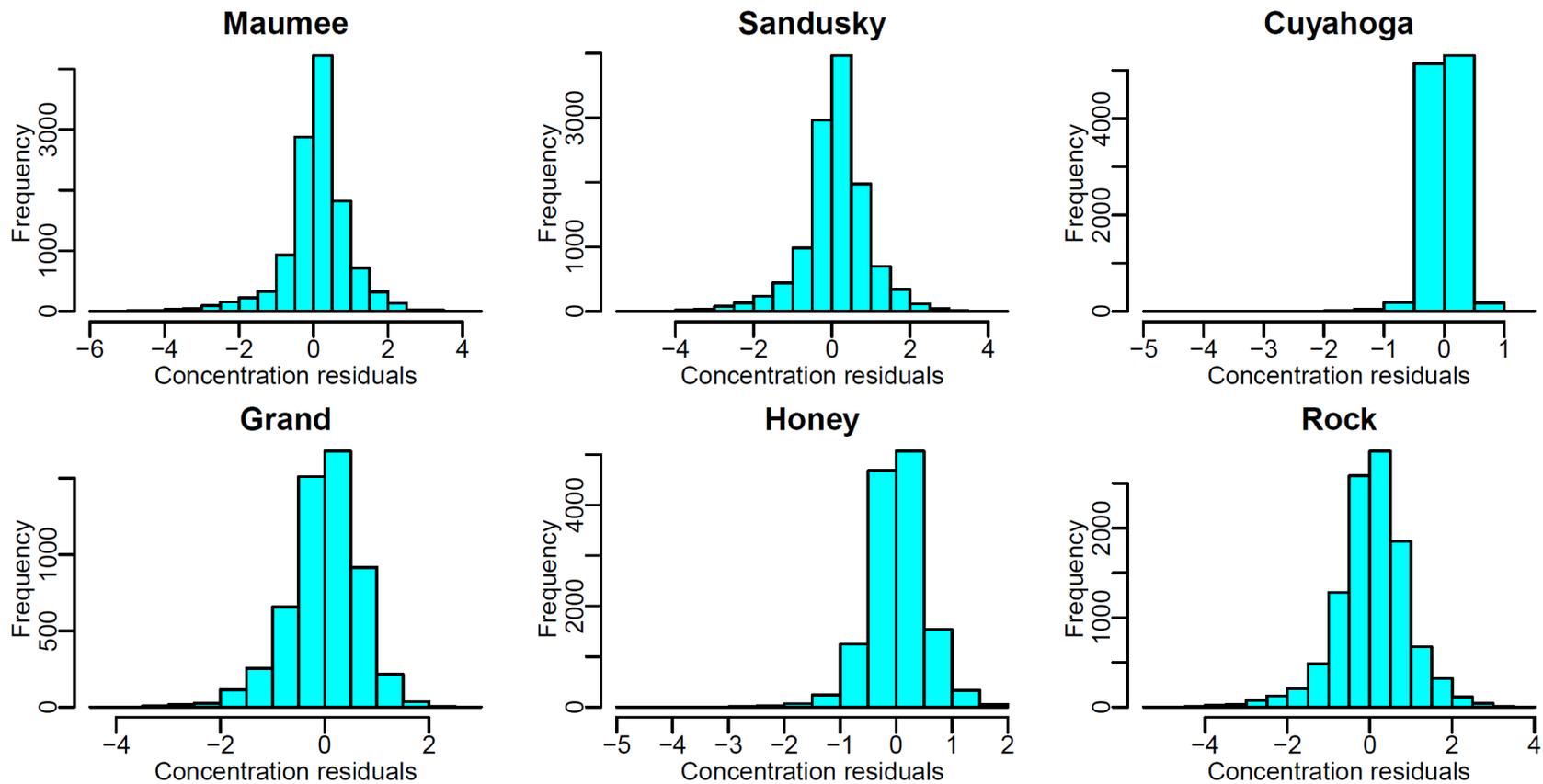


Figure H13. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for nitrate-plus-nitrite (NO<sub>x</sub>) at the six Lake Erie and Ohio River monitoring sites. See Table 10.1 for site and data details.

