Hydrologic flowpath and other natural and anthropogenic factors controlling nitrogen movement from the landscape to streams

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Scott Craig Kronholm

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Paul D. Capel, Adviser

August 2015

Acknowledgements

Abstract

The growth of food, fuel, feed, and fiber crops on agricultural land requires additions of nitrogen fertilizer to bolster crop yields. The portion of applied nitrogen that is not utilized by the crops is highly susceptible to transport to nearby streams. Elevated nitrogen concentrations in streams can lead to water quality concerns such as ecosystem degradation or drinking water contamination. Identification of geospatial, environmental, and watershed characteristics (variables) that are correlated with nitrogen concentration in streams will provide a greater understanding of the influence that certain variables have on nitrogen transport to streams. The route through which water travels from the landscape to streams (flowpath) is one of these variables. Movement of many forms of nitrogen is linked to the movement of water, therefore, understanding the fluxes of water to streams will help to expand the understanding of nitrogen transport and the affect that landscape management changes will have on the concentrations of nitrogen in streams.

In some areas, groundwater is an important flowpath for delivering water and nitrogen to streams. Hydrograph separation can be used to estimate the amount of total streamflow that is attributable to slowflow sources (flowpaths through which water moves slowly) such as groundwater, and fastflow sources (flowpaths through which water moves quickly) such as overland flow. Because flowpaths have an impact on water and nitrogen transport to streams, testing and improving hydrograph separation techniques is needed. Two independent methods of hydrograph separation, the graphical-based BFI program and chemical tracer-based end-member mixing analysis (EMMA), were used to estimate slowflow contributions to the same streams. The estimates of slowflow from the two separate methods of hydrograph separation were not identical, highlighting the differences in how each method works and the difficulty of accurately estimating slowflow.

A modified method of EMMA, referred to as a ratio-based EMMA, was created and tested using synthetic and real stream data. The ratio-based EMMA represents a new method of hydrograph separation, as it produced reasonably accurate slowflow estimates when tested against synthetic and real stream data.

The importance of flowpath was then tested in six highly modified streams in agricultural watersheds that had extensive data sets and well understood hydrology. The importance of flowpath on stream nitrogen concentrations in these streams aligned with expectations based on what is currently known.

Finally, this study was expanded to a large number of small streams where statistical methods were used to gain broader understanding of the controls on water and nitrogen movement to streams, and to allow for extrapolation of this information to unstudied streams.

In addition to flowpath, a small number of geospatial, environmental, and watershed variables were shown to be important for estimating total nitrogen loads and concentrations in a large number of small streams. These variables were used in the development of several multiple linear regression models, many of which performed well when applied to a set of validation streams, having reasonable high R^2 values and low normalized root mean squared errors.

Determination of the important flowpath(s) as well as other important variables which increase nitrogen movement to streams will allow watershed managers to more accurately implement beneficial land management practices in an effort to reduce nitrogen movement to streams. This knowledge will become increasingly important in an effort to maintain or reduce the amount of nitrogen in streams while increasing crop production to support the rising global population.

Table of contents

List of Tables

List of Figures

Chapter 1: Introduction

1. <u>Introduction</u>

Application of reactive forms of nitrogen to cropland (either through application of manure or industrially created forms) allowed for the dramatic increases in food production seen throughout a majority of the last century (USDA-NASS, 2015), and helps sustain the current global population of more than seven billion people (USCB, 2015). The use of nitrogen fertilizers to increase crop yields is not without its drawbacks. Water quality and environmental concerns relating to elevated nitrogen in water bodies are varied and include effects such as drinking water contamination and ecosystem degradation (see Appendix A for additional details). The amount and rate of nitrogen moved from the landscape to nearby streams is affected by a number of geospatial, environmental, and watershed characteristics (variables). The route through which water travels from the landscape to the stream (flowpath) is one of these variables. Common flowpaths include groundwater flow, overland flow, and subsurface drain flow. Important flowpaths in a watershed can play a large role in moving nitrogen to streams (Green, 2007) because many forms of nitrogen are highly soluble in water or travel with the water as particulates. The flowpath through which water travels to a stream affects the degree of interaction between the water and the soil, often resulting in water from different flowpaths having different nitrogen concentrations. Water from slow flowpaths (slowflow), such as groundwater flow, can have high nitrogen concentrations in areas where the groundwater is contaminated (Puckett et al., 2008). Water from fast flowpaths (fastflow), such as overland flow, can have low concentrations of nitrogen due to the short interaction between soil and water (Wang and Zhu, 2011). Estimating the fluxes of water to the stream from the respective flowpaths will help expand our understanding of nitrogen transport. In addition to flowpath, a number of other geospatial, environmental, and watershed variables play an important role in determining the amount of nitrogen that is moved with the water on its way to the stream. Although studies have attempted to determine the variables that have the greatest affect on the amount of nitrogen in streams,

those studies often examine a relatively small number of streams, or investigate streams that are within a single hydrologic setting.

In order to improve our understanding of the processes which govern nitrogen movement to streams, and thus, affect the amount of nitrogen measured in streams, four related studies were conducted. The first study compared two commonly used methods of hydrograph separation, a technique used for estimating the contributions of two or more flowpaths to total streamflow. Graphical hydrograph separation via the BFI program (Wahl and Wahl, 1995) and specific conductance-based end-member mixing analysis (SC-EMMA) were compared. An improvement upon the existing method of SC-EMMA for hydrograph separation was then proposed and examined with synthetic data as well as data from two real streams. To confirm and expand upon our current understanding of nitrogen dynamics in actual streams, a third study was conducted in which field observations from six hydrologically complex and anthropogenically modified agricultural watersheds were analyzed and interpreted. Finally, annual total nitrogen loads, flow-weighted average concentrations, and annual watershed yields, as well as data for 90 geospatial, environmental, and watershed variables, were collected from 636 small (<585 km^2) watersheds across the United States in an effort to determine which of the variables are fundamentally important for determining levels of nitrogen in wide variety of streams.

2. Research objectives

The important flowpath(s) that delivers water to a stream has a large effect on the amount of nitrogen that is delivered from the landscape (Hooper et al., 1990, Molénat et al., 2002, Peters, 1994, Rice and Bricker, 1995, Ross et al., 1994), as many forms of nitrogen are highly soluble in water or travel with the water in particulate form. Groundwater discharge to streams is an important flowpath for delivering nitrogen to streams in some hydrologic setting. Thus, it is important to gain a more complete understanding of the quantities and processes of groundwater discharge to streams.

Several methods of hydrograph separation are available to estimate the amount of water in the stream that is from each of the contributing flowpaths. **Chapter 2** presents

the comparison of two very different methods of hydrograph separation when applied to four hydrologically-challenging agricultural streams. The BFI program (Wahl and Wahl, 1995) is a graphical method of hydrograph separation which relies on variability in the stream hydrograph to separate the relatively stable baseflow portion of total streamflow from the short term spikes in streamflow that result from non-baseflow additions. Specific conductance-based end-member mixing analysis (SC-EMMA) is a chemical tracer-based method of hydrograph separation which relies not only on streamflow measurements, but also measurements of specific conductance. The specific conductance measurements from the stream and from each end-member (synonymous with flowpath) are used to mathematically separate the total streamflow into contributions from slowflow and fastflow sources.

Each method of hydrograph separation examined relies on certain assumptions. Although SC-EMMA was able to produce estimates of slowflow when applied to four diverse agricultural streams, it relied on several assumptions, some of which are known to be violated occasionally. As a result, a modified SC-EMMA method was created in an effort to eliminate the assumption of a constant chemical tracer concentration in the fastflow end-member. This new method introduces the constraint that two chemical mass budgets and a water volume budget must be met simultaneously at a daily time step, as opposed to a single chemical mass budget a water volume budget as used with the standard SC-EMMA. The new, ratio-based EMMA was tested in **Chapter 3**. Synthetic data were created to simulate seven scenarios of the timing of the streamflow peak to the minimum concentrations of the two chemicals within two hypothetical streams. One synthetic stream had large inputs of slowflow (ex. groundwater) and the other had large inputs of fastflow (ex. overland flow water). The new method was then applied to two real streams, one stream that had important groundwater inputs and one stream where overland flow inputs were important, to establish if the new method could perform under real world conditions.

The importance and influence of flowpath in determining stream nitrogen concentration was examined in **Chapter 4** among six diverse streams in agricultural watersheds**.** The steams were located in different hydrologic settings, where various

flowpaths were import. A majority of the water in two of the streams was from overland flow contributions, two other streams were dominated by additions of water from subsurface drainage, and groundwater additions to the stream were important in the last two streams. The observed spatial magnitude and temporal variability in nitrate loads, yields, and concentrations were interpreted from the perspective of the influence of flowpaths in these highly modified watersheds. High temporal resolution (15-minute interval) concentration data was also obtained from an additional three streams. These three streams were in similar geographic locations and had similar important flowpath additions as the previously mentioned streams (overland flow, subsurface drain flow, and groundwater flow), so the results were relatable. This high resolution data was examined to compare and contrast the timing of nitrate delivery to streams in relation to the streamflow peak, and to compare the results from the three hydrologically complex streams to current knowledge.

Finally, an extensive data set of small streams throughout the contiguous United States (636 stream) were analyzed in an effort to find the spatial variables (landscape, soil land use, climate, chemical use, and so forth) that strongly affect the amount of nitrogen that is measured in streams. In **Chapter 5**, recursive partitioning and random forest regression were applied to determine which of 90 variables are the most important for determining the amount of total nitrogen in streams. Variables extracted from each method were then used to create a series of multiple linear regression models, which were then applied to a set of validation watersheds. The models served as a test to determine if the extracted variables were significant for determining nitrogen in streams, and how applicable those variables could be when applied to unstudied streams.

References

Green, M.B, 2007. Hydrologic Control of Stream Water N:P Ratios, PhD Thesis, Water Resources Science, University of Minnesota, St. Paul, Minnesota

Hooper, R.P., N. Christophersen, and N.E. Peters. 1990. Modeling streamwater chemistry as a mixture of soil water end-members—an application to the Panola Mountain Catchment, Georgia, USA, J. Hydrol., 116(1-4), 321-343. doi:10.1016/0022-1694(90)90131-G

Molénat, J., P. Durand, C. Gascuel-Odoux, P. Davy, and G. Gruau. 2002. Mechanisms of nitrate transfer from soils to stream in an agricultural watershed of French Brittany, Water Air Soil Pollut., 133(1-4), 161−183. doi: 10.1023/A:1012903626192

Peters, N.E, 1994. Water-quality variations in a forested Piedmont catchment, Georgia, USA. J. Hydrol. 156(1-4), 73-90. doi:10.1016/0022-1694(94)90072-8

Puckett, L.J., C. Zamora, H. Essaid, J.T. Wilson, H.M. Johnson, M.J. Brayton, and J.R. Vogel. 2008. Transport and Fate of Nitrate at the Ground-water/Surface-water Interface. J. Environ. Qual. 37(3), 1034-1050. doi: 10.2134/jeq2006.0550

Rice, K.C. and O.P. Bricker. 1995. Seasonal cycles of dissolved constituents in streamwater in two forested catchments in the mid-Atlantic region of the eastern USA, J. of Hydrol. 170(1-4), 137-158. doi:10.1016/0022-1694(95)92713-N

Ross, D.S., R.J. Bartlett, F.R. Magdoff, and G.J. Walsh. 1994. Flow path studies in forested watersheds of headwater tributaries of Brush Brook, Vermont. Water Resour. Res. 30(9), 2611-2618. doi:10.1029/94WR01490

United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS). 2015. Quick Stats. Washington, D.C. Online at: http://quickstats.nass.usda.gov/?long_desc__LIKE=Corn&x=20&y=6#24D5A0F E-5F7C-3A4B-9523-8FCC678E6BA9. Accessed: 4-25-15.

United States Census Bureau (USCB). 2015. Total Midyear Population for the World: 1950-2050. Washington, D.C. Online at: http://www.census.gov/population/international/data/worldpop/table_population.p hp. Accessed 4-25-15.

Wahl, K.L. and T.L. Wahl. 1995. Determining the Flow of Comal Springs at New Braunfels, Texas. *Texas Water '95*. American Society of Civil Engineers, August 16-17, 1995, San Antonio, Texas, pp. 77-86.

Wang, T. and B. Zhu. 2011. Nitrate loss via overland flow and interflow from a sloped farmland in the hilly area of purple soil, China, Nut. Cyc. In Agroecosystems, 90(3), 309-319. doi: 10.1007/s10705-011-9431-7

Chapter 2: A comparison of high-resolution specific conductance-based end-member mixing analysis and a graphical method for baseflow separation of four streams in hydrologically challenging agricultural watersheds

Published in Hydrologic Processes:

Kronholm, S.C. and P.D. Capel. 2014. A comparison of high-resolution specific conductance-based end-member mixing analysis and a graphical method for baseflow separation of four streams in hydrologically challenging agricultural watersheds. Hydrological Processes.

<u>**Abstract**</u>

Quantifying the relative contributions of different sources of water to a stream hydrograph is important for understanding the hydrology and water quality dynamics of a given watershed. To compare the performance of two methods of hydrograph separation, a graphical program (BFI) and an end-member mixing analysis that used high-resolution specific conductance measurements (SC-EMMA) were used to estimate daily and average long-term slowflow additions of water to four small, primarily agricultural streams with different dominant sources of water (natural groundwater, overland flow, subsurface drain outflow, and groundwater from irrigation). Because the result of hydrograph separation by SC-EMMA is strongly related to the choice of slowflow and fastflow end-member values, a sensitivity analysis was conducted based on the various approaches reported in the literature to inform the selection of end-members. There were substantial discrepancies among the BFI and SC-EMMA, and neither method produced reasonable results for all four streams. Streams that had a small difference in the SC of slowflow compared to fastflow, or did not have a monotonic relationship between streamflow and stream SC posed a challenge to the SC-EMMA method. The utility of the graphical BFI program was limited in the stream that had only gradual changes in streamflow. The results of this comparison suggest that the two methods may be quantifying different sources of water. Even though both methods are easy to apply, they should be applied with consideration of the streamflow and/or SC characteristics of a stream, especially where anthropogenic water sources (irrigation and subsurface drainage) are present.

1. Introduction

Agricultural management alters the way in which water and agricultural chemicals travel through and across the landscape (Dubrovsky *et al*., 2010), oftentimes increasing the amounts of water and chemicals that are transported to surface water and groundwater. Nitrogen fertilizer is commonly applied to agricultural fields to bolster crop yield, but frequently not all of the applied nitrogen is incorporated into the crops (Bijeriego *et al*., 1979, Meisinger *et al.,* 1985, Olson, 1980, Reddy and Reddy, 1993, Townsend and Howarth, 2010). The unused nitrogen (often as nitrate) is exported from the field with the excess water that is moving through the landscape. Nitrate can travel from the field to the stream through many different water flowpaths including surface runoff, subsurface drains, shallow subsurface flow, and groundwater flow. Elevated nitrate concentrations are an issue in streams in agricultural areas across the Nation (Dubrovsky *et al*., 2010), potentially leading to eutrophication and hypoxia in waters nearby and downstream of the site of application (Ribaudoa *et al*., 2005). Determining the important flowpath(s) of water to a stream is important in the management of nitrogen in agricultural watersheds.

Estimating the baseflow of a stream is one way of separating sources of water to that stream. Baseflow is the portion of total streamflow that is not a result of direct surface runoff or storm interflow, and is often largely due to groundwater discharge to the stream. There are numerous methods for separating baseflow from total streamflow. Graphical methods such as the U.S. Geological Survey HYSEP program (Sloto and Crouse, 1996) and the U.S. Bureau of Reclamation Base-Flow Index (BFI) program (Wahl and Wahl, 1995) use daily streamflow measurements to separate baseflow from non-baseflow. The BFI program, for example, mathematically evaluates the hydrograph and separates the slowly varying component of the hydrograph (baseflow) from the more rapidly changing and elevated streamflow events (non-baseflow). Chemical tracer-based methods such as end-member mixing analysis (EMMA) are also frequently used to separate baseflow from total streamflow (Hooper and Shoemaker, 1986, McNamara *et al.,* 1997, Pellerin *et al*., 2008, Pilgrim *et al*., 1979, Robson and Neal, 1990, Sanford *et al*., 2012, Stewart *et al*., 2007). EMMA relies on continuous streamflow data plus

concentration data of one or more chemical constituents in the stream. As a result of the use of chemical data, EMMA may not specifically separate baseflow from non-baseflow, but instead may separate water that is moving slowly through the hydrologic system (slowflow) from water that is traveling quickly through the hydrologic system (fastflow) (Hooper and Shoemaker, 1986). As a portion of total streamflow, slowflow may be the same as baseflow under certain conditions, but it can also be a combination of baseflow water and other water that travels slowly to the stream from other sources (ponds, wetlands, subsurface drains, etc.). Therefore, EMMA identifies the water that has travelled slowly to the stream, whether that water is from groundwater discharge or from other sources. Each method of hydrograph separation can provide valuable information regarding the sources of water to a stream. However, each has a distinctly different way of estimating the sources, which can produce different results for the same stream.

Specific conductance (SC) has been demonstrated to be a useful basis for an EMMA for hydrograph separation into flowpath components (Ali *et al*., 2010, McCarthy and Johnson, 2009, McNamara *et al*., 1997, Pellerin *et al*., 2008, Sanford *et al*., 2012, Smith 2012, Stewart *et al*., 2007). Stream SC data must be measured across a range of flow conditions to accurately characterize the stream and its end-members. SC characterizes the extent of interaction between water and the soil and provides a chemical signature of the source of the water. Precipitation has a characteristic SC of about 10 µS/cm (NADP, 2013). Precipitation that falls on the landscape interacts with soil, resulting in an increase of SC (Pilgrim *et al*., 1979). Water that has had very little interaction with the soil will have a measured SC of less than 100 µS/cm, whereas groundwater will have SC values typically greater than 500 µS/cm (Hem, 1985). The outcome (distribution between fastflow and slowflow) of a specific conductance-based EMMA (SC-EMMA) is strongly influenced by the SC values of the slowflow and fastflow end-members chosen by the model user (McNamara *et al*., 1997, Pellerin *et al*., 2008, Robson and Neal, 1990).

Streams in four, small agricultural watersheds with different dominant flowpaths (Morgan Creek, MD – natural groundwater; Tommie Bayou, MS – overland flow; South Fork Iowa River, IA – subsurface drain outflow; and Granger Drain, WA – groundwater

from irrigation) were chosen as test examples for comparing the results of two different methods of hydrograph separation: graphical-based (using the Baseflow Index program, BFI) and end-member mixing analysis using high-resolution specific conductance measurements (SC-EMMA). Before any comparison of the two methods was conducted, the sensitivity of the SC-EMMA results were analyzed for the choice of the slowflow and fastflow end-members (using both fixed and time-varying values) following various approaches reported in the literature.

The use of high-resolution SC data, in conjunction with EMMA, for hydrograph separation has not been widely used, especially for agricultural watersheds and for time periods of a year or longer. The traditional collection of discrete stream samples for chemical analysis is expensive, time consuming, and provides limited information regarding the temporal nature of stream chemistry. High-resolution (daily, hourly, or less) chemical parameters, however, have the potential to provide high-resolution hydrograph separation when used in an EMMA model (Sanford *et al*., 2012) and could greatly improve the understanding of the sources and timing of water delivered to the stream. The two methods of hydrograph separation (BFI and SC-EMMA), although aimed at similar outcomes, are based on different assumptions and potentially characterize different parcels of water. Estimated slowflow from the two methods, as well as the deviation between the results from the two methods, provide insight into the sources of water to the four streams (slowflow/fastflow or baseflow/non-baseflow, respectively). Of the two methods of hydrograph separation examined, the more appropriate method of hydrograph separation can be chosen based on a priori knowledge regarding the dominant runoff processes in the watersheds.

2. Methods
2.1   Site descriptions

Four hydrologically different watersheds were chosen for this study. The Morgan Creek watershed (32.9 km^2) is located in Maryland. Agriculture comprises 85% of the land within the Morgan Creek watershed, most of which is non-irrigated. Mean annual precipitation within the watershed is 112 cm. Streamflow in Morgan Creek (Figure

2.1A) is supported by natural groundwater throughout the year with additions of storm event flow common (Hancock and Brayton, 2006). Annual streamflow within Morgan Creek has a mean of 0.3 m^3/s. The streamflow and SC from October 1, 2002, to September 30, 2004, were used in this study.

The Tommie Bayou, MS, watershed (15.3 km^2) is also primarily dedicated to agriculture with 86% of the land area planted with agricultural crops (McCarthy *et al.*, 2012). Tommie Bayou is a heavily irrigated watershed during the summer growing season, which supports much of the streamflow during that time (Figure 2.1B). Irrigation water is derived primarily from groundwater. One such irrigation well had a median SC of 1190 µS/cm (USGS Station identification (STAID) 334804090500801) during 2010. Most of the flow in the stream during the non-growing season is derived from storm events or shallow subsurface flow, as there is little connection with the deeper groundwater (McCarthy *et al.*, 2012). Mean annual precipitation within the watershed is 137 cm, with typically less than 30% of that falling during the growing season. Tommie Bayou has an annual average discharge of 0.2 m^3/s, but has occasional, short-term dry periods. The streamflow and SC from October 1, 2006, to September 30, 2008, were used in this analysis.

The South Fork Iowa River (SFIR) watershed (31.1 km^2) is located in central Iowa. Agriculture is the primary land use in the SFIR watershed with 96% of the land primarily in corn and soybeans (McCarthy *et al.*, 2012). The watershed receives an annual mean of 83 cm of precipitation. It is mostly non-irrigated and underlain with extensive networks of subsurface drainage to remove excess water from the shallow subsurface. The water from these drains is the primary source of water to the stream (Figure 2.1C), as there is little to no connection to the deeper groundwater system (McCarthy *et al.*, 2012, Thornburg, 2009). Average annual discharge in the stream is 0.5 m^3/s. The streamflow and SC from October 1, 2006, to September 30, 2008, were used.

Granger Drain watershed (161 km^2) is located in the Yakima Valley of Washington. Land use in the watershed is primarily agriculture (96%), >95% of which is irrigated from approximately mid-March through mid-October (Payne *et al.*, 2007). An annual average of 18.5 cm of precipitation falls within the watershed. Irrigation water in

the Granger Drain watershed is derived from irrigation canals, one of which had a median SC of 99 µS/cm (USGS STAID 12504508) during 2003 and 2004. Over a period of many decades of irrigation, the water table within the Granger Drain watershed has been elevated to the point where subsurface drainage systems have become necessary to remove excess water from the shallow subsurface. The source of streamflow is primarily this shallow groundwater during the non-irrigation season and a combination of groundwater and irrigation runoff during the irrigation season (Figure 2.1D), with much of the irrigation runoff travelling through subsurface drains (Payne *et al.*, 2007). Annual mean streamflow within Granger Drain is 1.0 m^3/s. The streamflow and SC from October 1, 2002, to September 30, 2004, were used.

Field studies in the four watersheds were carried out between 2002 and 2008 (Capel *et al.*, 2008), however, not all areas were sampled during the same time period. Continuous streamflow and SC were measured in the four streams at 15-minute intervals. SC measured in groundwater samples were collected from wells during the growing season. The annual median SC of precipitation was obtained from National Atmospheric Deposition Program sites that were nearest to each of the study watersheds (WA24, 2004; MD13, 2004; IA08, 2008; MS30, 2008; (NADP, 2013)). The SC was measured in overland flow samples (Washington and Iowa only) that were collected from a surface weir (Iowa, USGS STAID 423135093373301) and from the end of a culvert (Washington, USGS STAID 462023120075242) during the growing season (Capel *et al.*, 2008). These samples were collected sporadically, when there was sufficient rainfall/irrigation to produce overland flow.

2.2   SC-EMMA Hydrograph Separation

Streamflow (Q_S) is a combination of various sources of water from individual flowpaths. Specific conductance was used as a chemical tracer in the creation of a simple end-member mixing analysis (SC-EMMA) to separate streamflow into two end-members: water from slowflow (subscript SF) and water from fastflow (subscript FF). This separation can be represented by the following equations:

$Q_S = Q_{SF} + Q_{FF}$ 　　　　　　　　　　　　　　　　　　Eq. 1

and, therefore:

$$Q_S SC_S = Q_{SF} SC_{SF} + Q_{FF} SC_{FF} \qquad \text{Eq. 2}$$

where Q is the water discharge (m^3/s) and SC is specific conductance (µS/cm). By combining equations 1 and 2, the slowflow and fastflow contributions to overall streamflow can be calculated.

$$Q_{SF} = (Q_S SC_S - Q_S SC_{FF}) / (SC_{SF} - SC_{FF}) \qquad \text{Eq. 3}$$

$$Q_{FF} = (Q_S SC_S - Q_S SC_{SF}) / (SC_{FF} - SC_{SF}) \qquad \text{Eq. 4}$$

2.3   Data analysis

2.3.1   Slowflow and Baseflow Indices

For a given length of time, the water volume (V) is the product of streamflow (Q) and time (t). The slowflow index (SFI, in percent), for any length of time, was calculated as

$$SFI = (\Sigma V_{SF} / \Sigma V_S) * 100 \qquad \text{Eq. 5}$$

The BFI program (Wahl and Wahl, 2007) was also used to separate total streamflow within each of the watersheds into two components: baseflow and non-baseflow. The baseflow index (BFI, in percent), for any given length of time, was calculated as

$$BFI = (\Sigma V_{BF} / \Sigma V_S) * 100 \qquad \text{Eq. 6}$$

The BFI program has two input parameters (N and f) that are defined by the user. The method that was used for choosing these parameters is described in Wahl and Wahl (1995). The time increment variable (N) was selected as the value where any subsequent increase in N led to a less pronounced and more linear decrease in estimated baseflow index (BFI) calculated by the BFI program. For this study, N was determined to be 3 for

12

all four watersheds. The f value was set at 0.9 as recommended by the model developers.
The SFI and BFI were calculated for a 2-year period.

Daily values of slowflow were also calculated with SC-EMMA (as percent
slowflow, Eq. 7) and the BFI program (as percent baseflow, Eq. 8) to determine the
similarities and differences of the two methods for separating streamflow into two
components.

$$\% \ SF = (V_{SF} / V_S) * 100 \hspace{4cm} Eq. \ 7$$

$$\% \ BF = (V_{SF} / V_S) * 100 \hspace{4cm} Eq. \ 8$$

2.3.2   Sensitivity analysis of SC-EMMA

In an effort to minimize the quantity of data required for an SC-EMMA, the
sensitivity of the method to the interval between high-resolution data measurements was
analyzed. It was determined that data collected at a daily interval is sufficient to
represent the variability within streamflow and stream SC. The details of this analysis
are in Appendix B.

The sensitivity of SC-EMMA to the user-defined static values for SC_{FF} and SC_{SF}
(equations 3 and 4) was determined by systematically varying these inputs for each site to
show how the choice of static end-member can affect the results of the SC-EMMA and to
provide information regarding the choice of constant end-member values. The potential
values considered for the SC_{FF} input were the median SC of annual precipitation, SC
observed during the highest streamflow, lowest SC observed in the stream excluding
outliers, median SC of overland flow, and median SC of irrigation water (Table 2.1). The
potential values considered for the static SC_{SF} input were the highest SC observed in the
stream excluding outliers, median SC of shallow groundwater, and SC observed during
the lowest streamflow (Table 2.1). The SC values observed in the four streams are
shown in Figures 2.2A-D. Outliers were removed to prevent the results from being
influenced by one or a few anomalous measurements of SC. Outliers were defined as
values greater than the 75th percentile + 1.5 x IQR (interquartile range) and/or less than

the 25th percentile – 1.5 x IQR. A time-variable SC based on stream measurements was also created as a potential SC$_{SF}$.

Continuous, time-variable SC$_{SF}$ values were calculated for each stream following the method used by Sanford *et al.* (2012). The daily, time-variable SC$_{SF}$ values for input into the EMMA were generated by interpolating between the stream SC values from the first day of each month. For times when the stream SC profile over the month was concave down, a surrogate stream SC value was used such that the interpolated SC$_{SF}$ values were always greater than the stream SC values. The resultant values closely followed the trend and magnitude of the stream SC while remaining greater than the stream SC. Figures showing the time-variable SC$_{SF}$ can be seen in Appendix B (Figures B2.2A-D).

3. Results

3.1   Effect of input values on SC-EMMA results

The results of the hydrograph separations with SC-EMMA are presented in Figures 2.1A-D, based on the chosen values for the SC$_{FF}$ and SC$_{SF}$ end-members (Figures 2.2A-D, Table 2.1). The range of potential SC$_{FF}$ and SC$_{SF}$ input values for SC-EMMA in Table 2.1 represent the chemical conditions in the various potential end-members. The values chosen as SC$_{FF}$ and SC$_{SF}$ greatly affected the bi-annual slow-flow index (SFI).