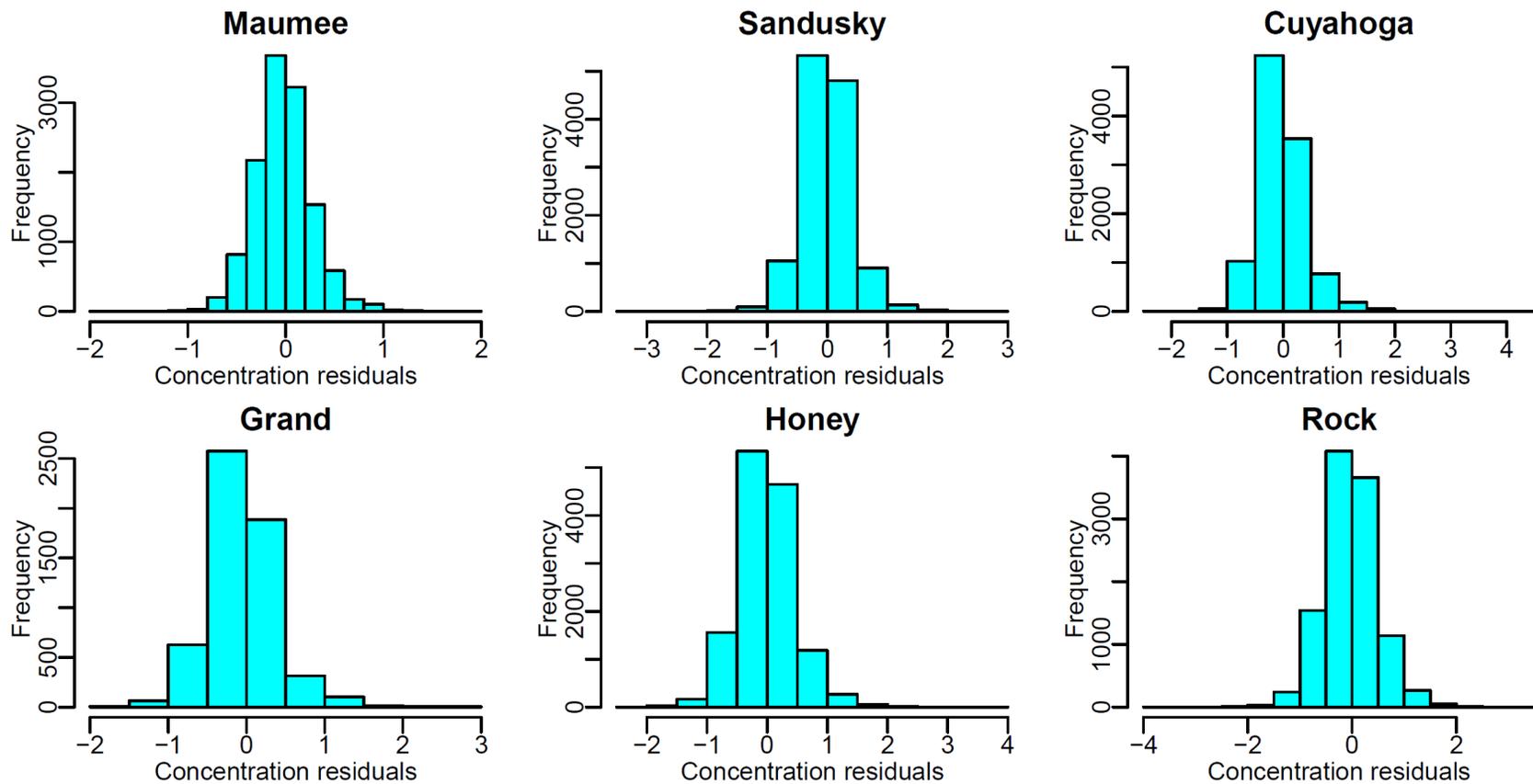


Figure H14. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for total phosphorus (TP) at the six Lake Erie and Ohio River monitoring sites. See Table 10.1 for site and data details.

*Page intentionally left blank*

## CURRICULUM VITAE

### ZHANG, QIAN

Department of Geography and Environmental Engineering  
Johns Hopkins University, Baltimore, MD  
Address: 3925 Beech Ave, Apt 115, Baltimore, MD 21211  
Email: [qzhang19@jhu.edu](mailto:qzhang19@jhu.edu); [zhangqian0324@gmail.com](mailto:zhangqian0324@gmail.com)

#### Research Interests

- Source, fate, and export of nitrogen, phosphorus, and sediment from watersheds
- Influences of changes in land use, climate, and input sources on nutrient export
- Evaluation of long-term river water-quality trends and the associated uncertainties
- Improvement of statistical methods for riverine flux estimation and trend analysis

#### Education

- Ph.D.** Johns Hopkins University, **Environmental Engineering**, 2011-2016.
- M.S.E.** Johns Hopkins University, **Applied Mathematics and Statistics**, 2013-2014.
- M.S.E.** Johns Hopkins University, **Environmental Engineering**, 2010-2011.
- B.Eng.** Nanyang Technological University (Singapore), **Environmental Engineering** (1<sup>st</sup> Class Honors), 2005-2009.

#### Appointments

- Research Assistant, Dept. of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD, 2010-present.
- Teaching Assistant, Dept. of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD, 2011-2012.
- Engineer, LBW Consultants LLP, Singapore, 2009-2010.

#### Skills & Expertise

- **Programming:** R, Matlab, Python
- **Computer:** Microsoft Office, ArcGIS, AutoCAD, HEC-RAS, Adobe Illustrator
- **Water Resources:** Hydrology, Watershed, Sediment, Nutrients, TMDL, BMPs, NPS, Reservoir
- **Engineering:** Physical/Chemical Processes, Biological Processes, Contaminant Transport
- **Analytics:** Statistics, Monte Carlo, Data Mining, Machine Learning, Time Series, Modeling

#### Honors & Awards

- **Recipient**, Johns Hopkins Graduate Representatives Organization (GRO) Travel Grant, 2016
- **Research Fellow**, Maryland Water Resources Research Center (MWRRC), 2015
- **Recipient**, Geological Society of America (GSA) Student Travel Grant, 2015
- **Research Fellow**, Maryland Sea Grant Graduate (MDSG), 2013 - 2014
- **Graduate Scholar**, Community Surface Dynamics Modeling System (CSDMS), 2013

- **Winner**, Chesapeake Water Environment Assoc. (CWEA) Student Paper Competition, 2012
- **Awardee**, Dean's List Award, Nanyang Technological University (NTU), Singapore, 2009
- **Awardee**, Dean's List Award, Nanyang Technological University (NTU), Singapore, 2008
- **Winner**, Singapore National Concrete Canoe Competition (SNCCC), 2008