Figures 2.3A-D show surface plots of the variability of the SFI over ranges of potential SC$_{FF}$ and SC$_{SF}$ values. When the SC$_{FF}$ input value was varied between 10 and 150 μS/cm, the change in the SFI for a given SC$_{SF}$ input ranged from 5 to 63% among the four watersheds. This change in the SFI corresponds to the black line parallel to the y-axis on Figures 2.3A-D. When SC$_{SF}$ input values were varied between the minimum and maximum, the change in the SFI for a given SC$_{FF}$ input in each watershed ranged from 16 to 55% among the four watersheds. This change in SFI corresponds to the black line parallel to the x-axis on Figures 2.3A-D. The rate of change in SFI increased as the difference between the chosen SC$_{FF}$ and SC$_{SF}$ values decreased. Table 2.1 reports the average change in the calculated SFI per unit change in input SC over the range of SC

values examined. Morgan Creek, MD, had a much greater change in the SFI per unit change of input SC than the other streams.

3.2   Comparison of hydrograph separations by SC-EMMA and BFI

Constant input values for SC_{SF} and SC_{FF} for each stream were used with the SC-EMMA (Table 2.1) for comparison with the results of the BFI program. The SC_{SF} was chosen to be the highest SC observed in the stream excluding outliers. The SC_{FF} value was selected as the value halfway between the measured precipitation SC (the driver of most fastflow) and the lowest SC measured in the stream. Based on the available data, the input for SC_{FF} was unknown and likely varied with time within fastflow events. The true value for SC_{FF} is expected to be greater than the precipitation SC and less than the lowest SC in the stream, however, in these four watersheds there was not a good rationale for choosing one SC value over the other. Although the SC of precipitation and the lowest SC in the stream are values that can be easily obtained, neither is likely to equal SC_{FF}, especially over an extended period of time.

The bi-annual SFI (from SC-EMMA) and bi-annual BFI (from the BFI program) differed by 5% in Morgan Creek, MD (natural groundwater). The SFI and BFI differed by almost 20% in Tommie Bayou, MS (the overland flow site) and SFIR, IA (subsurface drain outflow) and differed by 55% in Granger Drain, WA (groundwater from irrigation). The SFI was greater than the BFI for all watersheds except Granger Drain, WA (Table 2.1).

There were substantial differences in the daily slowflow (m^3/s) calculated by SC-EMMA and daily baseflow (m^3/s) from the BFI program (Figures 2.4A-D). Figures 2.4A-D compare 2 years of calculated daily slowflow and baseflow relative to 2:1, 1:1, and 1:2 lines of agreement. For both Morgan Creek, MD (Figure 2.4A) and Granger Drain, WA (Figure 2.4D), the slowflow from SC-EMMA was consistently less than baseflow from the BFI program (66% and 99% of the time, respectively). The two distinct groupings of points that appear on Figure 2.4D correspond to the irrigation season (mid-March through mid-October) and non-irrigation season within the Granger Drain, WA watershed, with the BFI program overestimating SC-EMMA to a greater extent during the irrigation season. For Tommie Bayou, MS (Figure 2.4B), 52% of the

15

time the estimated slowflow was greater than the estimated baseflow, but the results were highly scattered. For SFIR, IA (Figure 2.4C), slowflow from SC-EMMA was greater than baseflow from the BFI program 57% of the time, but the values fell much closer to the 1:1 line compared to Tommie Bayou.

A direct comparison of the daily percent of streamflow resulting from slowflow calculated using SC-EMMA (% slowflow) and daily percent of streamflow resulting from baseflow calculated using the BFI program (% baseflow) also showed substantial differences for the 2 years of record (Figures 2.5A-D). For Morgan Creek, MD (Figure 2.5A) and SFIR, IA (Figure 2.5C), the % slowflow showed less variability than the % baseflow. The standard deviation for % slowflow was 16%, but was 28% for % baseflow in Morgan Creek. In SFIR the standard deviation was 14% for % slowflow and 28% for % baseflow. In SFIR, IA, the % slowflow ranged from 60 - 100% for 89% of the days, but the % baseflow had a much larger range of between 20 and 100% for 94% of the days. For Tommie Bayou, MS (Figure 2.5B), the two methods produced widely scattered estimates of % slowflow and % baseflow. The mean and standard deviation of % baseflow was 51($\pm$ 36%), whereas % slowflow had a mean and standard deviation of 46($\pm$ 25%) over the same period. For Granger Drain (Figure 2.5D), the % slowflow had greater variability than the % baseflow. The BFI program estimated that the % baseflow was 97 $\pm$ 6% of the total streamflow throughout the entire length of the data record. In contrast, the % slowflow was much less, averaging 51 $\pm$ 24% over the 2-year period with two distinct clusters centered on 30% and 80%, corresponding to the irrigation and non-irrigation seasons.

The SC-EMMA and the BFI program generated hydrographs of slowflow and baseflow that were quite different (Figure 2.1A-D). In all four watersheds, two slowflow hydrographs were calculated from SC-EMMA using different assumptions (a constant SC_{SF} and a time-variable SC_{SF}). The SC-EMMA slowflow hydrographs calculated with both assumptions were much flashier than the baseflow hydrograph from the BFI program. In Morgan Creek, MD, Tommie Bayou, MS, and SFIR, IA, both of the slowflow hydrographs from SC-EMMA increased dramatically with the streamflow hydrographs during streamflow events greater than 1.0 m^3/s, whereas the baseflow

16

hydrograph from the BFI remained comparatively stable throughout many of these events (Figures 2.1A-C). (For example, see SFIR, IA between the dates of April 1 and July 1, 2007 Figure 2.1C). It should also be noted that in each watershed, the slowflow hydrograph using a time-variable SC_{SF} resulted in an overall increase in estimated slowflow compared to the hydrograph produced with a constant SC_{SF} (Figures 2.1A-D, Table 2.1).

For Granger Drain, both the SC-EMMA and BFI methods estimated that most of the streamflow was a result of slowflow/baseflow during the non-irrigation season ($78 \pm$ 12% from SC-EMMA and $98 \pm 5.0\%$ from BFI). During the irrigation season, the BFI program still calculated almost all of the water in the stream as baseflow contributions (97 ± 4.5 %), however, the SC-EMMA utilizing a constant SC_{SF} calculated that the percent slowflow decreased to about one-third of total streamflow ($32 \pm 4.7\%$). This finding suggests that much of the water in the stream during the irrigation season travelled through fast flowpaths to the stream. The SC-EMMA utilizing a time-variable SC_{SF} fell between the BFI results and the results from the constant SC_{SF} SC-EMMA, but more closely matched the BFI results suggesting that a majority of the water in the stream during the irrigation season travelled through slow flowpaths (Figure 2.1D).

4. Discussion

4.1   SC-EMMA as a tool for hydrograph separation – choice of fastflow end-member

As a technique for hydrograph separation, the success of EMMA depends heavily on the input parameters chosen by the user. The multiple choices for both SC_{FF} and SC_{SF} contribute to the uncertainty of the results from the SC-EMMA method for many streams, since there are numerous options (Table 2.1). The selection of the SC_{FF} value had a greater influence on the SC-EMMA results than did the selection of the SC_{SF} value. Given their relative magnitudes, a unit change in SC_{FF} has greater effect on SFI than did the same change in SC_{SF} (Table 2.1 and Figure 2.3).

The SC_{FF} end-member value was chosen as the average of precipitation SC and the lowest stream SC in an attempt to aggregate the various fastflow inputs. The true

SC$_{FF}$ is expected to be between those two values. SC$_{FF}$ input values for SC-EMMA used by others include SC of rainfall (Pellerin *et al.,* 2008 and Sanford *et al*., 2012), overland flow (McNamara *et al*., 1997), or the average stream SC during the highest 1% of stream flow (Miller et al., 2014). Precipitation, in the range of 5-20 µS/cm for the watersheds of this study, is usually the driver of fastflow and offers a baseline for the potential magnitude of SC$_{FF}$. However, a choice of SC$_{FF}$ equal to precipitation would generally underestimate the actual SC$_{FF}$ value, and thus, overestimate slowflow. SC increases rapidly after precipitation reaches the ground and begins to interact with the soil (Pilgrim *et al*. 1979, McNamara *et al*. 1997), resulting in fastflow components that have elevated SC in comparison to pure rainfall (Table 2.1). Irrigation can also be a driver of fastflow in a watershed and can be a potential candidate for SC$_{FF}$. Irrigation water that is derived from snowmelt and channeled to agricultural areas has a low SC. Granger Drain, WA, is an example of this situation (Figure 2.2D). In contrast, irrigation water can also be derived from high SC groundwater and would not be appropriate as the basis for the SC$_{FF}$ value because it cannot be differentiated from the SC$_{SF}$. Tommie Bayou, MS, is an example of this situation (Figure 2.2B) and is not a good candidate for hydrograph separation by SC-EMMA.

Furthermore, using SC values observed in the stream as the basis for SC$_{FF}$ can be problematic. At any time during a high flow event, the water in the stream is not solely fastflow water, but a combination of water from slow and fast flowpaths (Stewart *et al*., 2007). Pinder and Jones (1969) and Sklash and Farvolden (1979) showed that even during times of very high discharge in their study streams, groundwater inputs were not insignificant. Therefore, the SC in the stream during high flow events is often greater than the true SC$_{FF}$ due to mixing with the higher SC water from slowflow sources, causing SC-EMMA to underestimate the contributions of slowflow. This effect was seen in the Granger Drain, WA, watershed. When the SC$_{FF}$ input value was set equal to the lowest stream SC value, SC-EMMA indicated that there was almost no slowflow water contributed to the stream during the elevated flows corresponding with the irrigation season (see Appendix B, Figure B2.3). There is however, a substantial input of slowflow water to the stream, even during the irrigation season. McCarthy and Johnson (2009) and

Payne *et al.* (2007) explained that during the irrigation season in Granger Drain, WA, and in a sub-watershed within the Granger Drain watershed, the water in the stream is composed of considerable inputs from both slowflow and fastflow water.

All of the choices above rely on the assumption that SC_{FF} can be represented with a constant value, which is not realistic and can introduce substantial error in the estimation of slowflow. Another approach is the creation of a time-varying SC_{FF} over the course of the year, however, this approach would require knowledge (based on field observations), in both time and space, of all fastflow sources within the watershed – an undertaking that is not possible in most places. Therefore, for practical reasons, a constant representative value of SC_{FF} will be chosen in most situations.

4.2   SC-EMMA as a tool for hydrograph separation – choice of slowflow end-member

Slowflow sources to the stream are most commonly groundwater, but can include water from wetlands, ponds, subsurface drains, and other point sources (for time periods distant from rain or irrigation events). The potential SC_{SF} input values for SC-EMMA can have a much greater magnitude and higher variability compared to SC_{FF}. The SC_{SF} values used by other researchers as the input into a SC-EMMA include groundwater (Ogunkoya and Jenkins, 1993), SC in the stream at the lowest streamflow (Hooper and Shoemaker, 1986 and McNamara *et al.*, 1997), average SC in the stream during the lowest 1% of flows (Miller et al., 2014), and a time-variable value based on measured stream SC (Sanford *et al.*, 2012). For these four streams, these options yielded a wide range of potential SC_{SF} values (Table 2.1).

Each of the approaches to choosing the SC_{SF} value has its limitations. The use of the groundwater SC as the basis of SC_{SF} can be problematic due to the variability of groundwater chemistry. Groundwater well samples are point samples from a specific portion of an aquifer, and may not be representative of all groundwater moving into the streambed. McCarthy and Johnson (2009) and Stewart *et al.* (2007) found that the chemistry of groundwater sampled from wells varied substantially, even when the wells were separated by short distances. This difference in chemistry was also observed for the four watersheds included in this study. The approach of defining SC_{SF} as equal to the SC

measured during the lowest (or near lowest) streamflow has been successfully used in many streams, but can have limitations in some streams. In some cases, the stream SC, across a range of flows, can be greater than the SC during the lowest flow. Figure 2.2 shows that the lowest streamflows are not always associated with the highest stream SC. In SFIR, IA, the maximum stream SC did not occur at the lowest streamflow, but instead coincided with flows that were roughly two orders of magnitude greater. If the stream SC values during the lowest streamflow were used for the SC_{SF} in these streams, the SC-EMMA would overestimate slowflow (in some cases greater than 100%).

In this study, the highest measured SC in each stream (minus outliers) was chosen as the constant SC_{SF} value. This choice as SC_{SF}, while similar in magnitude compared with the other choices, prevented the SC-EMMA calculation from producing unrealistic results since the stream SC is never greater than the chosen SC_{SF}. According to Christophersen and Hooper (1992), the values chosen for end-member inputs should be extreme values, potentially outside of the values typically observed, in order to explain the mixture of waters (fastflow/slowflow) seen in the stream.

A time-variable SC_{SF} value was also tested as the slowflow input (SC_{SF}). The time-variable SC_{SF} produced slowflow estimates that were generally greater than the constant SC_{SF} approach (Figure 2.1A-D and Figure B2.4A-D in Appendix B). The time-variable SC_{SF} approach in effect broadens the interpretation of slowflow to include more, faster flowpath contributions to the stream and (or) multiple types of aquifer contributions to the stream. By varying SC_{SF} and holding it near the measured SC in the stream, the EMMA calculated a higher slowflow compared to using a constant SC_{SF} where the stream SC can diverge from SC_{SF} at times.

4.3   Comparison of the BFI program and SC-EMMA

One important aspect of exploring these two methods of hydrograph separation is to understand the differences and to recognize the unique information and insight that each method can provide about how water travels from the landscape to the stream. Neither method should be applied to a watershed without considering the limitations of the method with respect to the physical characteristics of the watershed. The differences in outcome between the two methods could represent the distinction between baseflow –

slowly varying flow (estimated by the BFI program) and slowflow - the aggregation of multiple slow flowpaths - estimated by SC-EMMA.

Generally, SC-EMMA estimated a greater portion of slowflow water during high flow events than did the BFI program. The BFI program, on the other hand, often estimated more baseflow than SC-EMMA during low flow events, frequently estimating 100% baseflow. Unlike the BFI program, SC-EMMA never estimated that 100% of streamflow was from slowflow (Figure 2.5A-D). The different mechanisms of the models account for these observations. The BFI program mathematically establishes a series of points on the streamflow hydrograph between peaks in streamflow when creating a baseflow hydrograph. The interpolation between these points, each corresponding to an estimation of 100% baseflow, produces a baseflow hydrograph. As a result, the baseflow hydrograph is relatively flat in comparison to the SC-EMMA slowflow hydrograph (Figure 2.1A-D). The "flashier" SC-EMMA slowflow hydrograph is a result of the variability of SC in the stream.

Under certain hydrological settings, one of the methods may be better suited for hydrograph separation than the other. The BFI was designed to be applied in streams where there is a strong, continual groundwater component, plus streamflow fluctuations resulting from storm event water (Wahl and Wahl, 1995). Streams with different hydrologic regimes can pose challenges to the BFI program. The BFI program did not work well when applied to Granger Drain, WA (Figure 2.1D). The annual hydrograph in the Granger Drain watershed exhibited a high-flow period that corresponded to the irrigation season and a low-flow period that corresponded to the non-irrigation season. Intra- and inter-daily variation in flow within either period was relatively small. The stability of the seasonal streamflow and the absence of high flow peaks caused the BFI program to incorrectly calculate that nearly 100% of the streamflow throughout the year was from baseflow (Figure 2.5D). McCarthy and Johnson (2009) showed that the irrigation water (moving to the stream as fastflow) had a distinctly different SC signature than the slowflow water derived from groundwater. The SC-EMMA method using a constant SC_{SF} responded to the dramatic fluctuation of SC in Granger Drain between the irrigation season and the non-irrigation season, showing that there were two major

sources of water to the stream during the irrigation season – a result not produced with the BFI program. As a result of the stability of the hydrograph, the BFI program is not an appropriate method for hydrograph separation in the Granger Drain watershed. Interestingly, the SC-EMMA method that utilized a time-varying SC_{SF} produced results more similar to those of the BFI program (Figure 2.1D), suggesting a time-variable SC_{SF} should also be carefully applied.

The SC-EMMA works best where there is a strong, monotonic relation between stream SC and streamflow. Though this condition is not found in any of the four streams in this study, it is best approximated in SFIR, IA, and Granger Drain, WA (Figures 2.2C and D). In contrast, Tommie Bayou, MS, (Figure 2.2B) shows no relation between stream SC and streamflow, and is therefore a very poor candidate for the use of SC-EMMA. The lack of relation is largely due to the use of groundwater for irrigation in the basin during a portion of the year (McCarthy *et al.*, 2012). This irrigation water, although travelling through relatively fast flowpaths, already had a high SC due to its residence time in the aquifer, and is contrary to the assumption that fastflow water has lower SC than slowflow water.

The utility of the SC-EMMA approach is limited in streams where there is not a substantial difference in the stream SC between low flow and high flow events (that is, there is not a large difference between SC_{SF} and SC_{FF}). As an example, the use of SC-EMMA is tenuous in Morgan Creek, MD, where there is comparatively little difference in stream SC during the lowest 5% of flows (181 $\pm$ 9.3 µS/cm) and the highest 5% of flows (111 $\pm$ 28 µS/cm, Figure 2.2A). The addition of low SC fastflow water does not result in a substantial decrease in stream SC, limiting the ability of the SC-EMMA to be used to perform a hydrograph separation. This was also observed by Sanford *et al.* (2012) in their studies, which were conducted in Virginia. In streams where the difference in SC_{SF} and SC_{FF} is small, small changes in the chosen input values will produce large changes in the estimated SFI (Figure 2.3A). The BFI program does not consider SC measurements, and therefore, could be the approach that is favored in streams where the difference between SC_{SF} and SC_{FF} is low.

Somewhat surprisingly, the SC-EMMA and the BFI program give similar results for SFIR, IA – the stream influenced by subsurface drains (Figure 2.4C). Streamflow in the SFIR during low-flow periods largely comes from artificial subsurface drainage, not groundwater (Thornburg, 2009), and streamflow during high-flow periods comes from a combination of discharge from surface ditches and subsurface drains. Although the BFI program is meant to estimate groundwater additions to the stream, not drainage source additions, the program is able to clearly differentiate and separate elevated flows in the stream hydrograph resulting from precipitation events (Figure 2.1C) because BFI works by filtering out the peaks in streamflow, no matter their source. The SC-EMMA is also able to effectively separate elevated streamflows. Much of the precipitation that falls within the Iowa watershed moves quickly through the landscape both through runoff flowpaths to surface ditches and through the soil to the subsurface drains. This fast movement through the landscape results in little interaction between the water and soil. Water that remained in the subsurface for longer periods of time substantially increased in SC (Smith, 2012). This water slowly drains to the stream and appears as baseflow/slowflow to both the BFI program and the SC-EMMA.

The two methods compared here are easy to apply and have long histories of use in the literature for the separation of stream hydrographs into component sources. The opportunities to utilize high-resolution SC data for an SC-EMMA hydrograph separation will continue to increase as the number of high-resolution water quality monitors increases. The SC-EMMA has the potential to allow for responsive, high-resolution hydrograph separation, but has substantial limitations in streams that have only a small difference in the SC in slowflow compared to fastflow, do not have monotonic relationships between streamflow and stream SC, or are located in areas where groundwater-derived irrigation is the driver of fastflow. The uncertainty and somewhat subjective choice of SC_{FF} and SC_{SF} for the SC-EMMA also adds to the uncertainty of its results. Graphical methods, like the BFI program, have limitations in streams that have only gradual changes in streamflow and are not recommended for the examination of short-term events. Neither method produced reasonable results for all four streams that were studied, although it should be noted that the four streams were chosen such that they

would present a challenge and test the limitations of both methods. Both methods should be applied to hydrograph separation with consideration of the flow and SC characteristics of the stream.

Differentiating baseflow or slowflow from total streamflow benefits water quality monitoring, modeling, and management because it aids our understanding of the movement of water and waterborne chemicals through the groundwater system and other slowflow pathways to the stream. The two methods, although aimed at similar outcomes, are based on different assumptions. The comparison of values of the Baseflow Index (BFI) from the graphical program and the analogous slowflow index (SFI) from SC-EMMA showed substantial difference in both the daily and the average long-term results for all four streams. The results, together with the previous knowledge of the hydrology of these well-studied watersheds, strongly suggest that the BFI and the SC-EMMA methods quantify different parcels of water; that is, baseflow is not always the same as slowflow.

References

Ali, G.A., A.G. Roy, M.C. Turmel, and F. Courchesne. 2010. Source-to-stream connectivity assessment through end-member mixing analysis. *Journal of Hydrology* **392**(3) : 119-135. doi: 10.1016/j.jhydrol.2010.07.049

Bijeriego, M, R.D. Hauch, and R.A. Olson. 1979. Uptake, translocation, and utilization of ^{15}N-depleted fertilizer in irrigated corn. *Soil Science. Society of America Journal* **43**(3) : 528-533. doi: 10.2136/sssaj1979.03615995004300030020x

Capel, P.D., K.A. McCarthy, and J.E. Barbash. 2008. National, holistic, watershed-scale approach to understand the sources, transport, and fate of agricultural chemicals: Supplemental material section. *Journal of Environmental Quality* **37**(3) : 983–993. doi: 10.2134/jeq2007.0226

Christophersen, N. and R.P. Hooper. 1992. Multivariate analysis of stream water chemical data: The use of principal components analysis for the end-member mixing problem. *Water Resources Research* **28**(1) : 99-107. doi: 10.1029/91WR02518

Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.M. Gronberg, P.A. Hamilton, K.J. Hitt, D.K. Mueller, M.D. Munn, B.T. Nolan, L.J. Puckett, M.G. Rupert, T.M. Short,

N.E. Spahr, L.A. Sprague, and W.G. Wilber. 2010. *The quality of our Nation's waters—Nutrients in the Nation's streams and groundwater, 1992–2004*. U.S. Geological Survey Circular 1350, Reston VA.

Hancock, T.C. and M.J. Brayton. 2006. *Environmental setting of the Morgan Creek Basin, Maryland, 2002–04*. U.S. Geological Survey Open-File Report 2006–1151, Reston VA.

Hem, J.D. 1985. *Study and interpretation of the chemical characteristics of natural water (Vol. 2254)*. Department of the Interior, US Geological Survey, Reston VA.

Hooper, R.P. and C.A. Shoemaker. 1986. A comparison of chemical and isotopic hydrograph separation. *Water Resources Research* **22**(10) : 1444-1454. doi: 10.1029/WR022i010p01444

McCarthy, K.A. and H.M. Johnson. 2009. *Effect of agricultural practices on hydrology and water chemistry in a small irrigated catchment, Yakima River basin, Washington*. U.S. Geological Survey Scientific Investigations Rep. 2009–5030, Reston, VA.

McCarthy, K.A., C.E. Rose, and S.J. Kalkhoff. 2012. *Environmental settings of the South Fork Iowa River Basin, Iowa, and the Bogue Phalia Basin, Mississippi, 2006–10*. U.S. Geological Survey Scientific Investigations Rep. 2012–5021, Reston, VA.

McNamara, J.P., D.L. Kane, and L.D. Hinzman. 1997. Hydrograph separations in an Arctic watershed using mixing model and graphical techniques. *Water Resources Research* **33**(7) : 1707-1719. doi: 10.1029/97WR01033

Meisinger, JJ, Bandel VA, Stanford G, and Legg JO. 1985. Nitrogen utilization of corn under minimal tillage and moldboard plow tillage: I. Four year results using labeled fertilizer on an Atlantic Coastal Plain soil. *Agron. Journal* **77**(4) : 602–611. doi:10.2134/agronj1985.00021962007700040022x

Miller, M.P., D.D. Susong, C.L. Shope, B.J. Stolp, V.M. Heilweil, and T. Marston. 2014. Continuous estimation of baseflow contribution to snowmelt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach. *Water Resources Research*

National Atmospheric Deposition Program (NADP) (NRSP-3). 2013. NADP Program Office, Illinois State Water Survey, 2204 Griffith Dr., Champaign, IL 61820. Accessed from: http://nadp.sws.uiuc.edu/data/ntndata.aspx

Ogunkoya, O.O. and A. Jenkins. 1993. Analysis of storm hydrograph and flow pathways using a three-component hydrograph separation model. *Journal of Hydrology* **142**(1) : 71-88. doi: 10.1016/0022-1694(93)90005-T

Olson, R.V. 1980. Fate of tagged nitrogen fertilizer applied to irrigated corn. *Soil Sci. Soc. Am. J.* **44**(3) : 514-517. doi: 10.2136/sssaj1980.03615995004400030015x

Payne, K.L., H.M. Johnson, and R.W. Black. 2007. *Environmental setting of the Granger Drain and DR2 Basins, Washington 2003–2004*. U.S. Geological Survey Scientific Investigations Rep. 2007–5102, Reston, VA.

Pellerin, B.A., W.M. Wollheim, X. Feng, and C.J. Vörösmarty. 2008. The application of electrical conductivity as a tracer for hydrograph separation in urban catchments. *Hydrological Processes* **22**(12): 1810-1818. doi: 10.1002/hyp.6786

Pilgrim, D.H., D.D. Huff, and T.D. Steele. 1979. Use of specific conductance and contact time relations for separating flow components in storm runoff. *Water Resources Research* **15**(2) : 329-339. doi: 10.1029/WR015i002p00329

Pinder, G.F. and J.F. Jones. 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. *Water Resources Research* **5**(2) : 438-445. doi: 10.1029/WR005i002p00438

Reddy, G.B. and K.R. Reddy. 1993. Fate of Nitrogen-15 Enriched Ammonium Nitrate Applied to Corn, *Soil Sci. Soc. Am. J*. **57**(1) : 111-115. doi:10.2136/sssaj1993.03615995005700010021x

Ribaudoa, M.O., R. Heimlich, and M. Peters. 2005. Nitrogen sources and Gulf hypoxia: potential for environmental credit trading. *Ecological Economics* **52**(2) : 159-168. doi:10.1016/j.ecolecon.2004.07.021

Robson, A. and C. Neal. 1990. Hydrograph separation using chemical techniques: an application to catchments in mid-Wales. *Journal of Hydrology* **116**(1) : 345-363. doi: 10.1016/0022-1694(90)90132-H

Sanford, W.E., D.L. Nelms, J.P. Pope, and D.L. Selnick. 2012. *Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis*. U.S. Geological Survey Scientific Investigations Report 2011–5198, Reston, VA.

Sklash, M.G. and R.N. Farvolden. 1979. The role of groundwater in storm runoff. *Journal of Hydrology* **43**(1) : 45-65. doi: 10.1016/0022-1694(79)90164-1

Sloto, RA and Crouse MY. 1996. HYSEP: A computer program for streamflow hydrograph separation and analysis: U.S. Geological Survey Water-Resources Investigations Report 96-4040, Reston, VA.

Smith, E.A. 2012. *Spatial and temporal variability of preferential flow in a subsurface-drained landscape in north-central Iowa*. Doctoral dissertation, Water Resources Science. University of Minnesota, St. Paul, MN

Stewart, M., J. Cimino, and M. Ross. 2007. Calibration of base flow separation methods with streamflow conductivity. *Ground Water* **45**(1) : 17-27. doi: 10.1111/j.1745-6584.2006.00263.x

Thornburg, J. 2009. *Temporal and spatial variability in nitrate and water quality parameters of subsurface drains in an agricultural stream*. MS Thesis. Water Resources Science. University of Minnesota, St. Paul, MN

Townsend, A.R. and R.W. Howarth. 2010. Fixing the global nitrogen problem. *Scientific American* **302**(2) : 64-71. doi:10.1038/scientificamerican0210-64

Wahl, K.L. and T.L. Wahl. 1995. Determining the Flow of Comal Springs at New Braunfels, Texas. *Texas Water '95*. American Society of Civil Engineers. August 16-17, 1995. San Antonio, Texas : 77-86.

Wahl, T.L. and K.L. Wahl. 2007. BFI: A Computer Program for Determining an Index to Base Flow (Version 4.15), [Software]. http://www.usbr.gov/pmts/hydraulics_lab/twahl/bfi/ Accessed on 4 November 2013.

Table 2.1: Possible SC-EMMA input values (SC_{FF} and SC_{SF}) for the four study watersheds. The bi-annual SFI and BFI as calculated by SC-EMMA (using the specified input values) and the BFI program, respectively, along with the change in SC-EMMA SFI resulting from a one unit change in SC_{FF} or SC_{SF} input.