### **Journal Publications**

- Zhang, Q.**, W.P. Ball., and D.L. Moyer. 2016. "Decadal-scale Export of Nitrogen, Phosphorus, and Sediment from the Susquehanna River Basin, USA: Analysis and Synthesis of Temporal and Spatial Patterns", *Science of the Total Environment*, 563-564: 1016-1029, [doi: 10.1016/j.scitotenv.2016.03.104](https://doi.org/10.1016/j.scitotenv.2016.03.104).
- Zhang, Q.**, R.M. Hirsch, and W.P. Ball. 2016. "Long-Term Changes in Sediment and Nutrient Delivery from Conowingo Dam to Chesapeake Bay: Effects of Reservoir Filling", *Environmental Science & Technology*, 50(4): 1877-1886, [doi: 10.1021/acs.est.5b04073](https://doi.org/10.1021/acs.est.5b04073).
- Zhang, Q.**, D.C. Brady, W.R. Boynton, and W.P. Ball. 2015. "Long-term Trends of Nutrients and Sediment from the Non-tidal Chesapeake Watershed: An Assessment of Progress by River and Season", *Journal of the American Water Resources Association*, 51(6): 1534-1555, [doi: 10.1111/1752-1688.12327](https://doi.org/10.1111/1752-1688.12327).
- Zhang, Q.**, D.C. Brady, and W.P. Ball, 2013. "Long-term Seasonal Trends of Nitrogen, Phosphorus, and Suspended Sediment Load from the Non-tidal Susquehanna River Basin to Chesapeake Bay", *Science of the Total Environment*, 452-453: 208-221, [doi: 10.1016/j.scitotenv.2013.02.012](https://doi.org/10.1016/j.scitotenv.2013.02.012).

### **Journal Manuscripts under Preparation or Review**

- Zhang, Q.**, C.J. Harman, and W.P. Ball. "An Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-quality Data", manuscript under review.
- Zhang, Q.**, C.J. Harman, and J.W. Kirchner. "Evaluation of Methods for Estimating Long-Range Dependence in Irregular Water Quality Time Series", manuscript under review.
- Zhang, Q.** and W.P. Ball. "Improving Riverine Constituent Concentration and Flux Estimation by Accounting for Antecedent Discharge Conditions", manuscript in preparation.
- Zhang, Q.** and W.P. Ball. "Non-stationary Concentration-Discharge Relationships: A Synthesis of Nutrient and Sediment Patterns in the Major Tributaries to Chesapeake Bay", manuscript in preparation.
- Zhang, Q.** and W.P. Ball. "What Can We Learn from Limited Data? Statistical Inferences and Uncertainties of Riverine Fluxes and Trends with Limited Sampling of Extreme-Flow Events", manuscript in preparation.

### **Data Publications**

- Zhang, Q.** and W.P. Ball. 2016. "Data associated with Improved Method for Interpretation of Concentration-Discharge Relationships in Riverine Water-Quality Data", *Johns Hopkins University Data Archive*, [dx.doi.org/10.7281/T18G8HM0](https://dx.doi.org/10.7281/T18G8HM0).
- Zhang, Q.** and W.P. Ball. 2016. "Data associated with Decadal-scale export of nitrogen, phosphorus, and sediment from the Susquehanna River basin, USA: Analysis and synthesis of temporal and spatial patterns", *Johns Hopkins University Data Archive*, [dx.doi.org/10.7281/T1QN64NW](https://dx.doi.org/10.7281/T1QN64NW).
- Zhang, Q.** and W.P. Ball. 2014. "Data associated with Long-term seasonal trends of nutrients and sediment from the nontidal Chesapeake Bay Watershed", *Johns Hopkins University Data Archive*, [dx.doi.org/10.7281/T1VD6WC7](https://dx.doi.org/10.7281/T1VD6WC7).

**Zhang, Q.** and W.P. Ball. 2014. “Data associated with Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay”, *Johns Hopkins University Data Archive*, [dx.doi.org/10.7281/T1KW5CX5](https://dx.doi.org/10.7281/T1KW5CX5).

### **Conference Presentations**

**Zhang, Q.**, R.M. Hirsch, and W.P. Ball, 2016. “Temporal Changes in Net Deposition of Sediment and Nutrients behind Conowingo Dam under Different Flow Conditions: Statistical Evaluations of Monitoring Data between 1987 and 2013”, oral presentation at *Chesapeake Modeling Symposium*, Williamsburg, VA, June 1-2, 2016.

**Zhang, Q.**, D.C. Brady, W.R. Boynton, and W.P. Ball, 2016. “Nutrient and Sediment Trends from the Nontidal Chesapeake Bay Watershed: Synthesis of Progress by Season for the Nine Major Tributaries”, oral presentation at *Chesapeake Modeling Symposium*, Williamsburg, VA, June 1-2, 2016.