		Morgan Creek, MD	Tommie Bayou, MS	South Fork Iowa River (SFIR), IA	Granger Drain, WA
	SC_{FF} (values used in this study)	**41**	**35**	**67**	**149**
	Median rainfall SC	21	10.3	10.6	5.4
Potential input values for SC_{FF} (μS/cm)	SC during highest flow	80	156	129	335
	Lowest stream SC	61	60	123	293
	Median overland flow SC	No data	No data	81 [1]	339 [2]
	Irrigation water SC	NA	1190	NA	99 [3]
	SC_{SF} (Highest stream SC, values used in this study)	**236**	**904**	**836**	**842**
Potential input values for SC_{SF} (μS/cm)	Median groundwater SC	168	1098	793	727
	SC during lowest flow	184*(185#)	392*(672#)	541*(574#)	674*(688#)
	2-year BFI (%)	46	7.6	45	97
	2-year SFI (%) [4]	51	28	64	42
	2-year SFI (%) w/ variable SC_{SF}	70	44	73	86
	Δ 2-year SFI per SC_{FF} μS/cm	0.55	0.11	0.05	0.08
	Δ 2-year SFI per SC_{SF} μS/cm	0.33	0.04	0.07	0.07

1 Range 33 to 395 μS/cm
2 Range 118 to 641 μS/cm
3 Range 79 to 126 μS/cm
4 Ranges of SFI for all possible input combinations can be seen in Table B2.1 in Appendix B
* median of all SC values if there were multiple SC measurements during times of equal flow
# highest SC value if there were multiple SC measurements during times of equal flow
NA: not applicable

(A) Morgan Creek, MD
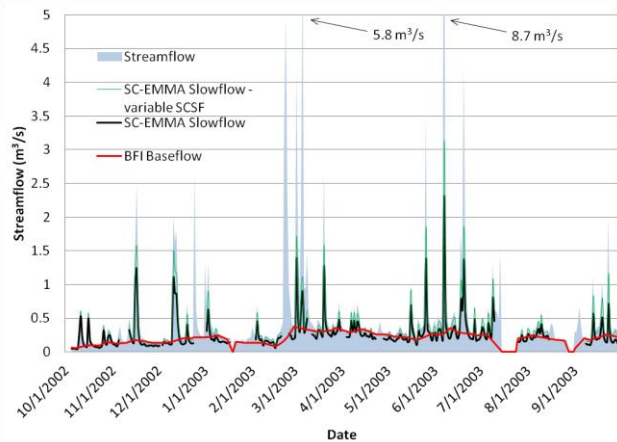
(B) Tommie Bayou, MS
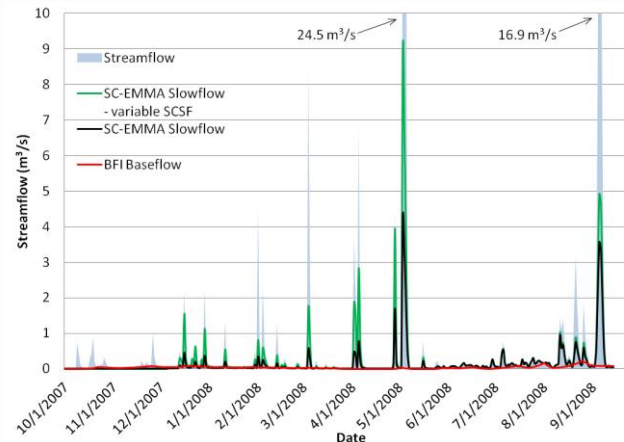
(C) South Fork Iowa River (SFIR), IA
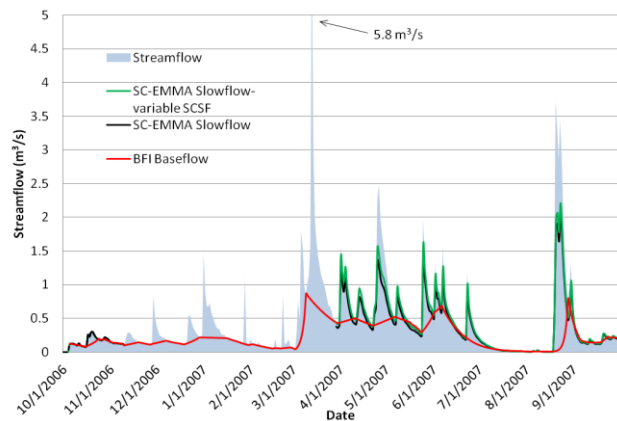
(D) Granger Drain, WA
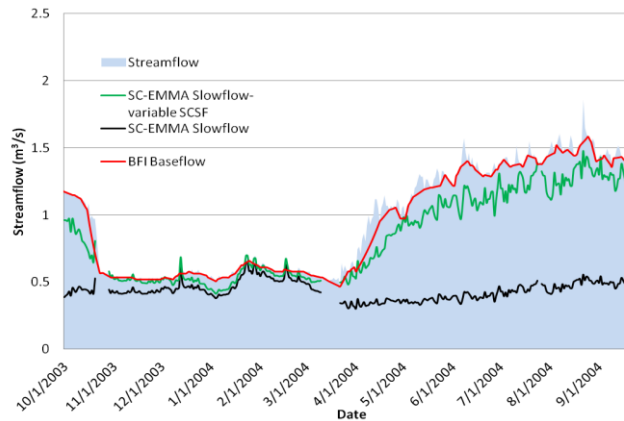
Figure 2.1: Total streamflow hydrographs compared to the slowflow/baseflow hydrographs calculated by SC-EMMA and the BFI program for the four agricultural streams (only one water year displayed). Input values (SC_{FF} and SC_{SF}) for SC-EMMA are provided in Table 2.1. Gaps in the hydrograph represent no data.

29

(A) Morgan Creek, MD
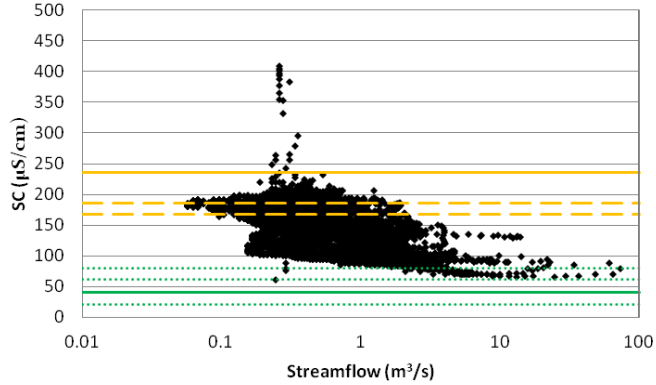
(B) Tommie Bayou, MS
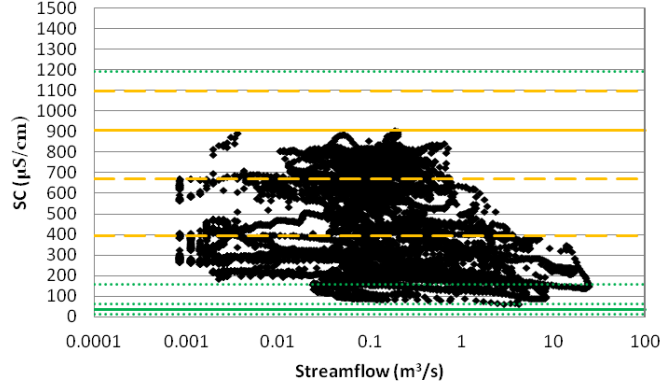
(C) South Fork Iowa River (SFIR), IA
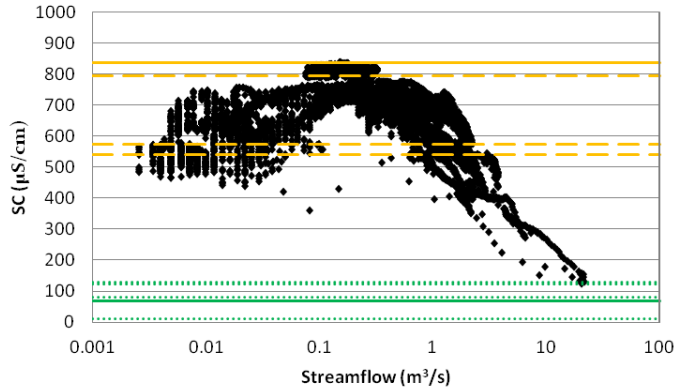
(D) Granger Drain, WA
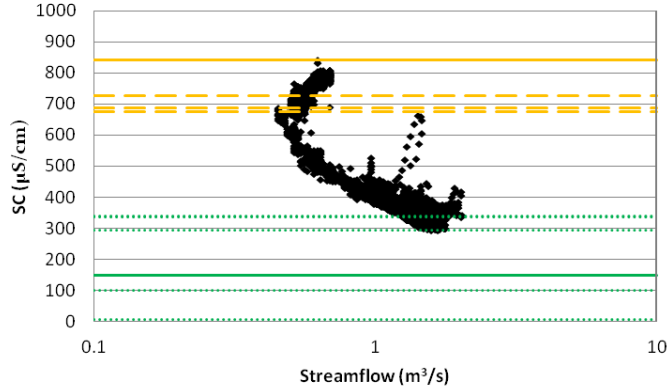
Figure 2.2: Hourly specific conductance (SC) in relation to discharge values over 2 years for four agricultural streams. Horizontal lines represent SC of potential input values for SC-EMMA (Table 2.1). Dotted lines are potential SC_{FF} inputs, dashed lines are potential SC_{SF} inputs (time-variable SC_{SF} not included), and solid lines are the chosen input values. Note the difference in y-axis range values.

(A) Morgan Creek, MD
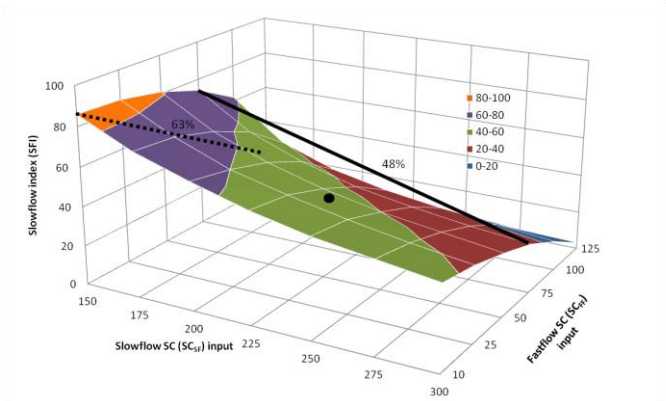
(B) Tommie Bayou, MS
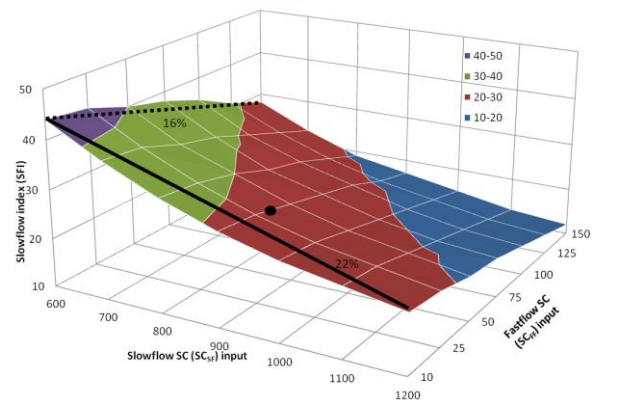
(C) South Fork Iowa River (SFIR), IA
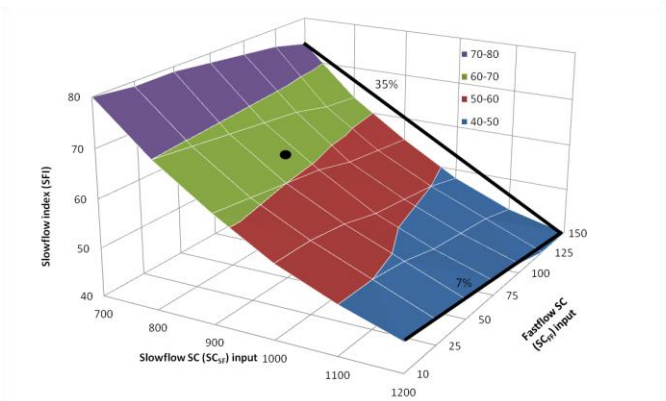
(D) Granger Drain, WA
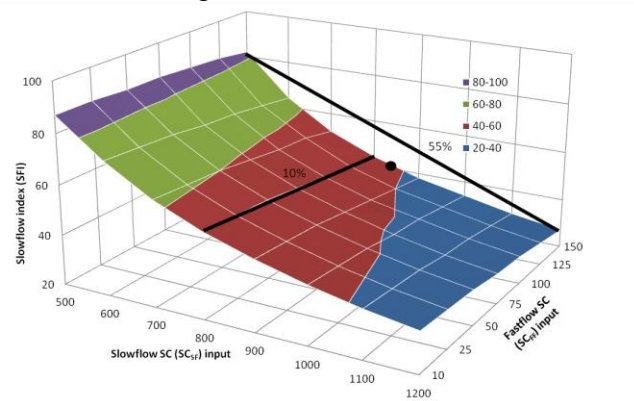
Figure 2.3: Surface plot of the slowflow index (SFI, in percent) for a 2-year period within four agricultural watersheds using a range of input values for SC_{SF} and SC_{FF}. The black dot on each graph is at SC_{SF} and SC_{FF} chosen in this study. The black lines and corresponding percentages represent the maximum change in SFI when holding one input constant and varying the other input between its minimum and maximum values. SFI was calculated using a 1-day time interval between SC values. Note the differing values on the x-axes, y-axes, and z-axes.

31

(A) Morgan Creek, MD
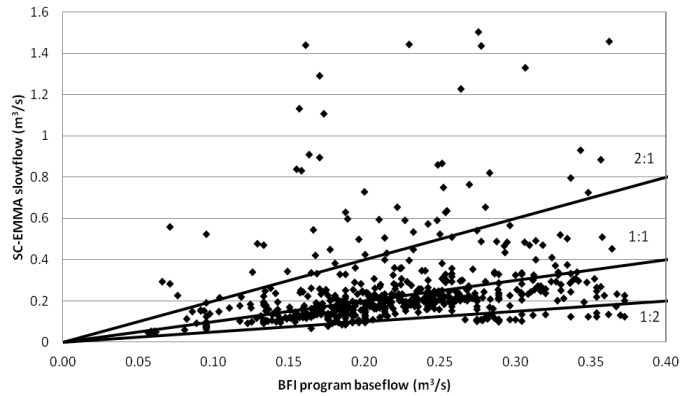
(B) Tommie Bayou, MS
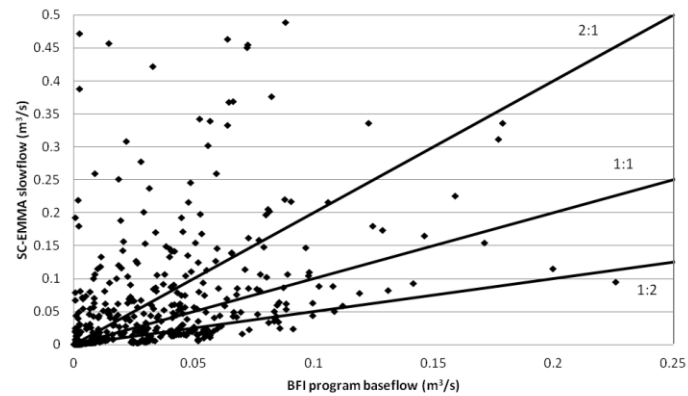
(C) South Fork Iowa River (SFIR), IA
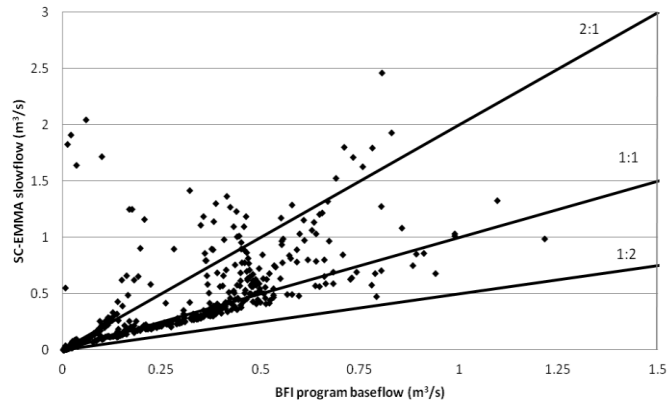
(D) Granger Drain, WA
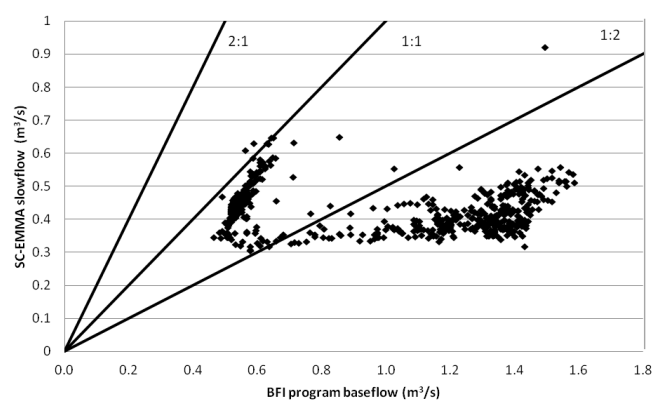
Figure 2.4: Daily SC-EMMA slowflow (m^3/s) compared to the BFI program baseflow (m^3/s) for a 2-year period for four agricultural steams. 1:1 line represents identical values calculated by SC-EMMA and BFI program. 2:1 line represents SC-EMMA over-estimating the BFI program by a factor of 2. 1:2 line represents the BFI program over-estimating SC-EMMA by a factor of 2. Input values (SC$_{FF}$ and SC$_{SF}$) for SC-EMMA are in Table 2.1.

32

(A) Morgan Creek, MD
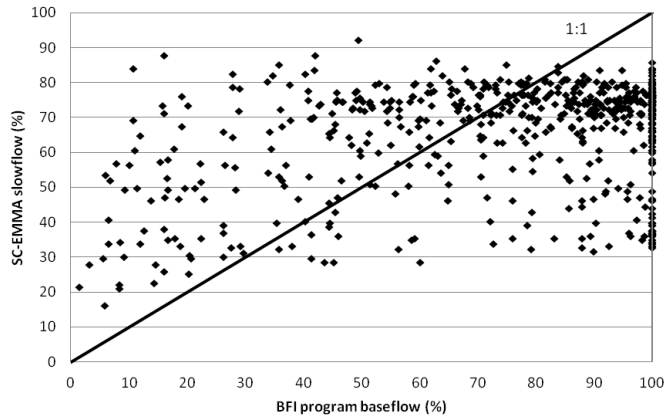
(B) Tommie Bayou, MS
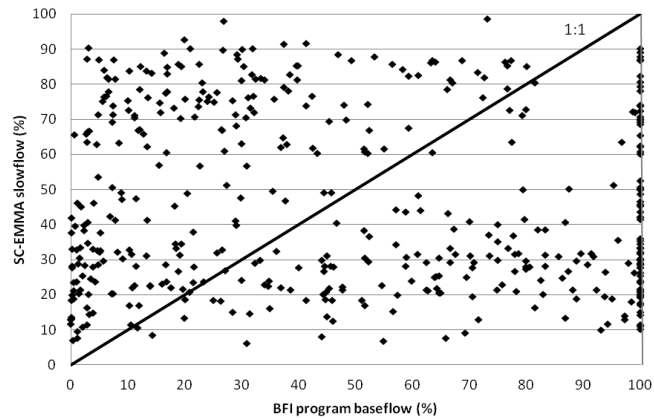
(C) South Fork Iowa River (SFIR), IA
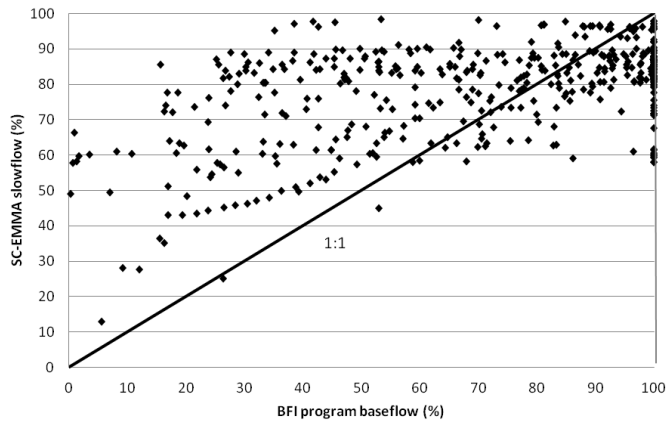
(D) Granger Drain, WA
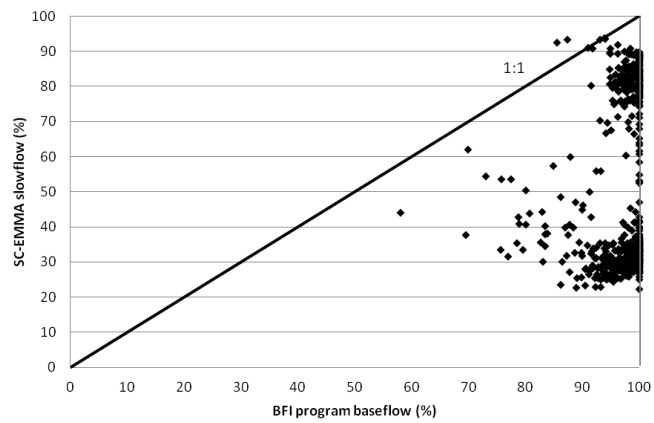
Figure 2.5: Daily SC-EMMA slowflow (%) compared to the BFI program baseflow (%) for a 2-year period for four agricultural streams. 1:1 line represents identical values calculated by SC-EMMA and BFI program. Input values (SC_{FF} and SC_{SF}) for SC-EMMA can be seen in Table 2.1.

Chapter 3: Estimation of time-variable fast flowpath end-member concentrations for application in chemical hydrograph separation analyses

Abstract

End-member mixing analysis (EMMA) is a commonly used method for hydrograph separation, but can be hindered by the subjective choice of end-member chemical concentrations. This work tests a new method of EMMA which relies on high frequency (continuous) measures of two chemicals and streamflow to separate total streamflow into water from slowflow and fastflow sources. The ratio between the concentrations of the two chemicals is used to create a time-variable estimate of the concentration of each chemical in the fastflow end-member. Multiple synthetic data sets and data from two hydrologically diverse streams were used to test the performance and limitations of the new model (ratio-EMMA). When applied to the synthetic streams under many different scenarios, the ratio-EMMA produced results that were reasonable approximations of the actual values of fastflow discharge ($\pm$0.20 m^3/s 3.4%) and fastflow concentration ($\pm$0.19 mg/L (24%) and $\pm$25 µS/cm (41%) for the two chemical respectively). With real stream data, the ratio-EMMA produced continuous estimates of slowflow and fastflow discharge that aligned with expectations for each stream based on their respective hydrological settings. The use of two chemicals with the ratio-EMMA provides an innovative and objective approach for estimating continuous fastflow concentration and contributions of fastflow water to the stream. This provides useful information for continuous tracking of chemical movement to streams, allowing for better selection and implementation of water quality management strategies.

1. Introduction

Waterborne chemicals and nutrients travel from the landscape to streams through many different hydrologic pathways (flowpaths). Excess nutrients transported from the landscape can lead to eutrophication and hypoxia when the nutrients enter the stream (Poor and McDonnell, 2007, David *et al*., 2010, Goodridge and Melack, 2012). Streamflow is a combination of water from multiple flowpaths, and changes in

contributions from the various flowpaths has been shown to dramatically alter nutrient concentration within a stream (Hooper et al., 1990, Molénat et al., 2002, Peters, 1994, Rice and Bricker, 1995, Ross et al., 1994, Goodridge and Melack, 2012, Baron *et al*., 2013) resulting in fluctuations in the chemical load delivered to the stream under various hydrologic conditions. This makes it difficult for watershed managers to estimate and mitigate inputs from chemical loadings.

The time required for water and chemicals to travel to the stream depends on the flowpath through which they move, affecting the degree of interaction between the water and the soil. During this water-soil contact time, chemicals dissolve and their concentrations increase in the water. Fastflow water is water that is transported rapidly to the stream. It is a combination of water from many sources including direct precipitation, overland flow, shallow subsurface flow, and subsurface drain outflow after irrigation or recent rain events. These flowpaths result in minimal soil/water interaction and usually have water with low specific conductance and low concentrations of soluble chemicals except in areas where soluble chemicals are applied to the land surface (Woodward et al., 2013, Raymond *et al*., 2012). Slowflow water, on the other hand, is composed of water that travels slowly to the stream through other flowpaths (subsurface drain flow distant from rain events or groundwater discharge) which results in greater soil/water interactions and often contains water with higher specific conductance and chemical concentrations (Woodward et al., 2013).

Hydrograph separation techniques can be used to estimate the contributions of water to the stream from two or more flowpaths. There are several methods of hydrograph separation. Graphical methods utilize a streamflow record to define the points where slowflow intersects the rising and falling limbs of the hydrograph, effectively separating the relatively stable slowflow from short-term elevated flows resulting from direct runoff (Sloto and Crouse, 1996, Wahl and Wahl, 1995). Recursive digital filters also utilize a record of streamflow, mathematically removing the high-amplitude runoff signal in the hydrograph, resulting in an estimate of the low-amplitude slowflow (Eckhardt, 2005, Rimmer and Hartmann, 2014). End-member mixing analysis (EMMA) incorporates additional information about the concentration of a chemical(s) to

mathematically separate the mixture of waters in the stream into their individual flowpaths (Hooper, 2013). Each method is not without its limitations. Results from graphical methods are often not reproducible from one person to the next (Sloto and Crouse, 1996), as some of the methods rely on subjective decisions. Some recursive digital filters require extensive calibration and produce results that have little hydrological basis (Schwartz, 2007). And, results from EMMA are sensitive to the end-member input parameters (Kronholm and Capel, 2014).

EMMA is often an attractive choice for hydrograph separation because it produces results that are based on known physical, chemical, and hydrological processes. An EMMA relies on measurements of streamflow and the concentration of a conservative (non-reactive and non-sorptive) chemical in the stream. Assumptions about the concentrations of the chemical in the slow flowpath and fast flowpath end-members are needed for the separation of total streamflow into the fraction of water from each end-member. Slowflow end-member concentration is often assumed to be equal to the measured concentration in the stream during lowflow conditions (Miller *et al*. 2014, Sanford *et al*. 2012), when most or all of the water in the stream is from slowflow sources. The slowflow end-member concentration is often assumed to be constant over short time periods because it responds slowly to changes within the watershed (Miller *et al*. 2014). One of the more difficult and debated aspects of an EMMA is the choice of an appropriate fastflow end-member concentration. A measured value of a chemical in the stream during the highest flows was used by Miller *et al*. (2014) as an estimate of the fastflow concentration. However, using the stream concentration during the highest flows may not accurately approximate the concentration of the chemical in fastflow, as there can still be significant inputs of slowflow water to the stream at the highest flows (Stewart *et al*., 2007), resulting in an elevated approximation of fastflow concentration. McNamara *et al*., (1997) utilized the concentrations of the chemical in overland flow water, but those were point samples which can vary greatly through time and space. Pellerin *et al.* (2008) and Sanford *et al*. (2012) utilized the concentration of a chemical in rainfall as the value for fastflow concentration, as rainfall is frequently a driver of fastflow water. However, due to the rapid interaction between rainwater and soil, and the

36

subsequent increase in chemical concentration in the water (Pilgrim *et al*., 1979), rainfall concentration underestimates the concentration of the chemical in fastflow. Kronholm and Capel, (2014) approximated the fastflow concentration as a value halfway between the concentration of a chemical in rainfall and the lowest measured concentration in the stream. These approaches are all valid and are based on easily obtained measurements, but rely on the rarely true assumption of a static value for fastflow concentration (Smith, 2012). By using a static fastflow concentration, the results can underestimate fastflow discharge at the beginning of a storm event (when the estimated fastflow concentration is too high) and overestimate fastflow discharge later in the storm (when the estimated fastflow concentration is too low) or vice versa.

In response to the difficulty of selecting an appropriate fastflow concentration for use with an EMMA model, a ratio-based EMMA (ratio-EMMA) has been created which utilizes high-frequency measurements of streamflow and the concentrations of two chemicals in the stream rather than one, as in the standard EMMA for hydrograph separation. By using the ratio of the concentration of one chemical to the other, a time-variable estimate of the concentrations of each chemical in fastflow is generated, which can then be used as part of an EMMA model to estimate a continuous record of slowflow and fastflow water entering the stream. Synthetic data were used to test the sensitivity of the ratio-EMMA model to various conditions of streamflow and chemical concentration. The sensitivity analysis provides understanding of the relationship between the inputs and outputs of the model while providing insights to its practicality and limitations. Data from two real streams were also used to test the model under real world conditions. Although concentration data for any two conservative chemicals can be used, specific conductance and nitrate were used as the two "chemicals" for this application of the ratio-EMMA. Nitrate is used in this work as an informative example, even though it is not ideal for the ratio-EMMA due to its reactive nature. As the number and variety of chemical sensors increases, the ratio-EMMA will be available to take advantage of the high frequency data to produce accurate continuous estimates of fastflow and slowflow water added to a stream. The ratio-EMMA eliminates a major assumption of an EMMA by not relying on a static input value for the fastflow end-member and allows for the

continuous separation of total streamflow and the continuous estimation of chemical concentrations in fastflow water without the use of subjectively defined fastflow end member concentration.

2. Methods

2.1 Standard EMMA

The total volume of water in a stream (S) is a combination of inputs of water from various flowpaths. The concentration of a chemical ([A]) is used in the creation of a standard end-member mixing analysis (EMMA) to separate streamflow into two end-members: water from fast flowpaths (FF) and water from slow flowpaths (SF). This separation is represented by the following equations:

$$Q_S = Q_{SF} + Q_{FF} \qquad \text{Eq. 1}$$

and, therefore:

$$Q_S[A]_S = Q_{SF}[A]_{SF} + Q_{FF}[A]_{FF} \qquad \text{Eq. 2}$$

where Q is the water discharge (m^3/s). Equation 1 represents a water budget and equation 2 is the corresponding chemical mass budget. By combining equations 1 and 2, the slowflow contributions to overall streamflow are calculated using equation 3.

$$Q_{SF} = (Q_S[A]_S - Q_S[A]_{FF}) / ([A]_{SF} - [A]_{FF}) \qquad \text{Eq. 3}$$

Q_{FF} is then calculated using equation 1.

For a given length of time (t), the volume (V) is

$$V = Q * t. \qquad \text{Eq. 4}$$

The slowflow index (SFI, slowflow volume as a percent of total stream volume) for a given length of t, is

$$SFI = (\Sigma V_{SF} / \Sigma V_S) * 100 \qquad \text{Eq. 5}$$

When using Equation 3 for hydrograph separation, Q_S and $[A]_S$ are measured values, whereas $[A]_{SF}$ and $[A]_{FF}$ are subjective estimates of the $[A]$ in each end-member. $[A]_{SF}$ is approximated as the measured concentration in the stream during lowflow conditions, however $[A]_{FF}$ is much more subjective and more variable through time.

2.2   Ratio-based EMMA

The ratio-based EMMA approaches hydrograph separation in the same manner as above. However, two simultaneous chemical mass budgets are used instead of one. Subscripts A and B represent values relating to chemicals A and B, respectively.

$$Q_{S,A} = Q_{SF,A} + Q_{FF,A} \qquad \text{Eq. 1,}_A$$

$$Q_{S,B} = Q_{SF,B} + Q_{FF,B} \qquad \text{Eq. 1,}_B$$

and, therefore:

$$Q_{S,A}\,[A]_S = Q_{SF,A}\,[A]_{SF} + Q_{FF,A}\,[A]_{FF} \qquad \text{Eq. 2,}_A$$

$$Q_{S,B}\,[B]_S = Q_{SF,B}\,[B]_{SF} + Q_{FF,B}\,[B]_{FF} \qquad \text{Eq. 2,}_B$$

and the assumptions that:

$$Q_{S,A} = Q_{S,B} \qquad \text{Eq. 6}$$

$$Q_{SF,A} = Q_{SF,B} \qquad \text{Eq. 7}$$

$$Q_{FF,A} = Q_{FF,B} \qquad \text{Eq. 8}$$

Each set of chemical data could be used on its own to accomplish separate EMMA analyses, but each would rely on static and subjective fastflow input concentrations. By combining both strings of unique chemical data into a single EMMA, the ratio between the concentrations of the chemicals in fastflow is used to more objectively define the fastflow end-member concentrations. This is accomplished by making the assumptions that both chemicals (A and B) are conservative (non-reactive and non-sorptive), that $[A]_{SF}$ and $[B]_{SF}$ are constant, and that A and B in fastflow travel in the same parcel of water to the stream. A finite range of ratios which encompasses the actual ratio of $[A]_{FF}$ to $[B]_{FF}$ is used to mathematically estimate time variable concentrations of the two

chemicals in the fastflow. In most applications of the ratio-EMMA, an iterative approach to ratio selection is recommended and will be discussed in detail. The iterative approach can be applied without prior knowledge of the fastflow conditions within a watershed and allows for application of the ratio-EMMA in a larger number of streams. The iterative approach removes the uncertainty of selecting a fixed upper and lower bound to the range of ratios, and is used exclusively in the remainder of the study. However, the range of possible ratios of $[A]_{FF}$ to $[B]_{FF}$ can be specified by the user based on knowledge of $[A]_{FF}$ and $[B]_{FF}$. The full sensitivity analysis of the ratio-EMMA results to a user defined fastflow ratio is found in Appendix C.

The time-varying $[A]_{FF}$ and $[B]_{FF}$ are calculated for each time interval (2-hours, in this case) according to the procedure presented in Figure 3.1. The data are entered into a macro enabled Microsoft Excel spreadsheet with a Visual Basic code to run the logical algorithm of the model (Visual Basic code presented in Appendix D). The estimated range of ratios of $[A]_{FF}$ to $[B]_{FF}$ is defined by the user based on the iterative approach or knowledge of the concentrations of $[A]_{FF}$ and $[B]_{FF}$. These are entered into the spreadsheet, and the model is run.

During each 2-hour time interval, the model enters the measured stream data (Q_S and $[A]_S$) and assumed chemical data ($[A]_{SF}$) from a single time interval (Figure 3.1-A), into Equation 3. Nitrate (mg N/L) is used as the example chemical A for the real stream examples.

Using a standard EMMA equation (Eq. 3), the model calculates 1000 individual estimates of Q_{SF} by inputting the measured values for Q_S, $[A]_S$, and $[A]_{SF}$, and 1000 possible values for $[A]_{FF}$. This produces 1000 different estimates for slowflow based on chemical A ($Q_{SF,A}$) (Figure 3.1-B). Possible values for $[A]_{FF}$ were chosen to range from 0.00 to 10 mg N/L in increasing increments of 0.01 mg N/L. The range of possible values of $[A]_{FF}$ can be altered in cases where the expected $[A]_{FF}$ is >10 mg N/L, and fewer (or more) than 1000 estimates can be calculated if desired. Although unrealistic, negative $Q_{SF,A}$ estimates, and $Q_{SF,A}$ estimates that are $>Q_S$ are possible as the model balances Equation 3 based on the input values. In the event that the estimated $Q_{SF,A} <0.0$ m^3/s or $>Q_S$, the ratio-EMMA sets $Q_{SF,A}$ equal to 0.0 m^3/s.

The model then enters the measured stream data (Q_S and $[B]_S$) and assumed, constant $[B]_{SF}$ for the same 2-hour time interval (Figure 3.1-C) into Equation 3. Specific conductance (in units of μS/cm) is used as the example chemical B for the real stream examples.

Using a standard EMMA equation (Eq. 3), the model calculates 1000 individual estimates of Q_{SF} by inputting the measured values for Q_S, $[B]_S$, and $[B]_{SF}$, and 1000 possible values for $[B]_{FF}$ into Equation 3. This produces 1000 different estimates for slowflow based on the EMMA calculations from Equation 3 using chemical B ($Q_{SF,B}$) (Figure 3.1-D). Possible values for $[B]_{FF}$ were chosen to range from 0.0 to 1000 μS/cm in increasing increments of 1.0 μS/cm. The range of possible values of $[B]_{FF}$ can be altered in cases where the expected $[B]_{FF}$ is >1000 μS/cm. In the event that the estimated $Q_{SF,B}$ <0.0 m^3/s or >Q_S, the ratio-EMMA sets $Q_{SF,B}$ equal to 0.0 m^3/s.

The model identifies the $Q_{SF,A}$ produced by the first $[A]_{FF}$ (Figure 3.1-E) and then searches for an equal value within the 1000 estimates of $Q_{SF,B}$ (Figure 3.1-F).

In Figure 3.1-G, the model searches for equivalent values of $Q_{SF,A}$ and $Q_{SF,B}$. If $Q_{SF,B}$ produces no match to the value of $Q_{SF,A}$ held in step E, the model reverts to step E and uses the next $[A]_{FF}$. If an equivalent value to $Q_{SF,A}$ is found, the program then refers back to the $[A]_{FF}$ used to produce $Q_{SF,A}$ and the $[B]_{FF}$ used to produce the $Q_{SF,B}$.

The model compares the ratio of $[A]_{FF}$ to $[B]_{FF}$ against a pre-defined range of ratios (Figure 3.1-H). Approaches to select the range of ratios are discussed below. If the ratio of $[A]_{FF}$ to $[B]_{FF}$ falls outside of the defined range, the model reverts to step E and uses the next $[A]_{FF}$ value (Figure 3.1-I). If the ratio of $[A]_{FF}$ to $[B]_{FF}$ falls within the range of ratios, these concentration values are set as possible values of $[A]_{FF}$ and $[B]_{FF}$ for the time interval. The model reverts to step A and begins calculating values for the next 2-hour time interval. If no solutions meet both criteria, the values for $[A]_{FF}$ and $[B]_{FF}$ are left blank and no solution is reached for the time interval. The model reverts to step A and begins calculating values for the next 2-hour time interval. At times, the estimated $[A]_{FF}$ can equal 0.0 mg/L, but $[B]_{FF}$ cannot be equal to zero due to mathematical constraints of the model.

There is the possibility of multiple solutions within each 2-hour time interval as a result of the range of concentration ratios necessary for the model to operate. This is especially (but not exclusively) true when not using the iterative method due to the likelihood of a wider range of ratios. Because of this, the model is then run a second time in reverse at each time interval. During the second run, the values of $[A]_{FF}$ and $[B]_{FF}$ used to create the $Q_{SF,A}$ and $Q_{SF,B}$ are reversed (10 to 0.00 mg N/L decreasing by increments of 0.01 mg N/L and 1000 to 0.0 μS/cm decreasing by increments of 1.0 μS/cm). The mean of the estimated $[A]_{FF}$ by running the model forward and the $[A]_{FF}$ by running the model in reverse is taken as the "true" $[A]_{FF}$ for the time interval. The same is done for chemical B. For both chemicals, there is a linear change in concentration between all of the possible fastflow solutions for a single 2-hour time interval. Since all possible values fall along a straight line and are bracketed by the first (forward) and last (reverse) solutions from the model, the mean is used as the measure of the central tendency (Figure 3.2). When using the iterative approach, the concentrations of each chemical from running the model forward and in reverse are only averaged after the last iteration. By running the model twice during each time interval, this model also quantifies variability in the estimated fastflow concentrations based on the range of selected ratios.

An iterative approach to ratio selection is the preferred method to objectively select the range of ratios necessary for the ratio-EMMA model, especially when little is known about $[A]_{FF}$ and $[B]_{FF}$. With the iterative method, the defined range of ratios is initially set from 0 to 1000 and the model is run (forward and in reverse). With possible ratios that are essentially unbounded on the upper and lower ends, the model finds the highest and lowest concentrations of $[A]_{FF}$ and $[B]_{FF}$ and ratios of $[A]_{FF}$ to $[B]_{FF}$ that can produce values during each 2 hour time interval. The resulting ratios for each 2 hour interval are used to define the range of ratios for the next iteration. The range is then narrowed (the lowest ratio is increased 10% and the highest ratio is decreased 10%), and the resulting values are input back into model as the defined range of ratios for the second iteration. As the selected range of ratios gets smaller, the likelihood that the model will not be able to find a solution that meets all necessary criteria gets larger. In the event that

the generated range of ratios is not able to produce results during a given iteration, or the range of ratios becomes narrower than 0.0020 (an arbitrary value used in this study), the previous range of ratios is selected for that time period. This process is repeated until the change in SFI for the entire period of record is less than 0.1% (an arbitrary value). The range of ratios (and fastflow concentrations) can likely be further narrowed during some time intervals, but the change in SFI will be minimal. This iterative approach produces a time-varying range of ratios which allows the ratio-EMMA to be applied in many hydrologically diverse streams.

Once the model has been run, $[A]_{FF}$, along with Q_S, $[A]_S$, and $[A]_{SF}$, are input into Equation 3 for the first 2-hour time interval (Figure 3.1-J), resulting in the estimated Q_{SF} for that time interval. The Q_S and estimated Q_{SF} are entered into Equation 1 to determine Q_{FF}. This process is then repeated for each time interval over the entire period of modeling, creating continuous estimates of Q_{SF} and Q_{FF} based on the time-variable $[A]_{FF}$. Using $[A]_{FF}$ or $[B]_{FF}$ along with the corresponding Q_S, $[A]_S$, and $[A]_{SF}$ or Q_S, $[B]_S$, and $[B]_{SF}$ will result in the same estimates of Q_{SF} and Q_{FF}, so Q_{SF} and Q_{FF} do not need to be calculated with both $[A]_{FF}$ and $[B]_{FF}$ separately.

2.3   Data for Model Characterization

To test the validity of this new model, synthesized and real stream data were utilized. Synthetic data were designed to test the precision and bias of the model against actual values. The model was also applied to real data from two hydrologically different streams. The analysis of the synthetic data provided information regarding the performance of the model, which aided in the interpretation of the results from the real data.

The measurements of streamflow and stream chemistry in the real stream sites were aggregated to 2-hour intervals for modeling purposes. The synthetic data were made at the same temporal resolution as the real measured data.

For this initial test, the model was run for a single storm event, or for a short time period containing multiple storm events (rather than an annual period) for the both synthetic data and real data. Certain assumptions are minimized when handling shorter data records that would not be possible to ignore when examining an annual or longer

time scale. Seasonal changes in the physical environment can be ignored, and the concentration of each chemical in slowflow water can be assumed to be relatively constant.

2.3.1   *Synthetic data creation*

Two separate synthetic watersheds were created in an effort to test the model under different hydrologic conditions. One synthetic stream was created to mimic a stream that has large inputs of groundwater to the stream (large slowflow inputs). The other synthetic stream is meant to mimic a stream that is dominated by overland flow additions of water (small slowflow inputs). Each data set was created to have balances of chemical mass and water volume. Synthetic data sets include values, at each 2-hour time interval, for total streamflow (m^3/s), fastflow (m^3/s), slowflow (m^3/s), and $[A]_S$, $[B]_S$, $[A]_{SF}$, $[B]_{SF}$, $[A]_{FF}$, and $[B]_{FF}$. Chemicals A and B are meant to represent nitrate (mg N/L) and specific conductance (μS/cm), respectively. The concentrations of both chemicals were held constant in slowflow water, whereas chemical concentrations in fastflow were varied through time.

Seven scenarios were created for each of the two synthetic streams to assess the sensitivity of the ratio-EMMA to the temporal alignment of the streamflow peak and the minimum (or maximum) $[A]_{FF}$ and $[B]_{FF}$. Scenarios include situations where (1) the streamflow peak occurs at the same time as the minimum $[A]_{FF}$ and $[B]_{FF}$ ($t_{Qmax}=t_{[A]min}=t_{[B]min}$),(2) streamflow peak occurs but the minimum $[A]_{FF}$ and $[B]_{FF}$ both occur before streamflow peak ($t_{Qmax}>t_{[A]min}=t_{[B]min}$), (3) streamflow peak occurs but the minimum $[A]_{FF}$ and $[B]_{FF}$ both occur after streamflow peak ($t_{Qmax}<t_{[A]min}=t_{[B]min}$), (4) streamflow peak and the lowest $[A]_{FF}$ occur at the same time but the lowest $[B]_{FF}$ occurs before the streamflow peak ($t_{Qmax}=t_{[A]min}>t_{[B]min}$), (5) streamflow peak and the lowest $[A]_{FF}$ occur at the same time but the lowest $[B]_{FF}$ occurs after the streamflow peak ($t_{Qmax}=t_{[A]min}<t_{[B]min}$), and (6) the streamflow peak occurs but the minimum $[A]_{FF}$ occurs before streamflow peak and the minimum $[B]_{FF}$ occurs after the streamflow peak ($t_{[A]min}<t_{Qmin}<t_{[B]min}$). Concentration minima in fastflow occurring before or after the streamflow peak were offset 12 time intervals (the equivalent of 24 hours when using 2-

hour data) from the streamflow peak. In the 7th scenario $[A]_{FF}$ increased with flow while at the same time $[B]_{FF}$ decreased with flow ($t_{Qmax}=t_{[A]max}=t_{[B]min}$).

2.3.2 Site descriptions

The ratio-EMMA was applied to two real streams to test the model in real world situations. Both streams have a gage to measure streamflow and are equipped with continuous (high-frequency) chemical sensors for measuring specific conductance and nitrate concentrations.

Chesterville Branch (15.9 km^2) is located in Maryland (USGS Station identification (STAID) 01493112). Agriculture comprises >90% of the land within the watershed, most of which is non-irrigated (Senus et al., 2005). Streamflow is supported by natural groundwater throughout the year with additions of overland flow during rainfall events being common (Ator et al., 2005). Mean annual precipitation is 112 cm. Annual streamflow in 2013 for within Chesterville Branch averaged 0.27 m^3/s (USGS, 2013b). Groundwater in the Chesterville Branch watershed is characterized by high nitrate concentration which is observed during times of lowflow (<0.30 m^3/s) in the stream (typically 4.5 to 8.5 mg N/L). Streamflow was measured using WaterLog H-3611 Radar Water Level Sensor (Design Analysis Associates, Inc., Logan UT). Nitrate in the stream was measured continuously with a SUNA V2 with a 5 millimeter optical window and an integrated wiper (Satlantic LP, Halifax, Nova Scotia, Canada). SC was measured with a YSI 6920 V2 (YSI Inc., Yellow Springs, Ohio). Streamflow, SC, and nitrate from June 10, 2013 to June 24, 2013 were used in this study. Linear interpolation was used to fill in any small gaps in the data record. Values for [nitrate]$_{SF}$ and [SC]$_{SF}$ used with all EMMA models were assumed to be constant over the study period and were set at 8.64 mg N/L and 203 μS/cm respectively, which represent the highest measured values in the stream.

Indian Creek (166 km^2) is located in eastern Kansas (STAID 06893390). Urbanized land comprises >75% of the area within the watershed with nearly 20% of the watershed having impervious land cover. Greater than 95% of the water in the stream during lowflow conditions comes from waste water treatment plant discharge (Lee et al., 2005), which will be considered slowflow by the model. Mean annual precipitation

within the watershed is 99 cm (Rasmussen et al., 2008). Annual average streamflow in 2013 for Indian Creek was 2.6 m^3/s (USGS, 2013c). Streamflow was measured using a Sutron Accubar Bubbler 5600-0131 (Sutron Corporation, Sterling, Virginia). Nitrate and SC measurements in the stream were measured continuously with a Nitratax plus SC sensor using a 5mm path length (HACH Inc., Loveland, Colorado) and a YSI 6560 (YSI Inc., Yellow Springs, Ohio) respectively. Streamflow, SC, and nitrate from April 6, 2013 to April 20, 2013 were used in this study. Values for [nitrate]$_{SF}$ and [SC]$_{SF}$ used with all EMMA models were assumed to be constant over the study period and were set at 8.36 mg N/L and 1480 μS/cm respectively, which represent the highest measured values in the stream.

3. Results

3.1   Use of the ratio-EMMA in real streams

The ratio-EMMA model was used to estimate slowflow (Q_{SF}), fastflow (Q_{FF}), [nitrate]$_{FF}$, and [SC]$_{FF}$ for two-week periods in the two real streams. The ratio-EMMA was compared to the standard EMMA models, one of which made use of specific conductance (SC-EMMA) and the other which used nitrate (nitrate-EMMA). The iterative approach to ratio selection was used with the ratio-EMMA for both streams.

The ratio-EMMA estimated large inputs of fastflow (~90% fastflow) to Chesterville Branch during storm events and smaller additions of fastflow (~20% fastflow) during the non-storm events, which is expected since it is a groundwater dominated stream (Figure 3.3A). The results from the ratio-EMMA were similar, but not identical to the results from the SC-EMMA and nitrate-EMMA models (Figure 3.3A). During 39% of the 14 day period of record, the ratio-EMMA produced a fastflow hydrograph that was between the SC-EMMA and nitrate-EMMA fastflow hydrographs (as seen during 6/23/2013 on Figure 3.3A). The ratio-EMMA produced fastflow discharge estimates that were nearly equivalent ($\pm$0.1 m^3/s) to the nitrate-EMMA during virtually all of the 14-day period of modeling, deviating mostly during periods of lowflow in the stream. The times when the ratio-EMMA hydrograph fell outside of the other two EMMA hydrographs (Figure 3.3A between 6-10-2013 and 6-13-2013)

46

demonstrate that the ratio-EMMA is not limited in its range relative to the results that the individual nitrate-EMMA and SC-EMMA yield. Figure 3.3B shows that the ratio of $[nitrate]_S$ to $[SC]_S$ decreased as flow increased, which indicates that water from fastflow has a smaller ratio of [nitrate] to [SC]. The smaller ratio may be a result of little available nitrate at the soil surface or less/slower uptake of nitrate than of other dissolvable ions which contribute to elevated $[SC]_{FF}$.

The ratio-EMMA estimated large inputs of fastflow (~90% fastflow) to Indian Creek during storm events, but larger fastflow inputs to the stream (~50% fastflow) during non-storm events than Chesterville Branch, as would be expected in an overland flow dominated stream (Figure 3.3C). The ratio-EMMA fastflow hydrograph fell between those produced by the individual SC-EMMA and nitrate-EMMA models only 12% of the time during the 14 day period of record (Figure 3.3C). The ratio-EMMA fastflow hydrograph was similar (± 0.1 m^3/s) to the nitrate-EMMA hydrograph during a majority of the 14 days, but deviated periodically during both lowflow and highflow events in the stream.

Continuous fastflow estimated by the ratio-EMMA was different than that of the nitrate-EMMA and the SC-EMMA models (Figures 3.3A and C), but the results were more closely related to the nitrate-EMMA in both real streams. The close relationship between the results from the ratio-EMMA and the nitrate-EMMA may be explained by the greater difference between the minimum and maximum normalized concentration (normalized to the highest concentration) of nitrate in the stream relative to the difference for specific conductance. And, the relationship between flow and nitrate concentration was more tightly and linearly monotonic than for specific conductance. A comparison of a plot of the ratio of $[nitrate]_S$ to $[SC]_S$ (Figure 3.3B and D) and a plot of streamflow (Figure 3.3A and C) provides insight into the ratio of the chemicals in fastflow. In Chesterville Branch and Indian Creek, as flow increased, the ratio of $[nitrate]_S$ to $[SC]_S$ decreased, suggesting that the ratio of $[nitrate]_{FF}$ to $[SC]_{FF}$ is less than the ratio measured in the stream. The ratio of $[nitrate]_{FF}$ to $[SC]_{FF}$ produced by the ratio-EMMA was less than the ratio of $[nitrate]_S$ to $[SC]_S$ in both streams, which fits expectations.

3.2 **Sensitivity analysis of the results to the streamflow peak and max/min concentrations in fastflow**

The temporal variability of flow and chemical concentration in the stream and in fastflow pose a challenge to the successful application of the ratio-EMMA model. To test the limitations of the model, seven different scenarios of the timing of the streamflow peak to the minimum $[A]_{FF}$ and $[B]_{FF}$ were tested in both of the synthetic streams. Chemicals A and B are assumed to travel to the stream in the same parcels of slowflow and fastflow water.

3.2.1 Scenarios in which the timing of the minimum $[A]_{FF}$ occurred concurrently with that of $[B]_{FF}$

For both synthetic streams, when the lowest measured $[A]_{FF}$ and $[B]_{FF}$ occurred simultaneously, whether that be before ($t_{Qmax}>t_{[A]min}=t_{[B]min}$), during ($t_{Qmax}=t_{[A]min}=t_{[B]min}$), or after peak streamflow ($t_{Qmax}<t_{[A]min}=t_{[B]min}$), the ratio-EMMA model estimates of fastflow (Q_{FF}) were nearly identical to the actual results. Bi-hourly fastflow estimates differed by a median absolute deviation of ±0.02 m^3/s (1.6%) in the groundwater dominated synthetic stream and ±0.40 m^3/s (5.0%) in the overland flow dominated synthetic stream with the largest deviations occurring during peakflow in both streams (Figures 3.4A, 3.5A, and Appendix C). The difference between actual and estimated Q_{FF} during lowflow conditions in the stream was nearly always <0.05 m^3/s (4.4%) in both synthetic streams. Total fastflow volume had an average percent error of $\pm0.6\%$ in the groundwater dominated synthetic stream and $\pm0.7\%$ in the overland flow dominated stream among the three scenarios (Table 3.1). Whether $[A]_{FF}$ and $[B]_{FF}$ reached their lowest values before, during, or after the streamflow peak, deviation from the actual SFI averaged 0.2% in the groundwater dominated stream and 0.9% in the overland flow dominated stream (Table 3.1).

The fastflow concentration estimates produced by the ratio-EMMA were nearly identical to the actual concentration values of the synthetic data (Figures 3.4A, 3.5A, and Appendix C). The bi-hourly estimates had a median absolute deviation of ±0.10 mg/L (12%) for chemical A and ±13 μS/cm (20%) for chemical B in the groundwater dominated synthetic stream; and ±0.27 mg/L (33%) for chemical A and ±34 μS/cm

48

(57%) for chemical B in the overland flow dominated synthetic stream over the entire 14 day period of modeling for the three scenarios. Estimated $[A]_{FF}$ and $[B]_{FF}$ were nearly identical to the actual values during times of highflow and deviated to a greater extent during lowflow (Figures 3.4A, 3.5A, and Appendix C). Over the 14 day period, the overestimates and underestimates of fastflow concentrations nearly negated each other, particularly in the groundwater dominated stream where the median deviations were - 0.04 mg/L and -5.5 µS/cm for chemicals A and B, respectively. A more detailed explanation of this and subsequent scenario testing results is found in Appendix C.

Chemical A loads in fastflow were also estimated for the entire length of record. The percent error for load estimates had a mean of $\pm$12% among the scenarios in the groundwater dominated synthetic stream and $\pm$26% in the overland flow dominated synthetic stream for the three scenarios (Table 3.1). Chemical B loads (calculated by converting SC (µS/cm) to total dissolved solids (TDS in mg/L); TDS ~ 0.64*SC by combining the Russell Equation and Langlier approximation) in fastflow were also estimated for the entire length of record. The percent error for load estimates had a mean of $\pm$19% among the scenarios in the groundwater dominated synthetic stream and $\pm$41% in the overland flow dominated synthetic stream for the three scenarios (Table 3.1).

3.2.2 *Scenarios in which the timing of the minimum $[A]_{FF}$ diverged from that of $[B]_{FF}$*

When the minimum $[A]_{FF}$ and $[B]_{FF}$ were offset in time ($t_{Qmax}=t_{[A]min}>t_{[B]min}$, $t_{Qmax}=t_{[A]min}<t_{[B]min}$, or $t_{[A]min}<t_{Qmax}<t_{[B]min}$), the fastflow volume, SFI, concentration, and chemical load estimates produced by the ratio-EMMA deviated from the actual values to a similar extent as the scenarios where the concentration minima occurred at the same time, particularly for the groundwater dominated synthetic stream (Figures 3.4B, 3.5B, and Appendix C). A few outliers in the estimated fastflow concentrations (an average of five 2-hour time intervals in each scenario) had a large effect on the estimated flows and loads in the groundwater dominated synthetic stream. These outliers were a result of the iterative ratio selection producing possible concentrations of both $[A]_{FF}$ and $[B]_{FF}$ that were highly varied (from running the method forward and in reverse), but the resulting range of ratios of $[A]_{FF}$ to $[B]_{FF}$ were very small. Because the iterative approach was not

able to narrow the range of ratios any further, the concentration estimates were averaged, which led to elevated concentration estimates relative to the surrounding time intervals. For the analysis of these scenarios, the outliers were removed and linear interpolation was used to replace the values.

The estimates of fastflow discharge (Q_{FF}) were accurately predicted. Fastflow estimates differed by a median absolute deviation of ±0.04 m^3/s (3.0%) in the groundwater dominated synthetic stream and ±0.33 m^3/s (3.9%) in the overland flow dominated synthetic stream with the largest deviations occurring during peakflow in both streams (Figures 3.4B, 3.5B, and Appendix C). The difference between actual and estimated Q_{FF} during lowflow in the stream was nearly always <0.05 m^3/s (6.8%) in the groundwater dominated synthetic stream and <0.5 m^3/s (7.6%) in the overland flow dominated synthetic stream. Total fastflow volume deviated from actual results by an average percent error of $\pm0.8\%$ in the groundwater dominated stream and $\pm0.8\%$ in the overland flow dominated stream among the three scenarios (Table 3.1). In these scenarios where $[A]_{FF}$ reached a minimum at the same time as the streamflow peak, but $[B]_{FF}$ reached a minimum either before or after the streamflow peak, the ratio-EMMA SFI deviated from the actual results by $\pm1.5\%$ in the groundwater dominated synthetic stream and by $\pm0.7\%$ in the overland flow dominated synthetic stream (Table 3.1).

The bi-hourly concentration estimates over the 14-day period had a median absolute deviation of ±0.18 mg/L (23%) for chemical A and ±22 μS/cm (38%) for chemical B in the groundwater dominated synthetic stream; and ±0.23 mg/L (28%) for chemical A and ±29 μS/cm (48%) for chemical B in the overland flow dominated synthetic stream. In both synthetic streams, when $[B]_{min}$ occurred before $[A]_{min}$ and Q_{max}, the greatest deviation occurred during lowflow and during the recession limb of the hydrograph (Figures 3.4B and 3.5B). When $[B]_{min}$ occurred after $[A]_{min}$ and Q_{max}, the greatest deviation occurred during lowflow and during the rising limb of the hydrograph in both synthetic streams (Appendix C). And, when $[A]_{min}$ occurred before Q_{max} and $[B]_{min}$ occurred after Q_{max}, the largest deviations were observed exclusively during lowflow in the synthetic streams (Appendix C). In spite of the temporal disconnect between the streamflow peak and $[A]_{FF}$ and $[B]_{FF}$ minimums, the estimated fastflow peak

aligned in time with the peak in streamflow in all scenarios in both synthetic streams (Figures 3.4B, 3.5B, and Appendix C).