**Zhang, Q.**, R.M. Hirsch, and W.P. Ball, 2016. “Effects of Reservoir Filling on Sediment and Nutrient Delivery from Susquehanna River to Chesapeake Bay: Input-Output Analyses based on Long-Term Monitoring”, **invited** oral presentation at *U.S. Geological Survey MD-DE-DC Water Science Center Seminars*, Baltimore, MD, April 12, 2016.

**Zhang, Q.**, R.M. Hirsch, and W.P. Ball, 2016. “Effects of Reservoir Filling on Sediment and Nutrient Removal in the Lower Susquehanna River Reservoir System: An Input-Output Analysis based on Long-Term Monitoring”, **invited** oral presentation at *Chesapeake Bay Program's Scientific and Technical Advisory Committee (STAC) Workshop: Conowingo Infill Influence on Chesapeake Water Quality*, Annapolis, MD, January 13-14, 2016.

**Zhang, Q.** and W.P. Ball. 2015. “Non-stationary Concentration-Discharge Relationships for Nitrogen, Phosphorus, and Sediment for Nine Major Tributaries of the Chesapeake Bay”, oral presentation at *American Geophysical Union Fall Meeting*, San Francisco, CA, December 14-18, 2015.

W.P. Ball, **Q. Zhang**, and R.M. Hirsch. 2015. “Effects of Reservoir Filling on Sediment and Nutrient Removal in the Lower Susquehanna River Reservoir: An Input-Output Analysis Based on Long-Term Monitoring”, oral presentation at *American Geophysical Union Fall Meeting*, San Francisco, CA, December 14-18, 2015.

**Zhang, Q.** and W.P. Ball. 2015. Concentration-Discharge Relationships for Nutrients and Sediment in Major Tributaries to Chesapeake Bay: Typical Patterns and Non-Stationarity, poster presentation at *Maryland Water Monitoring Council Annual Conference*, Linthicum Heights, MD, November 13, 2015.

**Zhang, Q.**, W.P. Ball., and D.L. Moyer. 2015. “Long-Term Export of Nitrogen, Phosphorus, and Sediment in the Susquehanna River Basin: Analysis of Decadal-Scale Trends and Sub-Basin Mass Balances”, oral presentation at *Geological Society of America Annual Meeting*, Baltimore, MD, November 1-4, 2015.

Ball, W.P., **Q. Zhang**, D.C. Brady, and W.R. Boynton. 2015. “Long-term Trends of Nutrients and Sediment from the Non-tidal Chesapeake Watershed: An Assessment of Progress by River and Season”, oral presentation at *Association of Environmental Engineering and Science Professors Annual Conference*, New Haven, CT, June 13-16, 2015.

**Zhang, Q.**, C.J. Harman, and W.P. Ball. 2014. “Evaluation of Methods for Estimating Long-Range Dependence in Water Quality Time Series with Missing Data and Irregular Sampling”, poster presentation at *American Geophysical Union Fall Meeting*, San Francisco, CA, December 15-19, 2014.

Ball, W.P., **Q. Zhang**, D.C. Brady, and W.R. Boynton. 2014. “Long-Term Loads of Nutrients and Sediment from Non-Tidal Regions of the Chesapeake Bay Watershed: An Assessment of Seasonal Trends and Progress”, poster presentation at *American Geophysical Union Fall Meeting*, San Francisco, CA, December 15-19, 2014.