The percent error for chemical A load estimates in fastflow had a mean of $\pm15\%$ among the scenarios in the groundwater dominated synthetic stream and $\pm20\%$ in the overland flow dominated synthetic stream (Table 3.1). The percent error for chemical B load estimates in fastflow had a mean of $\pm14\%$ among the scenarios in the groundwater dominated synthetic stream and $\pm23\%$ in the overland flow dominated synthetic stream (Table 3.1).

3.2.3 Scenarios in which the timing of the maximum $[A]_{FF}$ and minimum $[B]_{FF}$ occurred concurrently

In the third set of scenarios, $[B]_{FF}$ decreased as flow increased (as the concentration has done in all other scenarios), but $[A]_{FF}$ increased as flow increased ($t_{Qmax}=t_{[A]max}=t_{[B]min}$). This is the case for streams where the concentration of a chemical is lower in slowflow (groundwater) and higher in fastflow (runoff) water (Figure 3.6).

Fastflow (Q_{FF}) estimates differed by a median absolute deviation of ±0.35 m^3/s (42%) in the groundwater dominated synthetic stream and ±0.31 m^3/s (2.4%) in the overland flow dominated synthetic stream with the largest deviations occurring during peakflow in both streams (Figure 3.6A and B). Total fastflow volume for the 14-day period of record had a percent error of $\pm9.0\%$ in the groundwater dominated stream and $\pm1.4\%$ in the overland flow dominated stream (Table 3.1). Despite the difference in the behavior of the chemicals in the stream, the SFI from both synthetic streams differed from the actual values by an average of 2.3% (Table 3.1).

In these scenarios the concentration differences between the ratio-EMMA estimates and the actual values were the largest compared to the other scenarios that were tested. The bi-hourly concentration estimates had a median absolute deviation of ±1.2 mg/L (24%) for chemical A and ±228 µS/cm (290%) for chemical B in the groundwater dominated synthetic stream (Figure 3.6A); and ±0.15 mg/L (2.0%) for chemical A and ±18 µS/cm (71%) for chemical B in the overland flow dominated synthetic stream (Figure 3.6B) over the entire 14 day period. The largest deviations were observed during times of lowflow in the stream.

Chemical A load estimates in fastflow were nearly equivalent to actual values in the groundwater dominated synthetic stream (1.1% error) and in the overland flow dominated synthetic stream (0.2% error) (Table 3.1). Chemical B load estimates in fastflow diverged from the actual values to a greater extent than for chemical A. Load estimates from the ratio-EMMA for chemical B were 637% greater than actual in the groundwater dominated stream, and in the overland flow dominated synthetic stream the estimates for chemical B load were 88% greater than actual (Table 3.1).

4. Discussion

4.1   Accuracy and limitations of the ratio-EMMA model

The synthetic streams were created to represent two hydrologically different streams. Many scenarios of the chemical conditions within those streams were tested and showed that the ratio-EMMA was able to produce accurate estimates of fastflow volume (mean percent error of 0.7%), slowflow index (SFI) (2.1%), chemical concentrations (13% for chemical A and 23% for chemical B), and loads (18% for chemical A and 24% for chemical B) in nearly all of the scenarios. The accuracy of these values is important for understanding a stream and the watershed that is drained by that stream.

High-temporal resolution hydrograph separation provides a detailed estimate of the influence of both slow and fast flowpaths on stream discharge. Because the physical, chemical, and biological processes that occur in slow and fast flowpaths, as well as the residence time and discharge volume from each flowpath are so different, separating and understanding the influence of these distinct flowpaths is key to understanding current stream conditions, as well as future behavior of the stream under natural and/or anthropogenic pressures.

The ratio-EMMA provides hydrograph separation with a high degree of accuracy and at any temporal resolution that the data allow. In the tested scenarios within the synthetic streams, estimates of total fastflow volume were within 1.5% of the actual value in all but one scenario. The accuracy of this, as well as the estimated fastflow volume during each time interval and the time-varying fastflow concentration estimates from the ratio-EMMA will aid in the identification of conditions within the watersheds that may

lead to elevated inputs of water soluble, non-point source pollutants to the stream from slow and fast flowpaths.

SFI defines an important characteristic of the stream and the watershed it drains. Streams with a similar SFI will often behave in a similar manner (Johnes, 2007), and will respond to natural and anthropogenic stimuli similarly (Holman, 2011). Many currently used methods of hydrograph separation result in different estimates of SFI when applied to the same stream. In a one of their study watersheds, Kronholm and Capel (2014) showed that two commonly used methods of hydrograph separation produced SFI estimates that were significantly different; with one method resulting in an SFI of 42% (for the SC-based EMMA method) and the other a SFI estimate of 97% (for the BFI program method), highlighting the importance of a subjective and accurate method such as the ratio-EMMA. The ratio-EMMA objectively produced accurate estimates of SFI in all of the tested scenarios, providing the ability to apply the ratio-EMMA in a wide variety of real world watersheds.

For the most part, the ratio-EMMA also estimated chemical concentrations with a high degree of accuracy. Generally in the synthetic streams, the largest deviations of the estimated $[A]_{FF}$ and $[B]_{FF}$ compared to the actual occurred during lowflow conditions. During lowflow in the stream, when additions of water from fastflow sources are minimal, large fluctuations in $[A]_{FF}$ and/or $[B]_{FF}$ result in only slight changes in the concentrations of the chemicals in the stream. The smaller the fastflow volume, the smaller the effect $[A]_{FF}$ and $[B]_{FF}$ will have on the stream concentration, making it more difficult for the ratio-EMMA to find the actual $[A]_{FF}$ and $[B]_{FF}$ during lowflow conditions. Although this is an issue with the ratio-EMMA, variations in $[A]_{FF}$ and $[B]_{FF}$ during lowflow have less of an effect on the estimated Q_{SF} and Q_{FF} and, therefore, minimally affect the calculated SFI. Estimation of $[A]_{FF}$ and $[B]_{FF}$ during lowflow still pose a challenge to the ratio-EMMA model, but is a minor issue when estimating Q_{SF}, Q_{FF}, or SFI.

The ratio-EMMA produced accurate estimates of total fastflow volume, SFI, and chemical concentrations and loads whether the minimum concentration of each chemical in fastflow occurred before, during, or after the streamflow peak. In the seven different

scenarios tested in both of the synthetic streams, the largest percent error (although often still quite small) was generally produced by the ratio-EMMA when the concentrations of the two chemicals in the stream behaved in an opposite manner ($t_{Qmax}=t_{[A]max}=t_{[B]min}$). The estimates of chemical B load had the largest deviations in the groundwater dominated synthetic stream, while the estimates for the load of chemical A in both synthetic streams were very accurate. It appears that chemical A loads were estimated more accurately in the groundwater dominated synthetic stream due to the greater difference between the minimum and maximum normalized concentrations (normalized against the highest concentration) of chemical A in the stream compared to those of chemical B. The larger normalized difference between end-member chemical concentrations made it easier for the ratio-EMMA to produce accurate values for chemical A.

Because the ratio-EMMA was able to accurately estimate values from the synthetic streams, the results from the synthetic streams provide insight when applying the ratio-EMMA to real streams. However there are still considerations that must be taken before the model is applied to any real stream. A cursory evaluation of the stream data must be completed to determine if the ratio-EMMA is an appropriate model to use. The relationship between streamflow and the concentrations of each chemical in the stream should be monotonic and have a relatively large difference between the minimum and maximum concentrations. If either of these criteria is violated the model will not be able to properly explain the mixture of fastflow and slowflow water in the stream (Christophersen *et al*., 1990). As a result, care must be taken when interpreting the results.

4.2   Overall benefits of ratio-EMMA

The application of a ratio-EMMA can aid in the understanding of water and chemical movement from the landscape to a stream, eliminating the subjectivity of the choice of the fastflow end-member concentration and eliminating the assumption of a static fastflow end-member concentration, both of which are common to standard EMMA's. The model still assumes a constant concentration of the chemicals in slowflow, but over the short time periods tested, this assumption may not have introduced

substantial error. For longer periods of time (e.g., annual), this assumption could introduce error as the slowflow end-member concentration fluctuates. The elimination of this assumption will be a future challenge in the development of the ratio-EMMA method.

The ratio-EMMA also provides estimated fastflow concentrations which are a representation of the aggregation of all fastflow contributions to the stream (direct rainfall, overland flow, etc.). The aggregated fastflow concentrations are not easily obtained by field-based observations since these measurements are always spatially and temporally limited, expensive, time consuming, and uncertain because the data are a collection of point samples which must be volume, spatially, and temporally weighted. The model-produced aggregated concentrations of chemicals in fastflow might provide a better estimate than is generally available from field measurements. Also, the ratio-EMMA time-variable fastflow concentrations provide a more realistic portrayal of the natural system than is provided by the static fastflow concentration of a standard, non-ratio based EMMA.

The ratio-EMMA model is developed for any two chemicals that can be measured with continuous sensors and are minimally sorptive and minimally reactive. Specific conductance fulfills these requirements. In some locations, but not all, nitrate may also fulfill these criteria. Nitrate is used in this work solely as an informative example. As new sensors are developed and deployed for other chemicals that are better suited for an EMMA, the ratio-EMMA will be available to take advantage of the data.

The ratio-EMMA can provide useful, time-varying estimates of slowflow and fastflow as well as the concentrations of selected chemicals in fastflow water. It can also serve as a check for other hydrograph separation models (standard-EMMA, graphical, etc.) for the sites where data are available. This information has great value to water science and to watershed managers, and will allow for loads of chemicals to be traced from the stream back to their source from fastflow and slowflow with great accuracy and high temporal resolution. This can help to determine appropriate management strategies for reducing the transport of those chemicals to the stream.

References

Ator, S.W., J.M. Denver, and M.J. Brayton. 2005. Hydrologic and geochemical controls on pesticide and nutrient transport to two streams on the Delmarva Peninsula. US Department of the Interior, US Geological Survey.

Baron, J.S., E.K. Hall, B.T. Nolan, J.C. Finlay, E.S. Bernhardt, J.A. Harrison, F. Chan, and E.W. Boyer. 2013. The interactive effects of excess reactive nitrogen and climate change on aquatic ecosystems and water resources of the United States. *Biogeochemistry*, *114*(1-3), 71-92, doi: 10.1007/s10533-012-9788-y

Christophersen, N., C. Neal, R.P. Hooper, R.D. Vogt, and S. Andersen. 1990. Modeling streamwater chemistry as a mixture of soilwater end-members—a step towards second-generation acidification models. Journal of Hydrology, 116(1), 307-320, doi:10.1016/0022-1694(90)90130-P

David, M.B., L.E. Drinkwater, and G.F. McIsaac. 2010. Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, *39*(5), 1657-1667, doi:10.2134/jeq2010.0115

Eckhardt, K. 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes, 19(2), 507-515, doi: 10.1002/hyp.5675

Goodridge, B.M., and J.M. Melack. 2012. Land use control of stream nitrate concentrations in mountainous coastal California watersheds. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, *117*(G2), doi: 10.1029/2011JG001833

Holman, I.P., T.M. Hess, and S.C. Rose. 2011. A broad-scale assessment of the effect of improved soil management on catchment baseflow index. Hydrological Processes, 25(16), 2563-2572. doi: 10.1002/hyp.8131

Hooper, R.P. 2003. Diagnostic tools for mixing models of stream water chemistry. *Water Resources Research*, *39*(3), doi: 10.1029/2002WR001528

Hooper, R.P., N. Christophersen, and N.E. Peters. 1990. Modeling streamwater chemistry as a mixture of soilwater end-members—An application to the Panola Mountain catchment, Georgia, USA. Journal of Hydrology, 116(1), 321-343, doi:10.1016/0022-1694(90)90131-G

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(1), 241-258, doi:10.1016/j.jhydrol.2006.07.006

Kronholm, S.C., and P.D. Capel. 2014. A comparison of high-resolution specific conductance-based end-member mixing analysis and a graphical method for baseflow separation of four streams in hydrologically challenging agricultural watersheds. Hydrological Processes, doi: 10.1002/hyp.10378

Lee, C.J., D.P. Mau, and T.J. Rasmussen. 2005. Effects of nonpoint and selected point contaminant sources on stream-water quality and relation to land use in Johnson county, northeastern Kansas, October 2002 through June 2004: U.S. Geological Survey Scientific Investigations Report 2005-5144, 104 p.

McNamara J.P., D.L. Kane, L.D. Hinzman. 1997. Hydrograph separations in an Arctic watershed using mixing model and graphical techniques. *Water Resources Research* **33**(7) : 1707-1719. doi: 10.1029/97WR01033

Miller, M.P., D.D. Susong, C.L. Shope, V.M. Heilweil, and B.J. Stolp. 2014. Continuous estimation of baseflow in snowmelt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach. *Water Resources Research*, *50*(8), 6986-6999, doi: 10.1002/2013WR014939

Molénat, J., P. Durand, C. Gascuel-Odoux, P. Davy, and G. Gruau. 2002. Mechanisms of nitrate transfer from soils to stream in an agricultural watershed of French Brittany, Water Air Soil Pollut., 133(1-4), 161–183. doi: 10.1023/A:1012903626192

Pellerin B.A., W.M. Wollheim, X. Feng, and C.J. Vörösmarty. 2008. The application of electrical conductivity as a tracer for hydrograph separation in urban catchments. *Hydrological Processes* **22**(12) : 1810-1818. doi: 10.1002/hyp.6786

Peters, N.E. 1994. Water-quality variations in a forested Piedmont catchment, Georgia, USA. J. Hydrol. 156(1-4), 73-90. doi:10.1016/0022-1694(94)90072-8

Pilgrim, D.H., D.D. Huff, and T.D. Steele. 1979. Use of specific conductance and contact time relations for separating flow components in storm runoff. *Water Resources Research* **15**(2) : 329-339. doi: 10.1029/WR015i002p00329

Poor, C.J. and J.J. McDonnell. 2007. The effects of land use on stream nitrate dynamics. *Journal of Hydrology*, *332*(1), 54-68, doi:10.1016/j.jhydrol.2006.06.022

Rasmussen, T.J., C.J. Lee, and A.C. Ziegler. 2008. Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006: U.S. Geological Survey Scientific Investigations Report 2008–5014, 103 p.

Raymond, P.A., M.B. David, and J.E. Saiers. 2012. The impact of fertilization and hydrology on nitrate fluxes from Mississippi watersheds. *Current Opinion in Environmental Sustainability*, *4*(2), 212-218, doi: 10.1016/j.cosust.2012.04.001

Rice, K.C. and O.P. Bricker. 1995. Seasonal cycles of dissolved constituents in streamwater in two forested catchments in the mid-Atlantic region of the eastern USA, J. of Hydrol. 170(1-4), 137-158. doi:10.1016/0022-1694(95)92713-N

Rimmer, A. and A. Hartmann. 2014. Optimal hydrograph separation filter to evaluate transport routines of hydrological models. Journal of Hydrology, 514, 249-257, doi:10.1016/j.jhydrol.2014.04.033

Ross, D.S., R.J. Bartlett, F.R. Magdoff, and G.J. Walsh. 1994. Flow path studies in forested watersheds of headwater tributaries of Brush Brook, Vermont. Water Resour. Res. 30(9), 2611-2618. doi:10.1029/94WR01490

Sanford, W.E., D.L. Nelms, J.P Pope, and D.L Selnick. 2012. Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis. U.S. Geological Survey Scientific Investigations Report 2011–5198, Reston, VA.

Schwartz, S.S. 2007. Automated Algorithms for Heuristic Base-Flow Separation. *JAWRA Journal of the American Water Resources Association*, *43*(6), 1583-1594, doi: 10.1111/j.1752-1688.2007.00130.x

Senus, M.P., M.J. Langland, and D.L. Moyer. 2005. Nutrient and sediment concentrations, loads, and trends for four nontidal tributaries in the Chesapeake Bay watershed, 1997-2001. US Department of the Interior, US Geological Survey.

Smith, E.A. 2012. *Spatial and temporal variability of preferential flow in a subsurface-drained landscape in north-central Iowa*. Doctoral dissertation, Water Resources Science. University of Minnesota, St. Paul, MN

Sloto, R.A., and M.Y. Crouse. 1996. HYSEP: A computer program for streamflow hydrograph separation and analysis: U.S. Geological Survey Water-Resources Investigations Report 96-4040, Reston, VA.

United States Geological Survey (USGS) 2013b. *USGS Current Conditions for the Nation*. http://waterdata.usgs.gov/usa/nwis/uv?site_no=01493112

United States Geological Survey (USGS) 2013c. *USGS Current Conditions for the Nation*. Retrieved from http://waterdata.usgs.gov/usa/nwis/uv?06893390

Wahl, K.L., and T.L. Wahl. 1995. Determining the Flow of Comal Springs at New Braunfels, Texas. *Texas Water '95*. American Society of Civil Engineers. August 16-17, 1995. San Antonio, Texas: 77-86.

Woodward, S.J., R. Stenger, and V.J. Bidwell. 2013. Dynamic analysis of stream flow and water chemistry to infer subsurface water and nitrate fluxes in a lowland dairying catchment. *Journal of Hydrology*, *505*, 299-311, doi:10.1016/j.jhydrol.2013.07.044

Table 3.1: Comparison of the actual and ratio-EMMA estimates of fastflow volume, slowflow index (SFI), and chemical A and B loads under various scenarios of the timing of peak streamflow relative to the minimum (or maximum) $[A]_{FF}$ and $[B]_{FF}$.

	Fastflow $(\times 10^6\ m^3)^\dagger$	% Error[‡]	SFI[*] (%)	Difference in SFI[#] (%)	Chemical A load $(\times 10^3\ kg)$ **Actual** *Estimated*	% Error[‡]	Chemical B'' load $(\times 10^6\ kg)$ **Actual** *Estimated*	% Error[‡]
Groundwater dominated synthetic stream								
Actual Fastflow	**20.2**		**68.3**					
$t_{Qmax}=t_{[A]min}=t_{[B]min}$	20.3	0.2	68.3	0.0	**3.4** *3.6*	6.8	**170** *185*	8.9
$t_{Qmax}>t_{[A]min}=t_{[B]min}$	20.2	-0.3	68.4	0.1	**4.1** *3.9*	-4.5	**204** *187*	-8.0
$t_{Qmax}<t_{[A]min}=t_{[B]min}$	19.9	-1.4	68.8	0.5	**7.9** *5.9*	-26	**397** *233*	-41.3
$t_{Qmax}=t_{[A]min}>t_{[B]min}{}^+$	20.2	1.5	66.2	-2.1	**3.4** *3.0*	-11	**204** *172.5*	-15.2
$t_{Qmax}=t_{[A]min}<t_{[B]min}{}^+$	20.2	0.2	67.4	-0.9	**3.4** *3.5*	3.3	**397** *405.9*	2.2
$t_{[A]min}<t_{Qmax}<t_{[B]min}{}^+$	20.4	0.8	67.7	-0.6	**4.1** *5.3*	30	**397** *489.9*	23.4
$t_{Qmax}=t_{[A]max}=t_{[B]min}$	22.2	9.0	65.2	-3.1	**180** *182*	1.1	**170** *1252*	637
Overland flow dominated synthetic stream								
Actual Fastflow	**49.1**		**23.1**					
$t_{Qmax}=t_{[A]min}=t_{[B]min}$	48.8	-0.6	23.5	0.4	**11** *9.6*	-17	**574** *425*	-26
$t_{Qmax}>t_{[A]min}=t_{[B]min}$	48.5	-0.7	24.0	0.9	**15** *11*	-28	**749** *423*	-43
$t_{Qmax}<t_{[A]min}=t_{[B]min}$	48.2	-0.8	24.6	1.5	**20** *13*	-34	**982** *458*	-53
$t_{Qmax}=t_{[A]min}>t_{[B]min}{}^+$	48.4	0.6	23.7	0.6	**11** *6.7*	-42	**749** *377.2*	-50
$t_{Qmax}=t_{[A]min}<t_{[B]min}{}^+$	49.0	1.2	22.9	-0.2	**11** *11*	-6.4	**982** *926*	-5.7
$t_{[A]min}<t_{Qmax}<t_{[B]min}{}^+$	49.4	0.7	22.6	-0.5	**15** *17*	11	**982** *1105*	13
$t_{Qmax}=t_{[A]max}=t_{[B]min}$	50.1	1.4	21.6	-1.5	**449** *450*	0.2	**574** *1078*	88

† Total fastflow over the 14 day period of record
‡ Percent error between actual and estimated fastflow values over the 14 day period of record
* SFI is the Slowflow index (Equation 5)
# Difference between actual SFI and estimated SFI
+ Some outliers removed. Linear interpolation used to replace values.
" Converted from specific conductance to total dissolved solids

A
$Q_S, [A]_S, [A]_{SF}$

B
Calculate 1000 possibilities of $Q_{SF,A}$

C
$Q_S, [B]_S, [B]_{SF}$

D
Calculate 1000 possibilities of $Q_{SF,B}$

E
Store first (or next) $Q_{SF,A}$

F
Search for $Q_{SF,B} = Q_{SF,A}$

G
$Q_{SF,A} = Q_{SF,B}$?

Yes   No

H
Range of ratios of $[A]_{FF}$ to $[B]_{FF}$

I
Ratio of $[A]_{FF}$ to $[B]_{FF}$ within defined range of ratios?

Yes   No

J
Input $[A]_{FF}$ into standard EMMA
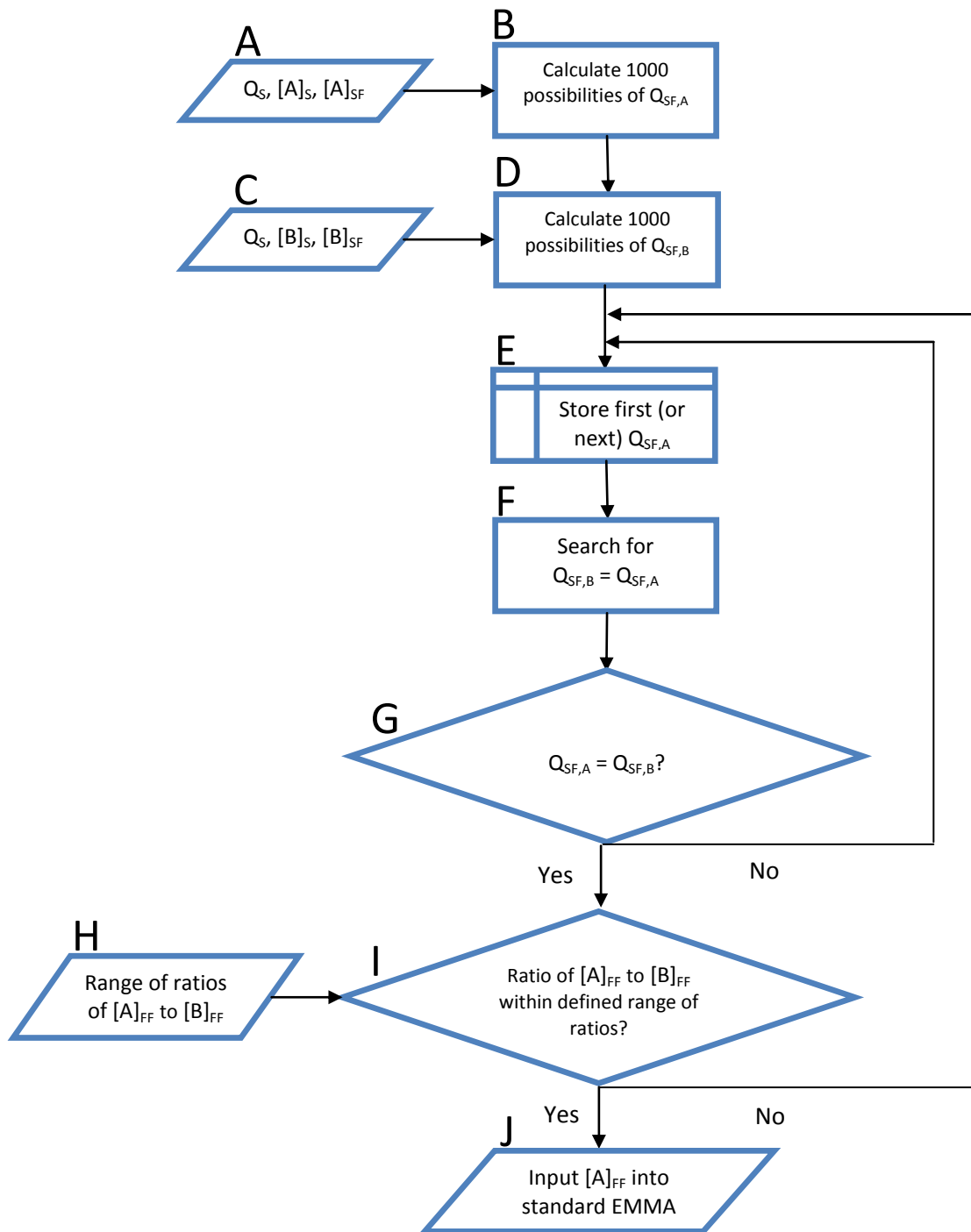
Figure 3.1: Ratio-EMMA logic flowchart for the selection of the time-variable fastflow concentrations of the chosen chemicals. Q is water discharge (m^3/s), [A] and [B] are the concentrations of chemical A (mg/L) and B (μS/cm), respectively. Subscripts: S = stream; SF = slowflow; FF = fastflow; SF,A = estimate of slowflow obtained using chemical A; SF,B = estimate of slowflow obtained using chemical B.
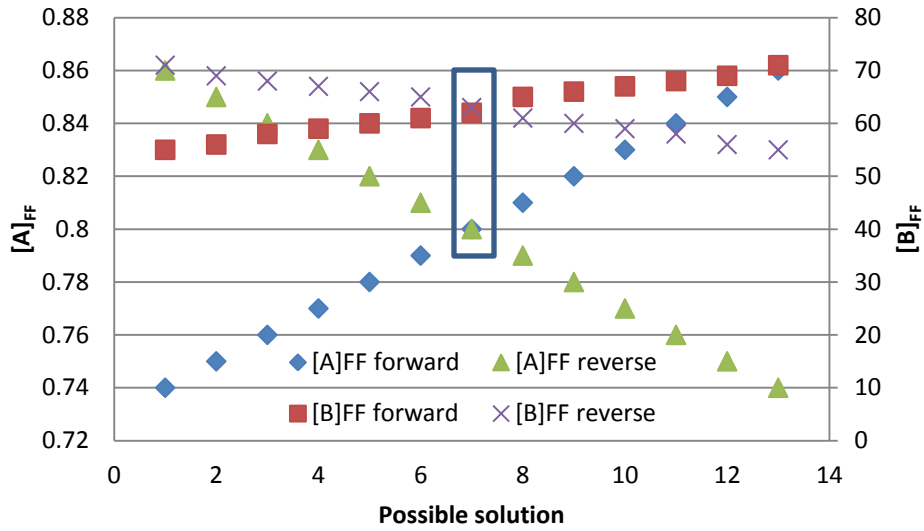
Figure 3.2: An example of all possible values of $[A]_{FF}$ and $[B]_{FF}$ for a single time interval when running the ratio-EMMA model forward and in reverse. The box encloses the values that are selected by the model as the estimated values for the time period.
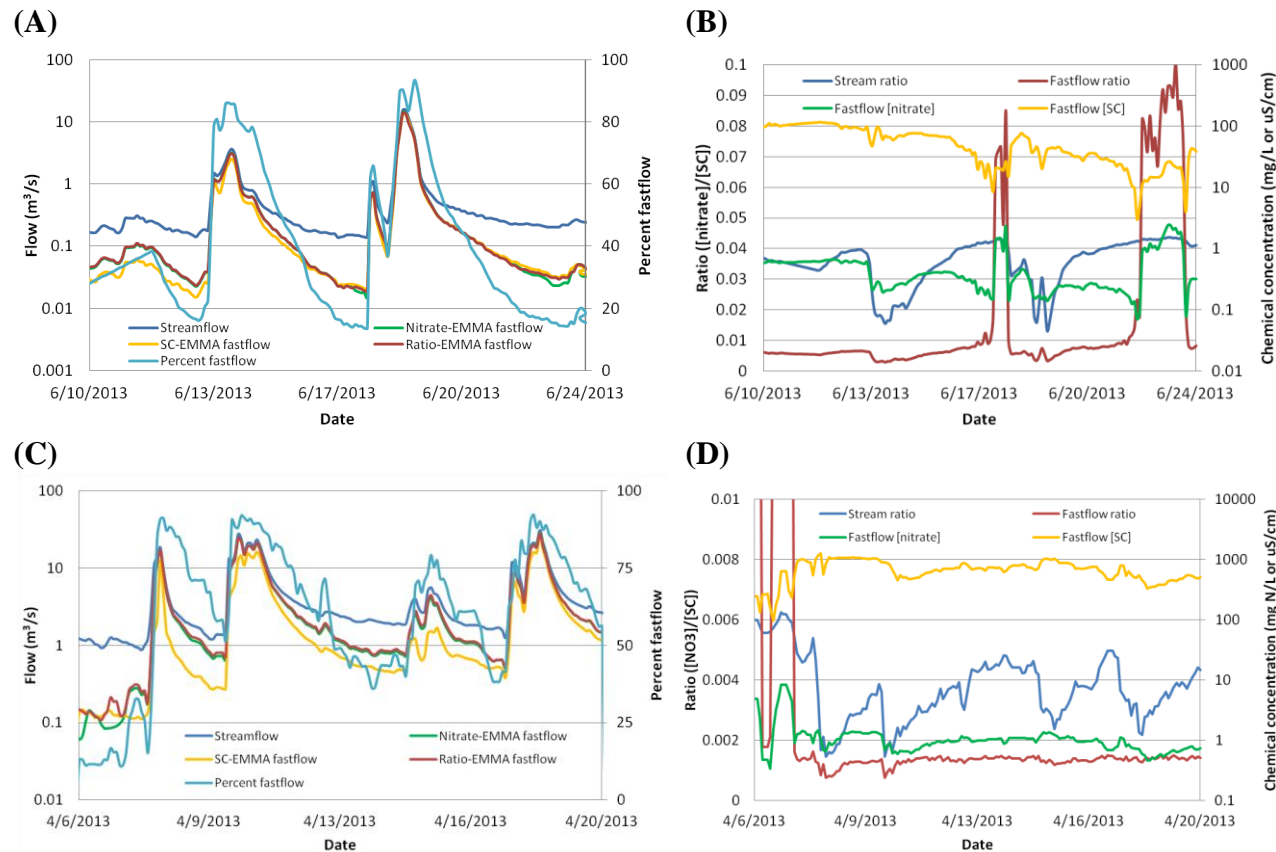
Figure 3.3: (**A**) Streamflow hydrograph, estimated fastflow hydrographs using a standard nitrate-EMMA, a standard SC-EMMA, and ratio-EMMA, and percent fastflow (as a percent of total streamflow) for Chesterville Branch, MD. (**B**) Actual ratio of [nitrate] to [SC] in the stream based on field measurements, and the ratio-EMMA estimated ratio of $[nitrate]_{FF}$ to $[SC]_{FF}$, together with $[nitrate]_{FF}$ and $[SC]_{FF}$. Similar figures are also shown for Indian Creek, KS (**C** and **D**). Large spikes in fastflow ratio (as seen in **B** and **D**) were a result of the ratio-EMMA producing possible concentrations of both $[A]_{FF}$ and $[B]_{FF}$ that were highly varied, but the resulting range of ratios of $[A]_{FF}$ to $[B]_{FF}$ was very small.

63

(A)

(B)
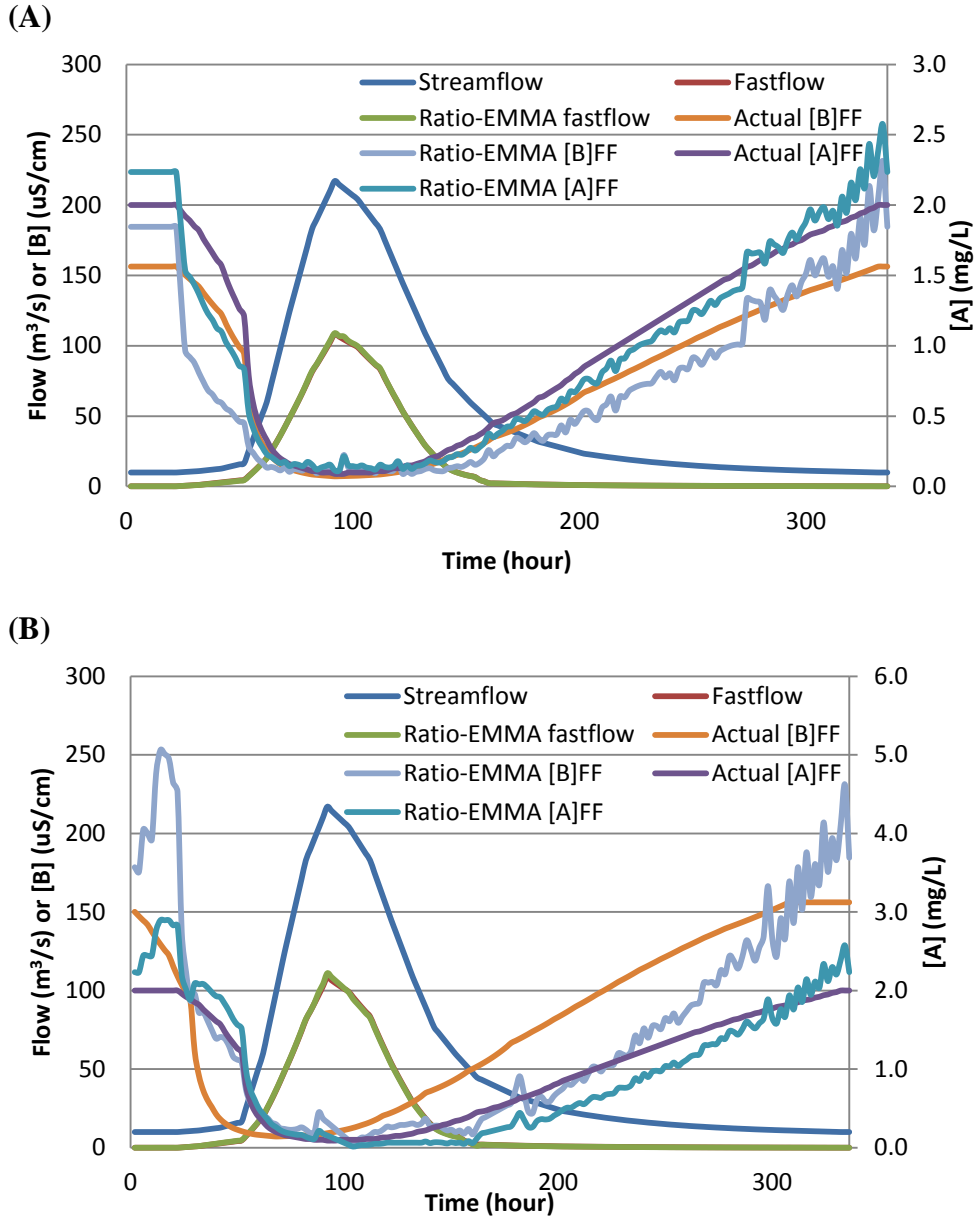
Figure 3.4: Ratio-EMMA estimated slowflow, $[A]_{FF}$, and $[B]_{FF}$ compared to actual values in the groundwater dominated synthetic stream where (**A**) the streamflow peak, minimum $[A]_{FF}$, and minimum $[B]_{FF}$ occurred concurrently ($t_{Qmax}=t_{[A]min}=t_{[B]min}$) and when (**B**) the streamflow peak and the minimum $[A]_{FF}$ occur concurrently, but minimum $[B]_{FF}$ occurred 24 hours earlier ($t_{Qmax}=t_{[A]min}>t_{[B]min}$). The slight fluctuations of estimated $[A]_{FF}$ and $[B]_{FF}$ for any given line are a result of the model needing to find a concentration of $[A]_{FF}$ and $[B]_{FF}$ that meet both necessary criteria ($Q_{SF}{:}A = Q_{SF}{:}B$ and $[A]_{FF}/[B]_{FF}$ between the defined range of ratios). Other figures showing the other temporal scenarios of streamflow, $[A]_{FF}$, and $[B]_{FF}$ minima can be seen in Appendix C.

(A)

(B)

Figure 3.5: Ratio-EMMA estimated slowflow, $[A]_{FF}$, and $[B]_{FF}$ compared to actual values in the overland flow dominated synthetic stream where (**A**) the streamflow peak, minimum $[A]_{FF}$, and minimum $[B]_{FF}$ occurred concurrently ($t_{Qmax}=t_{[A]min}=t_{[B]min}$) and when (**B**) the streamflow peak and the minimum $[A]_{FF}$ occur concurrently, but minimum $[B]_{FF}$ occurred 24 hours earlier ($t_{Qmax}=t_{[A]min}>t_{[B]min}$). The slight fluctuations of estimated $[A]_{FF}$ and $[B]_{FF}$ for any given line are a result of the model needing to find a concentration of $[A]_{FF}$ and $[B]_{FF}$ that meet both necessary criteria ($Q_{SF}:A = Q_{SF}:B$ and $[A]_{FF}/[B]_{FF}$ between the defined range of ratios). Other figures showing the other temporal scenarios of streamflow, $[A]_{FF}$, and $[B]_{FF}$ minima can be seen in Appendix C.

(A)

(B)

Figure 3.6: Ratio-EMMA estimated slowflow, $[A]_{FF}$, and $[B]_{FF}$ compared to known values in (**A**) the groundwater dominated synthetic stream and (**B**) the overland flow dominated stream when the streamflow peak, maximum $[A]_{FF}$, and minimum $[B]_{FF}$ occurred concurrently ($t_{Qmax}=t_{[A]max}=t_{[B]min}$).

Chapter 4: A comparison of nitrate concentrations, loads, and yields in six streams in different hydrologic settings

Abstract

Nitrate is a common contaminant in streams draining agricultural watersheds. The spatial magnitude and temporal variability of nitrate concentrations, loads, watershed yields, and loads as percent of nitrogen use (LAPU) were examined in six small, gaged streams in six different hydrologic settings. The sources of water to the six streams were determined based on hydrograph separation and prior knowledge of the hydrology, irrigation, and subsurface drainage. For two streams, overland flow was the most import source of water. Two others had substantial subsurface drainage in their watersheds, and two others had large groundwater contributions from aquifers that were contaminated with nitrate (~50% of total streamflow was from baseflow). The subsurface drain-fed streams had the highest concentrations of nitrate, and also had the greatest loads, watershed yields, and LAPUs. Streams where overland flow was the major water flowpath had the lowest concentrations, loads, and yields. The two groundwater-fed streams had intermediate nitrate concentrations that were inversely related to streamflow. The highest concentrations occurred during periods of low flow due to the high inputs from contaminated groundwater. For a single, typical storm even, high-resolution (15 minute interval) nitrate concentration and streamflow were compared for three different streams draining agricultural watersheds which have substantial contributions of subsurface drain flow, overland flow, or groundwater flow. The maximum nitrate concentration preceded the maximum streamflow for the overland flow watershed. The two maxima were concurrent in the subsurface drain watershed, but in the groundwater watershed, the peak nitrate was concurrent with the minimum streamflow. The observed differences in the nitrate concentrations, loads, and yields in the six watersheds are attributed, in part, to differences in the flowpaths of water to the stream, providing insight into the hydrologic process controlling the movement of nitrate.

1. <u>Introduction</u>

Nitrogen fertilizers are used to provide an essential nutrient to agricultural crops, but not all of the supplied nitrogen is utilized by the plants. Several studies (Bijeriego et al., 1979, Meisinger et al., 1985, Olson, 1980, Reddy and Reddy, 1993) have shown that when growing corn, for example, 43 to 76 percent of the applied nitrogen is not used by the plants during the first season after the fertilizer is applied. Some portion of the "surplus" nitrogen can be lost from the field through volatilization, runoff, or leaching. Nitrate is very mobile in the environment and is commonly a contaminant in streams and groundwater (Dubrovsky et al., 2010). Streams that drain predominantly agricultural land have higher nitrogen concentration when compared to streams draining land dedicated to other uses (Dubrovsky and Hamilton, 2010).

Excess nitrate within agricultural streams can negatively impact the overall health of the stream ecosystem by reducing the number and variety of algae, macroinvertebrates, and fish growing within the stream (Dubrovsky and Hamilton, 2010). Agriculturally applied nitrogen has also been identified as the key nutrient that supports the enlargement of the annual hypoxic region in the Gulf of Mexico (Ribaudoa et al., 2005); with a majority of that nitrogen coming from the Midwestern United States (Blann et al., 2009, Kladviko et al., 1991, Sauer et al., 2008). The excess nitrogen promotes extensive algae growth which leads to hypoxia as the algae decompose.

Nitrate concentration within agricultural streams can vary over short time periods such as storm events, and over longer, seasonal periods. Changes occur in response to past and current land management practices as well as physical, chemical, and (or) biological interactions between nitrogen and the surrounding environment (Poor and McDonnell, 2007 and Ruiz et al., 2002). Land management can change the quantity, form, and timing of nitrogen entering the field, and can alter the way in which water and waterborne chemicals move across and through the soil. Properties of the soil and soil environment, including biota, can also greatly alter the amount and species of nitrogen that are present in streams (Tesoriero et al., 2009).

The complex and interconnected system of routes (flowpaths) through which water moves regulates the movement of water and waterborne chemicals/nutrients from

the landscape to a stream (Green, 2007). Flowpaths such as overland flow, subsurface drainage, and groundwater discharge all convey water and nitrate to streams at different rates and in different quantities.   The temporal nature of each flowpath determines how quickly nitrate will travel from the land surface to the stream and in what quantity. The measured stream nitrate concentration is a combination of individual contributions of nitrate from each water flowpath (Cassell and Clausen, 1993). Fluctuations in water contributions from a given flowpath has been observed in many studies and has been shown to dramatically alter nutrient concentration within a stream over the course of a year (Hooper et al., 1990, Molénat et al., 2002, Peters, 1994, Rice and Bricker, 1995, Ross et al., 1994). The current understanding of nitrate loads and concentrations is largely based on discrete water samples collected over many years. In recent years, nitrate sensors with high temporal resolution (usually 15 to 60 minutes) have been developed and begun to be deployed to better define the changes in the time-variable nitrate concentrations and loads, and to better understand the processes that transport nitrate to and through the stream (*USGS*, 2013a).

The nitrate concentrations and loads observed in a stream are controlled, in part, by the hydrologic setting of the stream—the characteristic sources of water to that stream. Temporal nitrate concentration patterns and the magnitudes of concentrations, loads, watershed yields, and loads as a percent of nitrogen use (LAPU) were compared for six, small, well-studied streams in distinctly different, but common hydrologic settings. From previous studies, the flowpaths of water to the streams and the use of nitrogen fertilizer in the watersheds were known. Streamflow and concentrations and loads of nitrate were also known. The hydrologic settings for the six streams were characterized as perennial overland flow-fed, ephemeral overland flow-fed, irrigation-derived groundwater-fed, natural groundwater-fed, subsurface drain-fed with a shallow subsurface confining layer, and subsurface drain-fed without a shallow subsurface confining layer. High-resolution (15 minute interval) nitrate concentration and streamflow were also compared in streams in three of the hydrologic settings for a single storm event. These comparisons of time-variable nitrate concentrations in the stream, as a function of hydrologic setting, provide

insight into the controlling hydrologic processes and help set expectations for the changes of nitrate in the stream relative to changes in management practices in the watershed.

2. Methods

2.1   Characterization of water flowpath

The characteristic water flowpath to streams was used as the basis for analyzing variations in the magnitude and temporal patterns of nitrate concentration in six small, agricultural watersheds in different hydrologic settings (Table 4.1). The hydrology of these watersheds has been previously characterized (Capel et al., 2008, Gronberg and Kratzer, 2006, Hancock and Brayton, 2006, Lathrop, 2006, McCarthy et al., 2012, Payne et al., 2007).

The characteristic flowpaths contributing water to the streams were determined by previous studies of the hydrology of the watersheds and by hydrograph separation based on daily streamflow using the US Bureau of Reclamation BFI program (Wahl and Wahl, 1995, 2007). Within the BFI program there are two variables—length of the time increment chosen to divide the water year and percentage of the increment minimum streamflow that is used in the determination of turning points (default value used). The time increment variable was selected as the value at which baseflow index (BFI) became nearly constant. The six streams were characterized based on the streamflow, the calculated BFI, and knowledge of irrigation and(or) subsurface drainage practices in the watershed (Table 4.1).

2.2   Sampling and nitrate analysis

Streamflow was measured at a USGS gage (Table 4.1). Water samples for nitrate analysis were collected over a period of 2 years. The sampling years varied among the watersheds. Sampling, measurements, and analytical procedures were conducted according to U.S. Geological Survey standards (Capel et al., 2008, Shelton, 1994, Wilde et al., 1999). Water samples were collected approximately bi-weekly during the growing season and less frequently during the non-growing months. Manual stream water samples were collected using an isokinetic, equal-width-increment, depth-integrated, cumulative method. All individual vertical samples were combined, homogenized, and sub-sampled.

Auto-samplers also were used to collect hourly or bi-hourly stream water samples in response to increased flow due to precipitation events at all of the sites except for the watersheds in Washington and California.

Samples for nitrate analysis were filtered in the field through a 0.45-µm Supor hydrophilic polyethersulfone membrane filter (Pall Corporation, Port Washington, NY). Samples were stored in clean plastic bottles, and shipped overnight on ice in insulated coolers to the laboratory for analysis (*USGS*, 2006). Nitrate concentration was calculated mathematically by subtracting nitrite concentration from the combined concentration of nitrate and nitrite (USEPA Parameter Codes 00613 and 00631). Concentrations of nitrite and combined nitrate and nitrite were measured using an Alpkem rapid flow analyzer (OI Analytical, College Station, TX).

Nitrate concentration was assumed to be 0.0 mg N/L when the measured concentration was below the minimum analytical reporting limit. Quality-control samples were collected in the field. For every 20 water samples collected, one blank and one replicate were collected. Blank samples were composed of deionized water that was known to not contain the analyte of interest above the minimum reporting limit of the analytical method. Additional quality control measures were taken in the laboratory where the water samples were analyzed (Capel et al., 2008). Additional details of sample collection and processing can be found in Capel et al. (2008).

2.3   Calculation of loads and yields

Calculation of annual measurements (mean and median streamflow, annual water discharge, annual baseflow discharge, annual water yield, mean and median nitrate concentration, annual stream nitrate load, and annual watershed yield) was based on a water year (October 1 through September 30). The software program LOADEST, which uses seasonally defined relationships between the discrete nitrate concentration and streamflow measurements (Runkel et al., 2004), was used for estimating nitrate loads for the streams (Morgan Creek, MD, South Fork Iowa River, Blairsburg, IA, Leary Weber Ditch, IN, and Tommie Bayou, MS). The LOADEST program could not be used with the ephemeral overland flow-fed stream (Mustang Creek, CA) or the irrigation-derived groundwater-fed stream (Granger Drain, WA). In the ephemeral Mustang Creek,

continuous streamflow was recorded at 10 minute intervals and discrete measurements of nitrate concentration were collected during the few streamflow events. Linear interpolation was used to estimate nitrate concentration during the unmeasured periods. Total discharge during each 10 minute period was multiplied by the nitrate concentration during that time period to estimate the load during each 10 minute period. Annual load was calculated as the sum of all 10 minute loads within a given water year. In Granger Drain, linear interpolation was used to estimate daily nitrate concentrations between measured values. Loads were calculated in a similar manner as in Mustang Creek, with daily streamflow values used for calculations in Granger Drain.

The nitrate load in each of the streams was divided by the annual mass of nitrogen fertilizer that was used within the watershed and reported as the load as a percent of use (LAPU). There are definite limitations on the calculations and use of the LAPU value for comparative purposes. To do this calculation, the load in the stream for each year was assumed to be a direct result of the fertilizer use for that year. Given that the nitrogen residence time in soil and groundwater can be much longer than one year, this assumption is violated in the six watersheds to varying degrees. Nevertheless, it is a useful comparison tool.

2.4   Measurement of continuous nitrate concentration

Continuous nitrate and streamflow data were obtained for three additional watersheds. Two of these watersheds are the South Fork Iowa River, New Providence, IA (580 km^2, USGS Site 05451210) and Bogue Phalia, Leland, MS (1,250 km^2, USGS Site 07288650). The small watersheds in Iowa and Mississippi discussed earlier are sub-basins of these larger watersheds. The third watershed is the Chesterville Branch, Crumpton, MD (15.9 km2, USGS Site 01493112) which is an adjacent watershed to Morgan Creek, MD. The data from these additional streams is used because the six streams mentioned previously do not have high-temporal resolution nitrate data available. The high-temporal resolution nitrate concentrations were measured with a YSI EXO (YSI, Yellow Springs, OH) in Maryland and Mississippi, and with a Hydrolab Series 5 (Hach Hydromet, Loveland, CO) in Iowa. The instruments were calibrated with standards. The data was quality assured by comparing the nitrate concentration in discrete

water samples with the sensor data. Only a single storm event is reported here, but longer records are available (USGS, 2013d).

3. Results

3.1 Overview of watershed characteristics

The hydrologic setting, watershed characteristics, nitrogen use, streamflow, and the concentrations, loads, and yields of nitrate are summarized for the six streams in Table 4.1. The characteristic water flowpaths to the streams are characterized by the baseflow index (BFI) and prior knowledge of subsurface drainage (South Fork Iowa River, Blairsburg and Leary Weber Ditch) and irrigation (Granger Drain and Mustang Creek) in each watershed. The average BFI ranged from 14% for Tommie Bayou to 97% for Granger Drain and helped characterize the important flowpath to the stream (overland flow-fed and groundwater-fed, respectively). The precipitation was near normal for both years in the six watersheds except for 2008 in South Fork Iowa River at Blairsburg, IA. In April-June of 2008, precipitation in Iowa was much greater than normal. This caused flooding throughout much of Iowa including the South Fork Iowa River, Blairsburg watershed. The transport of chemicals was much greater that year compared to years with normal precipitation (Hubbard et al., 2011). Median streamflow ranged from <0.1 to 1 m^3/s, excluding the ephemeral Mustang Creek, which had no streamflow the majority of the days.

The nitrogen fertilizer application rate is the mass of nitrogen in fertilizer divided by the total area of the watershed, not the application rate on a cropped field. These values are displayed in Table 4.1 along with the total mass of applied nitrogen. Application rates ranged by nearly a factor of 12, from 12 to 139 kg N/ha. The watershed yields of nitrate ranged from 0.1 to 77 kg N/ha/yr for all watersheds (Table 4.1), and the LAPU values ranged from 0.1 to 93%.

There was a 35-fold difference between the highest and lowest median nitrate concentration in the six small streams. The median annual concentration and distribution of nitrate concentrations is given in Figure 4.1 for each of the streams. The two overland flow-fed streams had the lowest median concentrations and small distributions of

concentrations. The two groundwater-fed streams generally had somewhat greater median concentrations and larger distributions. The subsurface-drain-fed streams had the greatest median concentrations and largest distributions of nitrate concentrations. Streamflow and nitrate concentrations for one year are shown in Figure 4.2. Nitrate concentration as a function of streamflow is shown in Figure 4.3.

3.2   Overland flow-fed streams

The perennial Tommie Bayou, MS watershed is characterized as humid, but there is still a need for irrigation to provide adequate water to the crops during portions of the year. Much of the stream water within Tommie Bayou comes from runoff or overland flow as a result of precipitation, irrigation, or release of water from rice fields within the watershed (Table 4.1), as there is little connection between the stream and the shallow groundwater (McCarthy et al., 2012). Median streamflow was less than 0.2 m^3/s. The maximum streamflow was over 20 m^3/s during the two study years. Nitrate concentrations within Tommie Bayou were low when compared to the other streams. The maximum concentration during the two sampled years was 7.5 mg N/L with a median concentration of 0.38 mg N/L (Figure 4.1). Nitrate concentrations within Tommie Bayou reached maximum levels during the last part of April, decreased through the summer months, and remained low through the winter months during both years (Figure 4.2A). There was not a discernible trend in nitrate concentration relative to streamflow (Figure 4.3A).

The ephemeral Mustang Creek, CA watershed is naturally arid and requires irrigation to provide adequate water to the crops. All of the water in Mustang Creek comes from runoff, overland flow from irrigation, or, infrequently, from precipitation. This results in the stream flowing just a few times each year (Gronberg and Kratzer, 2006). Streamflow was only >1.0 m^3/s during the occasional storm flow peak. In Mustang Creek, the nitrate concentration was quite low, similar to Tommie Bayou (Figure 4.1). Due to the sporadic nature of flow within Mustang Creek, nitrate concentration in the stream came in the form of short duration pulses when the stream was flowing (Figure 4.2B). There was not a discernible trend in nitrate concentration relative to streamflow (Figure 4.3B).

74

3.3   Subsurface drain-fed streams

The South Fork Iowa River, Blairsburg, IA watershed is characterized by moderate temperatures and rainfall that is often adequate to meet the requirements of the crops. A substantial portion of the watershed is drained through subsurface drainage. At a depth of 1-3 m below land surface, there is a clay confining layer throughout much of the watershed. This prevents a majority of the soil water from reaching groundwater and limits the connection between the stream and the shallow groundwater (McCarthy et al., 2012). The drainage network efficiently moves excess soil water to the stream. During periods of low flow in the stream, the water discharging from the subsurface drains can account for almost all streamflow (Thornburg, 2009), whereas water at higher flows comes from a combination of subsurface drainage, surface drain networks, and overland flow. In 2008, which was an abnormally wet year, streamflow within the South Fork Iowa River, Blairsburg, IA reached a maximum of >15 m^3/s—over three times greater than the maximum streamflow during the 2007 water year. The South Fork Iowa River, Blairsburg, IA had the highest overall median concentration and range of nitrate concentrations of the six small streams (Figure 4.1). The median concentration was about five times greater than the groundwater-fed streams and 16 times greater than the overland flow-fed streams. The nitrate concentrations in the stream water remained elevated throughout the year (Figure 4.2C). The temporal pattern in nitrate concentration in the South Fork Iowa River, Blairsburg, IA saw an increase with increasing streamflow during the early spring through early summer, which was followed a rapid decrease in concentration during the late summer. Before and after the growing season, the concentrations were less than during the growing season (Figure 4.2C). Nitrate concentrations showed a general trend of increasing concentration with increasing streamflow, except at high flows (>1 m^3/s), where the concentration decreased (Figure 4.3C).

Leary Weber Ditch, Indiana is characterized by moderate temperatures and rainfall that is often adequate to meet the requirements of the crops. Nearly the entire watershed is drained by a network of subsurface drainage. During drier periods within the watershed, Leary Weber Ditch occasionally stopped flowing. This occurred several

times during the study period, totaling 97 days without flow during 2003 and 49 days without flow during 2004. Throughout the year, a majority of the streamflow is a result of subsurface drain discharge because there is very little input from overland flow or groundwater (Lathrop, 2006). The stream is generally disconnected from the groundwater, although there are small sand lenses in the subsurface throughout the watershed. Median streamflow was <0.1 m^3/s with maximum flows slightly >3.0 m^3/s. Median nitrate concentration in Leary Weber Ditch was elevated, second only to the other subsurface drain-fed stream (South Fork Iowa River, Blairsburg, IA). During both years, nitrate concentration peaked in early to mid-June (one higher concentration recorded in July 2003) and decreased over the remainder of the water year (Figure 4.2D). However, very few concentrations were recorded outside of the growing season. Nitrate concentration showed a weak to moderate positive relationship with streamflow (Figure 4.3D).

3.4 Groundwater-fed streams where groundwater is contaminated with nitrate

The Morgan Creek, MD watershed receives enough rain so that most farmers do not use irrigation. During much of the year a majority of the streamflow in Morgan Creek is from groundwater. About 51% of the annual streamflow was due to fastflow processes such as overland flow, in response to precipitation (Hancock and Brayton, 2006). In 2003, maximum streamflow was 6.0 m^3/s, but streamflow was <0.5 m^3/s for about 300 days during that water year. The magnitude and distribution of nitrate concentration in Morgan Creek was less than the subsurface drain-fed streams, but greater than the overland flow-fed streams (Figure 4.1). Groundwater underlying the Morgan Creek watershed is generally oxic. The stream had nitrate concentrations that were high during low-flow periods (from contaminated groundwater inputs) and decreased with increasing flow (Figure 4.3E). This trend was stronger in 2004 ($R^2 = 0.72$) than in 2003 ($R^2 = 0.26$, Figure 4.3E).

The Granger Drain, WA watershed requires irrigation to provide adequate water to the crops. Abundant irrigation water comes from snowmelt runoff channeled to the watershed. The long-term use of abundant irrigation water has increased the level of the

water table within the watershed. Engineered surface ditches, such as Granger Drain, keep the level of the shallow water table from rising to the soil surface (Payne et al., 2007). Throughout the non-irrigation season, streamflow within Granger Drain is a result of the ditch (stream) intercepting the water table. During the irrigation season, increased streamflow came from irrigation water that went unutilized by the crops, which flowed overland or travelled in the shallow groundwater to the stream. Streamflow was relatively constant, but substantially different in the irrigation season (1.4 $\pm$ 0.2 m^3/s) compared to the non-irrigation season (0.6 $\pm$ 0.1 m^3/s). Annual median nitrate concentrations were also relatively constant, but substantially different in the irrigation season (2.5 mg/L as N) compared to the non-irrigation season (6.6 mg/L as N). The temporal patterns of streamflow and nitrate concentration approximated the form of two opposite square waves (Figure 4.2F). During both years the nitrate concentration in the stream was high during the fall and winter (non-irrigation season, low flow period) resulting from older (>1 year), generally oxic groundwater with high nitrate concentrations. The stream nitrate concentration declined rapidly during the spring and was constant during the spring and summer as the irrigation water increased overall streamflow and diluted the groundwater. There was a very strong inverse relationship between stream nitrate concentration and streamflow (Figure 4.3F).