- Wei, H., D. Ha, **Q. Zhang**, and W.P. Ball. 2014. “Effectiveness of Nitrogen Assimilation in the Non-Tidal Chesapeake Bay Watershed: Evaluations Based on Thirty Years of Data”, poster presentation at *American Geophysical Union Fall Meeting*, San Francisco, CA, December 15-19, 2014.
- Ha, D., H. Wei, **Q. Zhang**, and W.P. Ball. 2014. Retrospective Analysis of Sediment-associated Phosphorus Concentration in the Non-Tidal Chesapeake Bay Watershed, poster presentation at *Maryland Water Monitoring Council Annual Conference*, Linthicum Heights, MD, November 21, 2014.
- Ha, D., H. Wei, **Q. Zhang**, and W.P. Ball. 2014. Nitrogen Source Input from the Non-Tidal Chesapeake Bay Watershed and Output in the Major Rivers: Evaluation of Changes Based on Long-term Data, poster presentation at *Maryland Water Monitoring Council Annual Conference*, Linthicum Heights, MD, November 21, 2014.
- Wei, H., D. Ha, **Q. Zhang**, and W.P. Ball. 2014. “Retrospective Analysis of Phosphorus Source Input and Riverine Output in the Chesapeake Bay Watershed”, poster presentation at *Maryland Water Monitoring Council Annual Conference*, Linthicum Heights, MD, November 21, 2014.
- Zhang, Q.** and W.P. Ball. 2014. “Decadal-scale Trends of Nutrients and Sediment from the Non-tidal Chesapeake Bay Watershed: Are We Making of Progress in Loading Reduction”, poster presentation at *Gordon Research Conference - Environmental Sciences: Water*, Holderness, NH, June 22-27, 2014.
- Zhang, Q.** and W.P. Ball. 2014. “Nutrient and Sediment Delivery from the Susquehanna River to Chesapeake Bay: Long-term Changes in Loading Trend and Reservoir Sedimentation”, oral presentation at *Chesapeake Modeling Symposium*, Annapolis, MD, May 28-29, 2014.
- Zhang, Q.** and W.P. Ball. 2014. “Long-term trends and mass-balance of nutrient and sediment loadings in the Lower Susquehanna River Watershed”, **invited** oral presentation at *Chesapeake Bay Program Modeling Quarterly Review Meeting*, Annapolis, MD, January 7, 2014.
- Zhang, Q.** and W.P. Ball. 2013. “Long-term Seasonal Trends of Nutrients and Sediment from the Non-tidal Chesapeake Bay Watershed: An Assessment of Progress in Loading Reduction”, oral presentation at *American Water Resources Association Annual Conference*, Portland, OR, November 4-7, 2013.
- Zhang, Q.** and W.P. Ball. 2013. “Nutrient and Sediment Loads Delivered through the Conowingo Reservoir”, **invited** oral presentation at *Maryland Department of Natural Resources Conowingo Dam Sediment Load Meeting*, Annapolis, MD, August 19, 2013.
- Zhang, Q.**, D.C. Brady, and W.P. Ball. 2013. “Application of a USGS Statistical Tool (WRTDS) toward Assessing Watershed Management and Reservoir Function in the Susquehanna River Basin”, poster presentation at *Association of Environmental Engineering and Science Professors 50th Anniversary Conference*, Golden, CO, July 14-16, 2013.
- Zhang, Q.**, D.C. Brady, and W.P. Ball. 2013. “Long-term Seasonal Trends of Nitrogen, Phosphorus, and Suspended Sediment Load from the Non-tidal Susquehanna River Basin to Chesapeake Bay”, poster presentation at *Community Surface Dynamics Modeling System Annual Meeting*, Boulder, CO, March 23-25, 2013.
- Zhang, Q.** and W.P. Ball. 2012. “Long-term Seasonal Nutrient Trends for the Non-tidal Portions of the Major Tributaries to Chesapeake Bay”, poster presentation at *Chesapeake Modeling Symposium*, Annapolis, MD, May 21-22, 2012.

### **Synergistic Activities**

- Member, Geological Society of America (GSA), 2015-present.
- Member, American Geophysical Union (AGU), 2014-present.
- Member, American Water Resources Association (AWRA), 2013-present.
- Member, Assoc. of Environ. Engineering and Science Professors (AEESP), 2013-present.
- Member, Community Surface Dynamics Modeling System (CSDMS), 2012-present.
- Member, American Water Works Association (AWWA), 2012-present.