3.5   Summary of nitrate loads in the streams

Daily estimates of nitrate load were summed to equal annual loads (Table 4.1). Annual loads of nitrate in the streams were highly tied to watershed area, which made useful comparisons regarding flowpath difficult between watersheds. As a result, the cumulative percent of daily nitrate load and stream volume are presented as a function of sorted streamflow (Figure 4.4). Cumulative load plots show the percent of annual nitrate load that occurred under various streamflow conditions, and relate the load in the stream to the important flowpaths which transport nitrate. These figures will be discussed in greater detail in the Discussion section.

3.6   Continuous nitrate concentrations in the streams

Nitrate concentration and streamflow collected at 15 minute intervals over single storm events in three watersheds are plotted in Figure 4.5. Each watershed where 15-

minute data were collected is closely related (spatially and hydrologically) to previously mentioned watersheds. The natural groundwater-fed Morgan Creek is adjacent to, and is about twice as large as the Chesterville Branch. The perennial overland flow-fed Tommie Bayou is a sub-basin, comprising 1.2% of the Bogue Phalia watershed. The subsurface drain-fed South Fork Iowa River, Blairsburg is a sub-basin, comprising 5.4% of South Fork Iowa River, New Providence watershed. The same flowpaths are important in each of the related watersheds.

In the Bogue Phalia, MS (Figure 4.5A), the increase in streamflow is from a 2-day rain that occurred three days before the peak of streamflow. The nitrate concentration increased abruptly with an increase in streamflow and reached a maximum concentration before peak streamflow. In the South Fork Iowa River, New Providence, IA (Figure 4.5B), the increase in streamflow is from rain that started two days prior to the peak of streamflow. Much of the water moved off of the landscape within a few days, but some water took more than a week to get to the stream through the network of surface and subsurface drains. The nitrate concentration decreased with the initial increase in streamflow, then increased and peaked one day after the streamflow peak, and finally declined but at a slower rate than the decline in streamflow. In the Chesterville Branch, MD (Figure 4.5C), the increase in streamflow is from rain that occurred the day of the peak of streamflow. Streamflow prior to and after the peak is baseflow from groundwater which is contaminated with nitrate (Böhlke and Denver, 1995). The nitrate concentration decreased with an increase in streamflow and reached a minimum concentration just after peak streamflow, then increased again as the streamflow decreased.

4. Discussion

4.1 Comparisons among the six watersheds

The concentration of nitrate in a small stream draining an agricultural watershed is influenced by many factors. The fraction of total nitrogen in soil which exists as nitrate is the balance between its formation (nitrification) and loss (denitrification and uptake). The nitrate in the soil that comes in contact with moving water is generally the major source to streams and groundwater. Olarewaju et al. (2009) showed that most of

the soil nitrate is in the upper 0.45m of the soil column. Water that moves only across the surface comes in less contact with the soil compared to water which infiltrates the soil and moves downward. Domagalski et al. (2008) and Wang and Zhu (2011) have shown that the movement of nitrate from the landscape to the stream is small where overland flow is the major water flowpath. Infiltrated water can move through the soil column and recharge the groundwater or, if subsurface drainage is present, water can move to the drain and discharge to the stream. Subsurface drains are purposed to remove excess water from the soil and control the level of the water table. Dubrovsky et al. (2010) and others have shown that streams fed by subsurface drains generally have higher concentrations of nitrate than streams without subsurface drainage. Puckett et al. (2008) discussed that nitrate contaminated, oxic, shallow groundwater can be a very long-term source of nitrate to streams. However, if the groundwater is anoxic, then a portion of the nitrate is removed through denitrification. These differences in hydrologies produce the variability in nitrate concentrations, loads, and yields that were observed.

The water flowpaths, streamflow, nitrate concentrations, and nitrogen fertilizer inputs are known for the six streams in different hydrologic settings. These data and characterizations of the hydrologies provide the opportunity to compare and contrast the concentrations, loads, and yields of nitrate and provide insight into the important processes for water and nitrate movement.

The comparison of the six streams and their watersheds can be summarized as:

Nitrogen use:	OF(p) ~ SD(w/o cl) < OF(e) < SD(cl) ~ GW(n) < GW(i)
Median streamflow:	OF(e) ~ SD(w/o cl) < OF(p) < SD(cl) < GW(n) < GW(i)
Annual water discharge:	OF(e) < SD(w/o cl) < OF(p) < SD(cl) ~ GW(n) < GW(i)
Median nitrate concentration:	OF(p) < OF(e) < GW(n) < GW(i) << SD(w/o cl) < SD(cl)
Nitrate load:	OF(e) < OF(p) < SD(w/o cl) < GW(n) < GW(i) << SD(cl)
Nitrate yield:	OF(e) < OF(p) < GW(i) < GW(n) < SD(w/o cl) << SD(cl)
Nitrate LAPU:	OF(e) < GW(i) < GW(n) < OF(p) < SD(w/o cl) < SD(cl)

where OF(e) is the ephemeral overland flow-fed stream (Mustang Creek, based on data collected in water year 2004), OF(p) is the perennial overland flow-fed stream (Tommie Bayou, 2008), GW(i) is the irrigated groundwater-fed stream (Granger Drain, 2003), GW(n) is the natural groundwater-fed stream (Morgan Creek, 2003), SD(cl) is the subsurface drain-fed stream whose watershed is underlain by a shallow confining layer (South Fork Iowa River, Blairsburg, 2007), and SD(w/o cl) is the subsurface drain-fed stream whose watershed is not underlain by a shallow confining layer (Leary-Weber Ditch, 2003). LAPU is the load of nitrate as a percentage of annual nitrate fertilizer use.

The first three comparisons (nitrogen use, median streamflow, and annual water discharge) are descriptors of the streams that are largely based on the physical (area, amount of precipitation and(or) irrigation) and agricultural (nitrogen) characteristics of the watersheds. As would be expected, the order in the comparison above for the physical and agricultural characteristics does not have a relation to water flowpath. Nitrate concentration and nitrate yield from the watershed, on the other hand, are grouped according to flowpath.

4.2   Overland flow-fed streams

The variability in the magnitude of nitrate concentration was considerable among these six streams in the different hydrologic settings. The lowest median concentrations and smallest concentration distributions were found in the overland flow-fed streams (Figure 4.1). Measured concentrations varied randomly through the year within these streams and showed no relation to streamflow (Figures 4.2 and 4.3). The temporal patterns of nitrate concentration observed in the two overland flow-fed streams were also much more episodic compared to the other streams.

The relations between nitrate concentration and streamflow differed greatly among the six watersheds (Figure 4.3). Nitrate concentrations in the overland flow-fed streams were quite low at high or low flow, but very briefly spiked with increases in streamflow. This led to essentially no relation between nitrate concentration and streamflow for the two overland flow-fed streams. The episodic nature of rainfall, increased streamflow, and nitrogen application in these watersheds likely contributed to this lack of relation. Vidon et al. (2012) also found very little relationship between

streamflow and nitrate concentration in a stream within a watershed in which overland flow was an important flowpath.

The smallest annual nitrate loads were in the overland flow-fed streams. In the perennial Tommie Bayou, the annual load was 7,000 kg N/yr, whereas loads in the ephemeral Mustang Creek were only 230 kg N/yr. Much of the nitrate load in the overland flow-fed streams travelled during or immediately after highflow events. Almost 80% of the annual load in Mustang Creek occurred during one large rain event. Other times of the year had little or no nitrate moving to the stream. In Tommie Bayou, the percent of the cumulative load (Figure 4.4A) increased in a relatively even fashion across nearly all streamflow conditions due to the comparatively consistent nitrate concentration. Even so, only ~12% of the total annual load was transported under baseflow conditions.

The ephemeral, overland flow-fed Mustang Creek had the lowest nitrate yield and the perennial, overland flow-fed Tommie Bayou had the second lowest yield. Low yields from these overland flow dominated watersheds were likely a result of the minimal interactions between the run-off water and the soil, thus preventing extensive subsurface nitrate movement to the streams.

The perennial, overland flow-fed Tommie Bayou had a high LAPU of 36% in 2008 (Table 4.1), partially explained by the relatively small amount of nitrogen fertilizer applied compared to the other watersheds. The LAPU in Mustang Creek was very small (0.1%) which would be expected given the lack of flow to, and in the stream.

Figure 4.5 illustrates the substantial differences in the nitrate concentration responses for a single storm event for the three larger streams examined. Although each stream and each event are different, important generalities can be observed (particularly with this high-temporal resolution data) that provide insight into the important processes controlling the rate and amount of water and nitrate which move to the stream after a rain event. The continuous nitrate concentrations in Bogue Phalia, the large, overland flow-fed Mississippi stream, had an abrupt increase and rapid decrease in nitrate concentration preceding the increase in streamflow (Figure 4.5), suggesting that a small mass of nitrate on the soil surface is readily available for transport to the stream. As the runoff water

moves over the land surface and through the shallow soil, the available nitrate in this surface layer is quickly transported to the stream. The later, smaller nitrate peak suggests that there are additional flowpaths which transport nitrate to the stream more slowly.

4.3   Subsurface drain-fed streams

By far, the subsurface drain-fed streams had the highest concentrations and largest range of concentrations (Figure 4.1). The rapid movement of water through the soil to the drains and then to the stream allowed for ample interaction between the water and soil, but little time for denitrification to occur (Green et al., 2008). The concentration of nitrate remained elevated (usually >5 mg N/L) throughout most of the year and through all flow conditions in the subsurface drain-fed streams. Nitrate concentration decreased rapidly between the months of June and September in both streams. This was a result of a higher percentage of the water in the drains coming from the nitrate depleted deeper soils beneath the drains due to less water percolating vertically to the drains (Smith and Kellman, 2011), as well as plant uptake of nitrate (Tufekcioglu et al., 2003) in the upper soil layers.

The subsurface drain-fed streams had a general inverted "U-shaped" relationship between streamflow and nitrate concentration (Figure 4.3). Low nitrate concentrations occurred at low flows, largely during the fall and winter months. The highest nitrate concentrations occurred at intermediate streamflows. At the high streamflows the nitrate concentrations generally decreased, suggesting that the reservoir of nitrate stored in the soil was being washed-out during the large rainfall events (Thornburg, 2009) or the volume of low nitrate precipitation and overland flow was great enough to decrease the concentration in the stream.

The subsurface drain-fed South Fork Iowa River, Blairsburg, Iowa had the greatest loads of the six small watersheds. The Leary Weber Ditch had loads that were an order of magnitude smaller than the Iowa stream, but Leary Weber Ditch was the smallest watershed examined. When examining the cumulative load plots for the subsurface drain-fed streams, the steep slope between 20 and 80% of the cumulative load illustrates that most of the annual load in the streams occurred during intermediate streamflow conditions. The leveling off of the plot at the highest streamflows suggests

that the very high flow events were less important. About 50 and 8% of the total annual nitrate load was transported under BFI-calculated baseflow conditions in the South Fork Iowa River, Blairsburg and Leary Weber Ditch, respectively.

The average watershed yields, just as with concentrations, from the subsurface drain-fed streams were the highest yields among the six watersheds. These yields were substantially greater compared to the other watersheds (Table 4.1). The movement of water through the nitrate rich layer of top soil, in concert with the efficient removal of water from the landscape via surface and subsurface drains enhances the transport of nitrate to the stream.

The calculated LAPUs for the subsurface drain-fed watersheds were extremely high. In the South Fork Iowa River, Blairsburg during 2007, a year which the rainfall was close to normal, the LAPU was calculated as 66%, meaning the equivalent of 66% of the applied nitrogen fertilizer moved out of the watershed through the stream. In 2008 within the South Fork Iowa River, Blairsburg, a year of unusually heavy rainfall throughout the spring and early summer, the calculated LAPU was extremely high at 93%. The 2008 LAPU likely has a contribution from the nitrogen applied in 2008 and contributions from nitrogen used in previous years that was flushed from the soil through the drainage network with the large volume of water from precipitation. Thornburg (2009) found that the concentration in water discharging from the drains after these large rains was consistently depleted in nitrogen suggesting that large masses of nitrate had been removed during times of elevated drain flow.

In the large, subsurface drain-fed South Fork Iowa River, New Providence, the peak in nitrate concentration coincided with the peak in streamflow. The recession of streamflow and nitrate concentration after the peak had long, slowly decreasing tails compared to the overland flow-fed and groundwater-fed streams (Figure 4.5), but there was not a second peak in nitrate concentration as there was in the overland flow-fed Bogue Phalia, MS. In the South Fork Iowa River, New Providence basin, water moves through a combination of subsurface drain flow, overland flow, and to a lesser extent, groundwater. Both types of flowpaths transport water and nitrate quickly to the stream. The combination of these two flowpaths exposes the nitrate at the soil surface and the

nitrate in the shallow soil column (above the depth of the drain) to the moving water. Water that infiltrates into the soil is engineered to drain quickly through the subsurface drain network. The soils nearest the drains lose their water quickly to the subsurface network. Water from soils progressively further from the subsurface drain network is removed from the soil over longer periods of time. This is reflected in the long-tail recession curves for both nitrate and water. Consistent with the observations over two years in the smaller subsurface drain-fed streams, South Fork Iowa River, New Providence, IA had high nitrate concentrations.

4.4 Groundwater fed-streams where groundwater is contaminated with nitrate

The two groundwater-fed streams had similar distributions of nitrate concentration. The highest concentrations at both sites were during periods of low-flow where the nitrate-contaminated groundwater comprised all or most of the flow. Since the nitrate concentration in these streams is a reflection of the nitrate concentration in groundwater, the variability in median annual concentration among groundwater-fed stream is expected to be large. Where the nitrate concentration in groundwater is relatively low (either not contaminated or with denitrifying conditions), the concentrations of nitrate in the stream is expected to be systematically low. Where the nitrate concentration in groundwater is relatively high (contaminated and oxic conditions), the concentrations of nitrate in the stream is expected to be systematically high. However, if substantial denitrification is occurring at the groundwater-surface water interface, it is possible to have a stream with low nitrate concentration at low flows, even if the groundwater has a high nitrate concentration (Bachman et al., 2002 and Böhlke and Denver, 1995).

The two groundwater-fed streams generally had lower nitrate concentrations with higher streamflow. The strong inverse relation between nitrate concentration and streamflow in Morgan Creek, MD (in 2004) and Granger Drain (in 2003 and 2004) had an $R^2 > 0.7$. This relation is quite striking in Granger Drain where the data from the irrigated season forms a distinct cluster away from the non-irrigated season. The two distinct data clusters in Granger Drain may artificially increase the R^2 value, but the

84

relationship is strong nonetheless. In both cases, the highest stream concentrations were during low flow periods when the high-nitrate groundwater discharging to the stream comprised a very large fraction of streamflow. Water entering the streams at higher flows had less nitrate and acted to dilute the nitrate in the stream from groundwater.

The two groundwater-fed streams had intermediate annual nitrate loads, compared to the other streams. The slope of the cumulative load curve for Morgan Creek became almost linear above the baseflow line (Figure 4.4E), which straddles the transition from the BFI-calculated baseflow to non-baseflow (0.6 m^3/s). This is in agreement with the conclusions reached from the concentration data; there are distinct differences in nitrate transport to the stream by overland flow compared to groundwater. The relation between percent cumulative load and sorted discharge is quite different in Granger Drain compared to the other streams. The plot for Granger Drain shows the very different water and nitrate movement during the irrigation and non-irrigation seasons (Figure 4.4F). When the BFI program calculated value for baseflow is used, the total annual nitrate load transported under baseflow conditions is estimated to be ~100%, but also assumes that 98% of the annual discharge in from baseflow. The cumulative load figure suggests that during the non-irrigation season (streamflows <~1.0 m^3/s) approximately 28% of the annual water discharge occurred, but roughly 52% of the annual nitrate load was carried in the stream. If the groundwater nitrate concentration and discharge to the stream during the non-irrigation season are assumed to be constant throughout the entire year, then the excess irrigation water that travels to the stream during the irrigation season must have a very small nitrate load. McCarthy and Johnson (2009) used an inverse end-member mixing model to estimate the concentration of nitrate in the excess irrigation water in the DR2 Drain (a sub-basin of Granger Drain). Their estimate suggests that the concentration of nitrate in the excess irrigation water was about 70% lower than the groundwater nitrate concentration, which could lead to small nitrate loads in the excess irrigation water.

The two groundwater-fed streams had intermediate annual nitrate yields, compared to the other streams. The irrigated, groundwater-fed Granger Drain had a similar nitrate yield to the overland flow-fed Tommie Bayou, but lower than might be

expected from the concentration data. Although the median nitrate concentration in Granger Drain was higher than in Tommie Bayou, the average water yield from Granger Drain was three times less than from Tommie Bayou, resulting in much less water per hectare for transporting nitrate to Granger Drain. The natural, groundwater-fed Morgan Creek watershed had nitrate and water yields that were about double that of the irrigated, groundwater-fed Granger Drain partially due to greater annual precipitation on the Morgan Creek watershed. The high water yield from Morgan Creek, coupled with the slow rate of denitrification in the groundwater within the watershed (Green et al., 2008) has allowed the accumulation of nitrate in groundwater, which is constantly delivered to the stream throughout the year in baseflow.

In the natural groundwater-fed Morgan Creek, where the stream nitrate concentration is sustained by contaminated groundwater, the calculated annual LAPUs were 12 and 17%. However, some of the nitrate load observed in Morgan Creek was from nitrogen that was applied during previous years. The travel times of water from rainfall on the landscape to baseflow discharge in the stream range from months to years to decades (Sanford and Pope, 2013). The high concentrations of nitrate observed in the stream during baseflow were from nitrogen that was applied years to decades ago. This would give an overestimation of these LAPUs, which are based on only nitrogen use during the study years.

The high resolution streamflow and nitrate concentration data for the Chesterville Branch, MD stream (natural groundwater-fed) produced a very different nitrate response to a storm event than the two other watersheds with continuous data. Streamflow prior to and after the peak is a result of baseflow from groundwater. The nitrate concentration was highest when the stream was comprised largely of baseflow as a result of the nitrate-contaminated, oxic groundwater (Böhlke and Denver, 1995). After the onset of the precipitation event, the nitrate concentration in the stream decreased with the increase in streamflow, reached a minimum concentration just after peak streamflow, then increased again as the streamflow decreased. The increased flow in the stream was a result of water that had quickly found its way to the stream after the rainfall. Because the storm water

was moved very quickly off of the landscape in this watershed (Figure 4.5C), it contained relatively little nitrate, thus diluting the nitrate in the stream.

5. <u>Conclusions</u>

The magnitude and temporal patterns of nitrate concentrations, loads, and watershed yields varied widely for the six streams, and their watersheds. The streams, located in six different hydrologic settings, showed very different responses of nitrate. The high-temporal resolution streamflow and nitrate concentration data over a single storm event showed the very different nitrate responses in larger streams in three of the hydrologic settings. Based on these characteristic streams examined here, subsurface drain-fed streams could be expected to have the highest concentrations, largest distribution of concentrations, and highest nitrate loads when comparing similar sized watersheds. These watersheds will probably also have the greatest yields of nitrate. Streams where overland flow is the major water flowpath could be expected to receive less nitrate from the watershed, and generally have lower concentrations, loads, and watershed yields of nitrate. Groundwater-fed streams, depending on the degree of contamination of the groundwater, could have widely varying concentrations, loads, and yields. The two groundwater-fed streams included in this study had the highest stream nitrate concentrations during periods of low flow because both aquifers were contaminated with nitrate. The observed differences in nitrate between these watersheds are attributed, in part, to differences in flowpaths of water to the stream, which provides insight into the hydrologic process controlling the movement of nitrate.

References

Bachman, L.J., D.E. Krantz, and J.K. Böhlke. 2002. Hydrogeologic framework, ground-water geochemistry, and assessment of nitrogen yield from base flow in two agricultural watersheds, Kent County, Maryland. U.S. Environmental Protection Agency Report EPA/600/R-02/008, 79 p.

Bijeriego, M., R.D. Hauch, and R.A. Olson. 1979. Uptake, translocation, and utilization of ^{15}N-depleted fertilizer in irrigated corn. Soil Sci. Soc. Am. J. 43(3), 528-533. doi: 10.2136/sssaj1979.03615995004300030020x

Blann, K.L., J.L. Anderson, G.R. Sands, and B. Vondracek. 2009. Effects of Agricultural Drainage on Aquatic Ecosystems: A Review. Critical Reviews in Environmental Science and Technology 39(11), 909-1001. doi: 10.1080/10643380801977966

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 Resources Research, 31(9), 2,319-2,339. doi:10.1029/95WR01584

Capel, P.D., K.A. McCarthy, and J.E. Barbash. 2008. National, holistic, watershed-scale approach to understand the sources, transport, and fate of agricultural chemical, J. Environ. Qual., 37(3), 983–993. doi: 10.2134/jeq2007.0226

Cassell, E.A. and J.C. Clausen. 1993. Dynamic Simulation Modeling for Evaluating Water Quality Response to Agricultural BMP Implementation, Wat. Sci. Tech. 28(3-5), 635-648.

Domagalski, J.L, S. Ator, R. Coupe, K. McCarthy, D. Lampe, M. Sandstrom, and N. Baker. 2008. Comparative Study of Transport Processes on Nitrogen, Phosphorus, and Herbicides to Streams in Five Agricultural Basins, USA, J. Environ. Qual., 37(3), 1158-1169. doi: 10.2134/jeq2007.0408

Dubrovsky, N.M. and P.A. Hamilton. 2010. Nutrients in the Nation's streams and groundwater: National Findings and Implications, U.S. Geological Survey Fact Sheet 2010-3078, 6 p.

Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.M. Gronberg, P.A. Hamilton, K.J. Hitt, D.K. Mueller, M.D. Munn, B.T. Nolan, L.J. Puckett, M.G. Rupert, T.M. Short, N.E. Spahr, L.A. Sprague, and W.G. Wilber. 2010. The quality of our Nation's waters—Nutrients in the Nation's streams and groundwater, 1992–2004, U.S. Geological Survey Circular 1350, 174 p.

Green, M.B. 2007. Hydrologic Control of Stream Water N:P Ratios, PhD Thesis, Water Resources Science, University of Minnesota, St. Paul, Minnesota.

Green, C.T., L.J. Puckett, J.K. Böhlke, B.A. Bekins, S.P. Phillips, L.J. Kauffman, J.M. Denver, and H.M. Johnson. 2008. Limited Occurrence of Denitrification in Four Shallow Aquifers in Agricultural Areas of the United States, J. Environ. Qual., 37(3), 994-1009. doi: 10.2134/jeq2006.0419

Gronberg, J.M. and C.R. Kratzer. 2006. Environmental setting of the Lower Merced River Basin, California, U.S. Geological Survey Scientific Investigations Rep. 2006–5152, Reston, VA.

Hancock, T.C. and M.J. Brayton. 2006. Environmental setting of the Morgan Creek Basin, Maryland, 2002–04, U.S. Geological Survey Open-File Report 2006–1151, Reston VA.

Hooper, R.P., N. Christophersen, and N.E. Peters. 1990. Modeling streamwater chemistry as a mixture of soil water end-members—an application to the Panola Mountain Catchment, Georgia, USA, J. Hydrol., 116(1-4), 321-343. doi:10.1016/0022-1694(90)90131-G

Hubbard, L., D.W. Kolpin, S.J. Kalkhoff, and D.M. Robertson. 2011. Nutrient and sediment concentrations and corresponding loads during the historic June 2008 flooding in eastern Iowa. J. Environ. Qual., *40*(1), 166-175. doi:10.2134/jeq2010.0257

Kladivko, E. J., G.E. Van Scoyoc, E.J. Monke, K.M. Oates, and W. Pask. 1991. Pesticide and Nutrient Movement into Subsurface Tile Drains on a Silt Loam Soil in Indiana, J. Environ. Qual., 20(1), 264–270. doi: 10.2134/jeq1991.00472425002000010043x

Lathrop, T.R. 2006. Environmental setting of the Sugar Creek and Leary Weber Ditch Basins, Indiana, 2002–04, U.S. Geological Survey Scientific Investigations Report 2006–5170, Reston, VA.

McCarthy, K. A., and H.M. Johnson. 2009. Effect of agricultural practices on hydrology and water chemistry in a small irrigated catchment, Yakima River basin, Washington. U.S. Geological Survey Scientific Investigations Report 2009–5030, Reston, VA.

McCarthy, K.A., C.E. Rose, and S.J. Kalkhoff. 2012. Environmental settings of the South Fork Iowa River Basin, Iowa, and the Bogue Phalia Basin, Mississippi, 2006–10, U.S. Geological Survey Scientific Investigations Rep. 2012–5021, Reston, VA.

Meisinger, J.J., V.A. Bandel, G. Stanford, and J.O. Legg. 1985. Nitrogen utilization of corn under minimal tillage and moldboard plow tillage: I. Four year results using labeled fertilizer on an Atlantic Coastal Plain soil, Agron. J, 77(4), 602–611. doi: 10.2134/agronj1985.00021962007700040022x

Molénat, J., P. Durand, C. Gascuel-Odoux, P. Davy, and G. Gruau. 2002. Mechanisms of nitrate transfer from soils to stream in an agricultural watershed of French Brittany, Water Air Soil Pollut., 133(1-4), 161–183. doi: 10.1023/A:1012903626192

Olarewaju, O., M. Adetunji, C. Adeofun, and I. Adekunle. 2009. Nitrate and phosphorus loss from agricultural land: implications for nonpoint pollution, Nutrient Cycling in Agroecosystems. 85(1), 79-85. doi: 10.1007/s10705-009-9249-8.

Olson, R.V. 1980. Fate of tagged nitrogen fertilizer applied to irrigated corn. Soil Sci. Soc. Am. J. 44(3), 514-517. doi: 10.2136/sssaj1980.03615995004400030015x

Payne, K.L., H.M. Johnson, and R.W. Black. 2007. Environmental setting of the Granger Drain and DR2 Basins, Washington 2003–2004, U.S. Geological Survey Scientific Investigations Rep. 2007–5102, Reston, VA.

Peters, N.E. 1994. Water-quality variations in a forested Piedmont catchment, Georgia, USA. J. Hydrol. 156(1-4), 73-90. doi:10.1016/0022-1694(94)90072-8

Puckett, L.J., C. Zamora, H. Essaid, J.T. Wilson, H.M. Johnson, M.J. Brayton, and J.R. Vogel. 2008. Transport and Fate of Nitrate at the Ground-water/Surface-water Interface. J. Environ. Qual. 37(3), 1034-1050. doi: 10.2134/jeq2006.0550

Poor, C.J. and J.J. McDonnell. 2007. The effects of land use on stream nitrate dynamics. J. of Hydro. 332(1-2), 54-68. doi:10.1016/j.jhydrol.2006.06.022

Reddy, G.B. and K.R. Reddy. 1993. Fate of Nitrogen-15 Enriched Ammonium Nitrate Applied to Corn, Soil Sci. Soc. Am. J. 57(1), 111-115. doi:10.2136/sssaj1993.03615995005700010021x

Ribaudoa, M.O., R. Heimlichb, and M. Peters. 2005. Nitrogen sources and Gulf hypoxia: potential for environmental credit trading, Ecological Economics 52(2), 159-168. doi:10.1016/j.ecolecon.2004.07.021

Rice, K.C. and O.P. Bricker. 1995. Seasonal cycles of dissolved constituents in streamwater in two forested catchments in the mid-Atlantic region of the eastern USA, J. of Hydrol. 170(1-4), 137-158. doi:10.1016/0022-1694(95)92713-N

Ross, D.S., R.J. Bartlett, F.R. Magdoff, and G.J. Walsh. 1994. Flow path studies in forested watersheds of headwater tributaries of Brush Brook, Vermont. Water Resour. Res. 30(9), 2611-2618. doi:10.1029/94WR01490

Ruiz, L., S. Abiven, C. Martin, P. Durand, F. Vertes, V. Beaujouan, and J. Molenat. 2002. Effect of nitrate concentration in stream water of agricultural practices in small catchments in Brittany. II – Temporal variations and mixing processes, Hydrol.Earth Syst. Sci. 6(3), 507–513. doi:10.5194/hess-6-507-2002

Runkel, R.L., C.G. Crawford, and T.A. Cohn. 2004. Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. U.S. Geological Survey Techniques and Methods book 4, chapter A5, Reston, VA.

Sanford, W.E. and J.P. Pope. 2013. The Role of Groundwater in Delaying Chesapeake Bay Restoration, Environmental Science and Technology, in press.

Sauer, T. J., R.B. Alexander, J.V. Brahana, and R.A. Smith. 2008. The Importance and Role of Watersheds in the Transport of Nitrogen. Nitrogen in the Environment, 203.

Shelton, L.R. 1994. Field Guide for Collecting and Processing Stream-Water Samples for the National Water-Quality Assessment Program, U.S. Geological Survey Open-File Report 94-455, 50 p., Reston, VA.

Smith, E.L. and L.M. Kellman. 2011. Nitrate loading and isotopic signatures in subsurface agricultural drainage systems. J. Environ. Qual., 40(4), 1257-1265. doi:10.2134/jeq2010.0489.

Tesoriero, A.J., J.H. Duff, D.M. Wolock, N.E. Spahr, and J.E. Almendinger. 2009. Identifying Pathways and Processes Affecting Nitrate and Othophosphate Inputs to Streams in Agricultural Watersheds, J. Environ. Qual. 38(5), 1892-1900. doi: 10.2134/jeq2008.0484

Thornburg, J. 2009. Temporal and spatial variability in nitrate and water quality parameters of subsurface drains in an agricultural stream. MS Thesis. Water Resources Science. University of Minnesota, St. Paul, MN

Tufekcioglu, A., J.W. Raich, T.M. Isenhart, and R.C. Schultz. 2003. Biomass, carbon and nitrogen dynamics of multi-species riparian buffers within an agricultural watershed in Iowa, USA. *Agroforestry Systems*, *57*(3), 187-198. doi: 10.1023/A:1024898615284

U.S. Geological Survey (USGS) 2006. Collection of water samples (ver. 2.0): U.S. Geological Survey Techniques of Water-Resources Investigations. Book 9, chap. A4. Accessed January 21, 2011 at http://water.usgs.gov/owq/FieldManual/chapter4/html/Ch4_contents.html.

U.S. Geological Survey (USGS) 2013a. Continuous Nitrate Monitoring in Rivers. Accessed October 24, 2013 at http://water.usgs.gov/nasqan/docs/RT-BriefingSheet.pdf

U.S. Geological Survey (USGS) 2013d. USGS Water Data for the Nation. Accessed October 24, 2013 at http://waterdata.usgs.gov/usa/nwis/.

Vidon, P., H. Hubbard, P. Cuadra, and M. Hennessy. 2012. Storm NO_3^- and NH_4^+ Exports in Stream, Overland Flow, and Tile Drains of the US Midwest. Annals of Environmental Science, 6(1), 5.

Wahl, K.L. and T.L. Wahl. 1995. Determining the Flow of Comal Springs at New Braunfels, Texas. *Texas Water '95*. American Society of Civil Engineers. August 16-17, 1995. San Antonio, Texas : 77-86.

Wahl, T.L. and K.L. Wahl. 2007. BFI: A Computer Program for Determining an Index to Base Flow (Version 4.15), [Software]. http://www.usbr.gov/pmts/hydraulics_lab/twahl/bfi/

Wang, T. and B. Zhu. 2011. Nitrate loss via overland flow and interflow from a sloped farmland in the hilly area of purple soil, China, Nut. Cyc. In Agroecosystems, 90(3), 309-319. doi: 10.1007/s10705-011-9431-7

Wilde, F.D., D.B. Radtke, J. Gibs, and R.T. Iwatsubo. 2004 with updates through 2009. Processing of water samples (version 2.2). U.S. Geological Survey Techniques of Water-Resources Investigations. Book 9, chap. A5, Accessed February 13, 2011, at http://pubs.water.usgs.gov/twri9A5/.

Table 4.1: Watershed information with annual water and nitrate concentration, load, and yield data for six agricultural streams during the sampled years. Included are data for three additional streams with high temporal resolution streamflow and concentration data. The table was split into two sections to decrease overall table width.

Watershed information

State (USGS Site ID)	Study area	Watershed area (km²)	Hydrologic setting	Water year	Percent in agriculture	Precipitation (cm)	Nitrogen fertilizer use (kg N)	Nitrogen fertilizer application rate (kg N/ha)
Mississippi (07288636)	Tommie Bayou[1]	15.3	Overland flow (perennial)	2007"	86	130	28,243	18
	Tommie Bayou[1]			2008"	86	145	18,440	12
California (373112120382901)	Mustang Creek[2]	17.5	Overland flow (ephemeral)	2003	97	No data	209,078	120
	Mustang Creek[2]			2004	97	27	209,078	120
Iowa (05451080)	SFIR, BB[+1]	31.1	Subsurface drainage (w/confining layer)	2007	96	114	264,000	85
	SFIR, BB[+1]			2008	96	124	229,000	74
Indiana (03361638)	Leary Weber Ditch[3]	7.23	Subsurface drainage (wo/confining layer)	2003"	93	133	48,000	66
	Leary Weber Ditch[3]			2004	93	124	48,000	66
Maryland (01493500)	Morgan Creek[4]	32.9	Groundwater (natural)	2003	85	162	265,000	81
	Morgan Creek[4]			2004	85	100	226,000	69
Washington (12505450)	Granger Drain[5]	161	Groundwater (irrigated)	2003	96	No data	2,248,110	140
	Granger Drain[5]			2004	96	22	2,248,110	140
Mississippi (07288650)	Bogue Phalia[1]	1250	Overland flow (perennial)					
Iowa (05451210)	SFIR, NP[#1]	580	Subsurface drainage (Wo/confining layer)					
Maryland (01493112)	Chesterville Branch	15.9	Groundwater (natural)					

	Water						Concentration			Load and Yield		
Study area	Mean streamflow (m³/s)	Median streamflow (m³/s)	Annual water discharge (m³/yr)	Annual baseflow discharge (m³/yr)	Percent of annual discharge from baseflow (%)	Annual water yield (m³/ha/yr)	Mean concentration (mg N/L)	Median concentration (mg N/L)	# of samples	Annual stream load (kg N)	Nitrate load as a percentage of fertilizer use (LAPU)	Annual watershed yield (kg N/ha/yr)
Tommie Bayou[1]	0.2	0.1	6,654,120	998,118	15	4,340	1.2	0.52	14	No data	No data	No data
Tommie Bayou[1]	0.4	0.1	13,644,870	1,773,833	13	8,900	0.6	0.35	39	6,600	36	4.3
Mustang Creek[2]	< 0.1	< 0.1	29,835	No data	No data	17	0.7	0.67	3.0	No data	No data	No data
Mustang Creek[2]	< 0.1	< 0.1	200,371	No data	No data	114	1.6	1.7	6.0	230	0.11	0.13
SFIR, BB[+1]	0.5	0.2	15,252,466	6,939,872	46	4,910	14.9	14	35	240,000	66	77
SFIR, BB[+1]	0.6	0.2	18,430,626	8,699,255	47	5,930	10.5	11	18	213,000	93	68
Leary Weber Ditch[3]	0.1	< 0.1	3,670,000	367,000	10	5,080	7.6	7.8	16	25,000	53	35
Leary Weber Ditch[3]	0.1	< 0.1	2,700,000	604,000	22	3,730	4.7	4.5	23	11,000	24	16
Morgan Creek[4]	0.5	0.3	16,409,885	6,875,742	42	4,990	2.1	2.1	27	32,000	12	10
Morgan Creek[4]	0.5	0.3	14,238,881	8,059,206	57	4,330	2.5	2.5	22	39,000	17	12
Granger Drain[5]	1.0	1.0	30,587,087	29,210,669	96	1,910	3.5	2.4	22	100,000	4.4	6.2
Granger Drain[5]	1.0	1.0	31,091,082	30,375,987	98	1,940	3.5	2.7	22	No data	No data	No data
Bogue Phalia[1]	21											
SFIR, NP[#1]	5.2											
Chesterville Branch	0.21											

\# South Fork Iowa River northeast of New Providence, IA   + South Fork Iowa River near Blairsburg, IA

" Calendar year used for nitrate load calculation   1 McCarthy et al. (2012)

2 Gronberg and Kratzer (2006)   3 Lathrop (2006)

4 Hancock and Brayton (2006)   5 Payne et al. (2007)
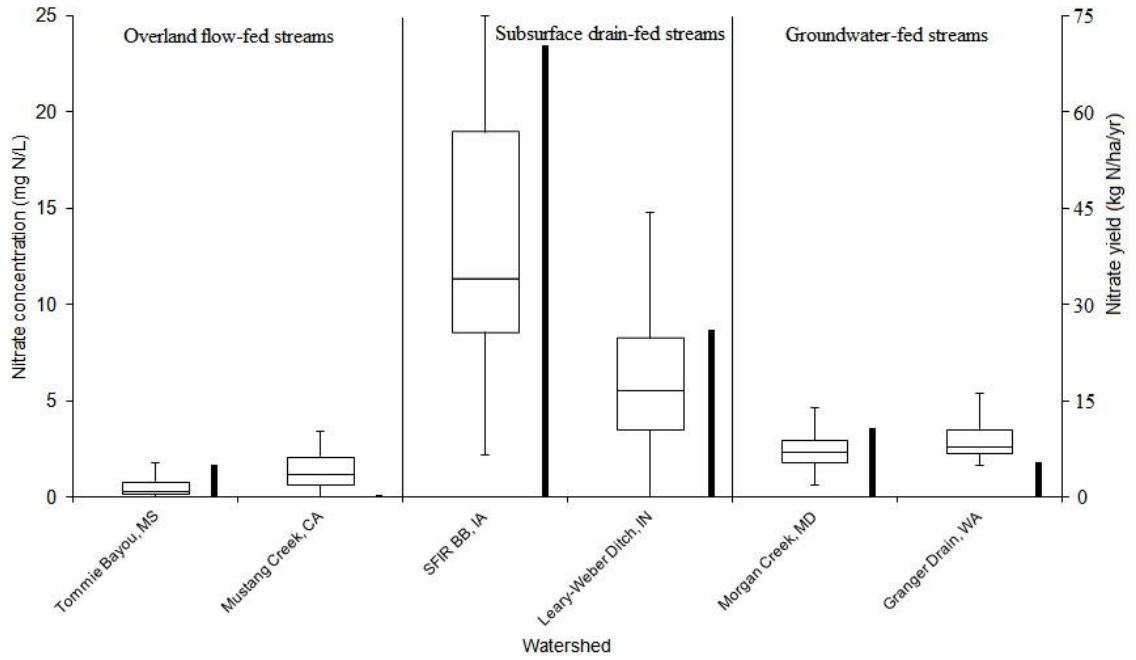
Figure 4.1: Boxplots of two years of nitrate concentrations in stream water for the six watersheds. Black bars to the right of boxplots represent annual nitrate yield from the watershed. Outliers were removed and were defined as data points greater than 1.5 times the inter quartile range above the 75th percentile or data points less than 1.5 times the inter quartile range below the 25th percentile. SFIR BB, IA is South Fork Iowa River, Blairsburg, IA

(A) Perennial, overland flow-fed
   Tommie Bayou, MS
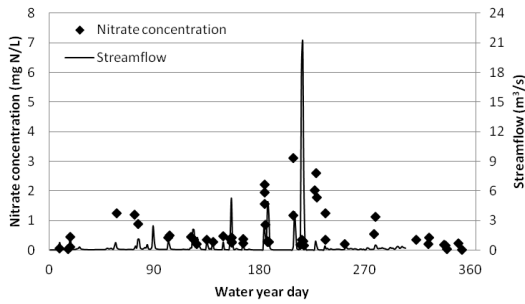
(B) Ephemeral, overland flow-fed
   Mustang Creek, CA
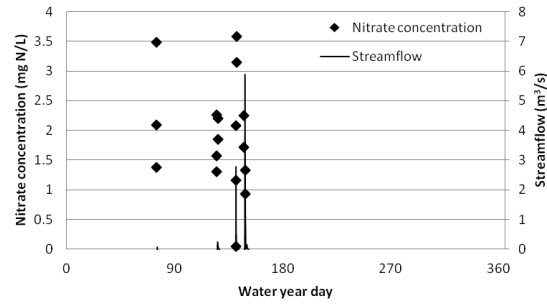
(C) Subsurface drain-fed
   South Fork Iowa River, Blairsburg, IA
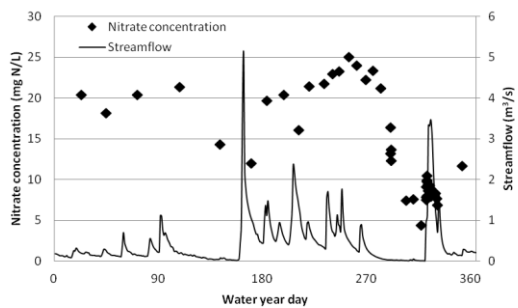
(D) Subsurface drain-fed
   Leary Weber Ditch, IN
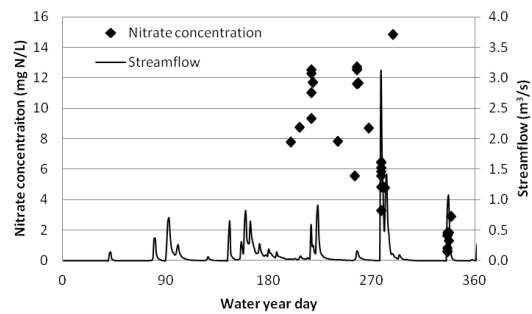
(E) Natural groundwater-fed
   Morgan Creek, MD
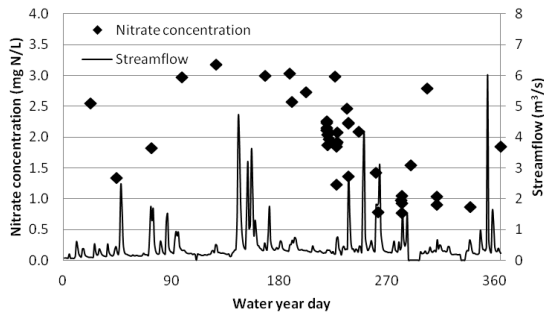
(F) Irrigated groundwater-fed
   Granger Drain, WA
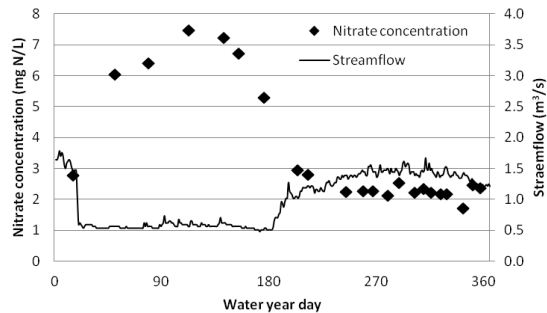
Figure 4.2: Nitrate concentration and streamflow through the 2003 or 2007 water year (2004 in Mustang Creek, CA and 2008 in Tommie Bayou, MS). Day 0 corresponds to October 1, which is the first day of the water year. Note the differences in Y-axis scales.
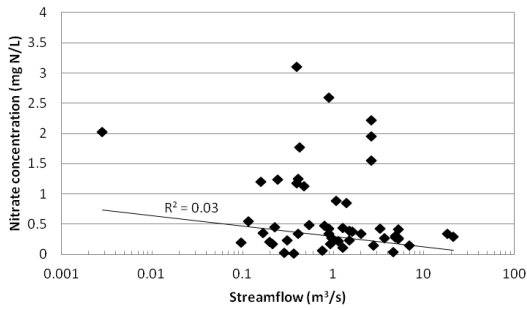
95

(A) Perennial, overland flow-fed
   Tommie Bayou, MS

(B) Ephemeral, overland flow-fed
   Mustang Creek, CA
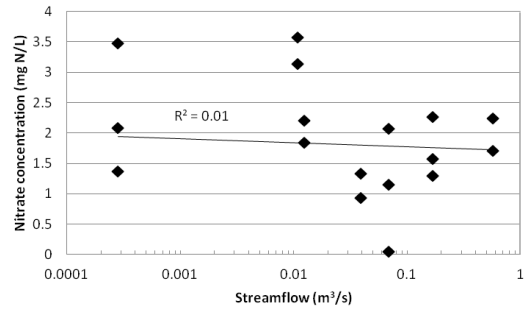
(C) Subsurface drain-fed
   South Fork Iowa River, Blairsburg, IA
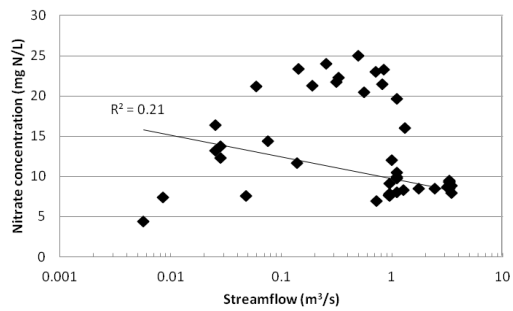
(D) Subsurface drain-fed
   Leary Weber Ditch, IN
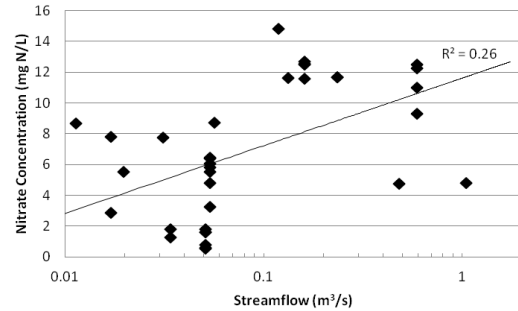
(E) Natural groundwater-fed
   Morgan Creek, MD
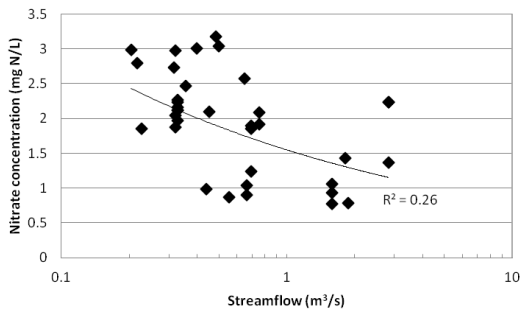
(F) Irrigated, groundwater-fed
   Granger Drain, WA
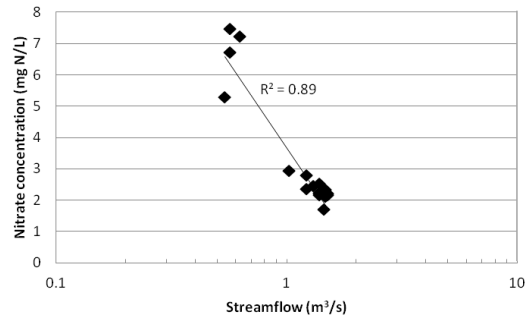
Figure 4.3: Relationship between nitrate concentration and streamflow through the 2003 or 2007 water year (2004 in Mustang Creek, CA and 2008 in Tommie Bayou, MS). Note the differences in X-axis and Y-axis scales.
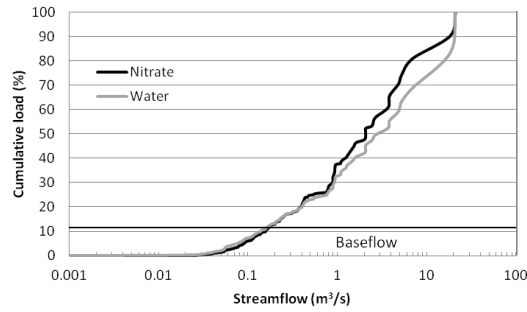
96

(A) Perennial, overland flow-fed
     Tommie Bayou, MS

(B) Ephemeral, overland flow-fed
     Mustang Creek, CA

(C) Subsurface drainage-fed
     South Fork Iowa River, Blairsburg, IA
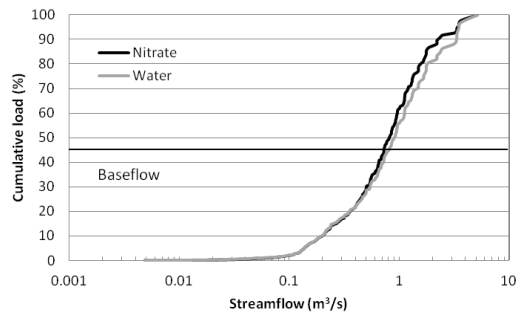
(D) Subsurface drain-fed
     Leary Weber Ditch, IN
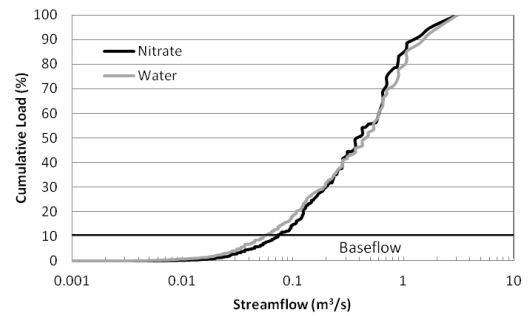
(E) Natural groundwater-fed
     Morgan Creek, MD
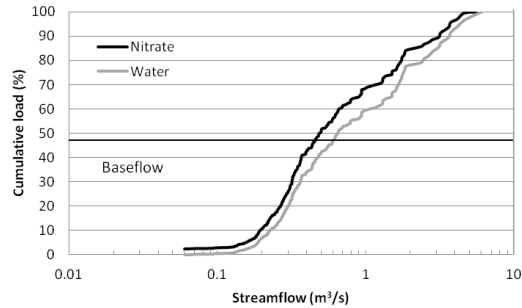
(F) Irrigated, groundwater-fed
     Granger Drain, WA
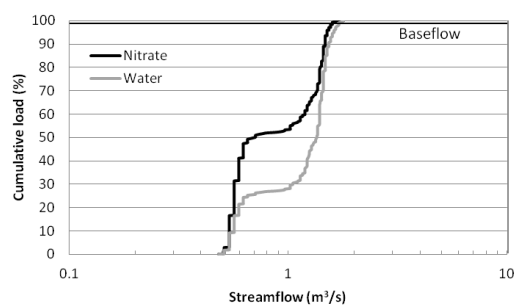
Figure 4.4: Cumulative load of nitrate in relation to streamflow through the 2003 or 2007 water year (2008 in Tommie Bayou, MS). The horizontal line represents the percentage of total annual streamflow that is from baseflow. The intersection of the horizontal baseflow line and the cumulative load curve is the transition from baseflow to non-baseflow as the source nitrate and water to the stream. Mustang Creek was excluded because streamflow occurred on a small number of days, such that a calculation of cumulative load was not appropriate.

97

(A) Perennial, overland flow-fed Bogue Phalia, MS
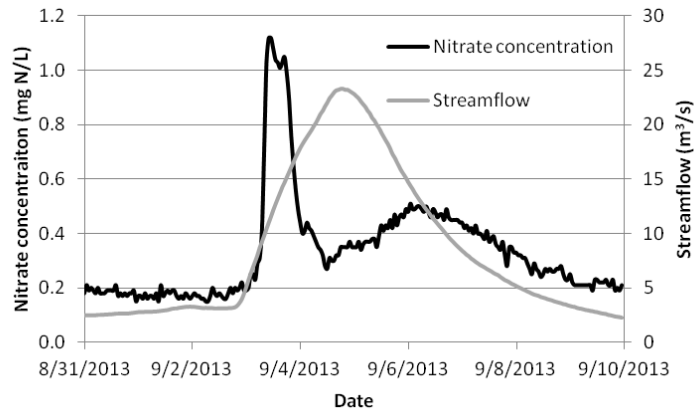
(B) Subsurface drainage-fed South Fork Iowa River, New Providence, IA
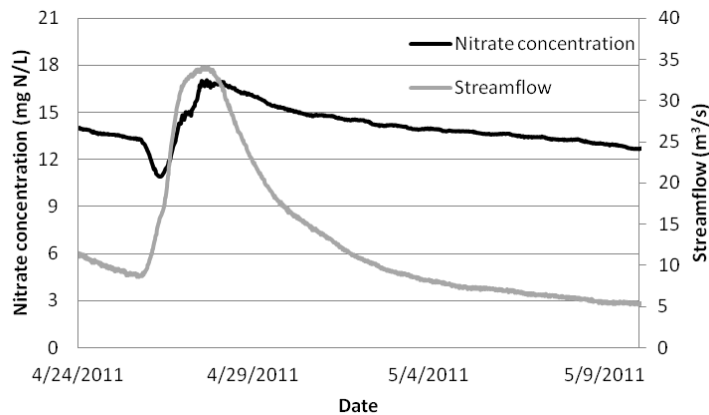
(C) Natural groundwater-fed Chesterville Branch, MD
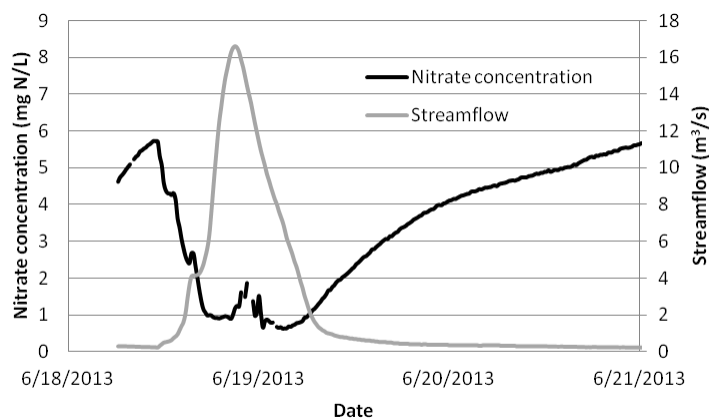
Figure 4.5: Nitrate concentration and streamflow for a single storm event in the three larger streams. The perennial overland flow-fed Tommie Bayou is a tributary to Bogue Phalia, the subsurface drain-fed South Fork Iowa River, Blairsburg is a tributary to South Fork Iowa River, New Providence, and the natural groundwater-fed Morgan Creek is adjacent to Chesterville Branch.

Chapter 5: Fundamental watershed factors influencing the transport of nitrogen to streams

Abstract

Accurate estimation of total nitrogen loads, yields, and concentrations is essential for evaluating conditions in the aquatic environment. The ability to estimate those values within unstudied watersheds is greatly beneficial. Recursive partitioning and random forest regression were used to assess 85 geospatial, environmental, and watershed variables across 636 small (<585 km^2) watersheds to determine which variables have an important influence on the annual load and flow-weighted mean concentration in streams, and annual watershed yield of total nitrogen. Initial analysis led to the splitting of watersheds into three groups based on dominant land use (agricultural, developed, and undeveloped). Nitrogen application, agricultural land area, forested canopy area, and impervious buffer area were some of the most frequently extracted variables by both recursive partitioning and random forest regression. A series of multiple linear regression equations utilizing the extracted variables were developed and applied to the watersheds. As few as three variables explained as much as 76% of the variability in total nitrogen loads for watersheds with predominantly agricultural land use. Estimates of nitrogen in agricultural streams were the most accurate when compared against stream measurements, whereas estimates for undeveloped streams were often the least accurate. Gaining as much information with the fewest number of variables will allow for total nitrogen in streams to be estimated without extensive knowledge of a watershed or use of more complex modeling techniques. The estimates provided by models which use fundamentally important variables can inform and isolate areas where more in-depth research may be beneficial.

1. Introduction

The movement of nitrogen from the landscape to streams plays a large role in the global nitrogen cycle (Howarth et al., 1996, Vitousek et al., 1997). This waterborne nitrogen comes from natural and anthropogenic sources. Plant senescence and abscission, atmospheric deposition, and fertilizer and manure application all provide

reactive forms of nitrogen to the landscape surface and eventually to water resources. Nitrogen is a necessary macronutrient that is essential for all life. However, elevated nitrogen concentrations have wide ranging consequences, particularly in water bodies. Elevated nitrogen concentration can cause surface water contamination (Dubrovsky et al., 2010) as well as hypoxia and ecosystem degradation at, or near the source of the nitrogen addition, as well as in downstream ecosystems (Alexander et al., 2000, Ribaudoa et al., 2005).

Nitrogen is present in many forms in water. Movement of organic nitrogen, nitrate, nitrite, and ammonia - collectively referred to as total nitrogen - is strongly linked to the movement of water. Nitrate is particularly soluble and typically makes up >50% of total nitrogen (Johnes, 1996). Because these forms of nitrogen are easily dissolved or transported by water, natural or anthropogenic changes to the amount of water that reaches the land surface, or changes to the pathways through which that water travels can alter the flux of total nitrogen to streams.

Determining the environmental and watershed variables that influence the amount of nitrogen that has moved to streams has been the focus of research for decades (Benning, 2013, Billy et al., 2013, Hively et al., 2011, Mueller et al., 1997, Oehler et al. 2011, Onderka et al., 2012, Peterson and Preston et al., 2011, Reynolds and Edwards, 1995, Schilling and Libra, 2000, Stålnacke et al., 2009). Most studies, however, include comparatively few watersheds or examine watersheds from a small region or of a specific land use type. The particular variables determined to be important for estimating nitrogen in streams by one study often only apply to the watersheds within that study (Billy et al., 2013, Onderka et al., 2012, Stålnacke et al., 2009).

This study examines 636 streams with reported annual total nitrogen loads that are distributed widely across the contiguous United States in an effort to identify fundamental geospatial, environmental, and watershed variables that influence the amount of nitrogen in streams. The associated watersheds (<585 km^2) encompass a wide range of physical, climate, and land use characteristics. Data mining was accomplished with the use of recursive partitioning and random forest regression. Ninety different variables were tested as possible factors that play a role in controlling annual total

nitrogen loads and flow-weighted mean concentrations in the streams, as well as annual total nitrogen yields from the watersheds (henceforth referred to as loads, concentrations, or yields). Particular attention was paid to the importance of flowpath related variables, those variables that directly relate to the route through which water travels to the stream (baseflow, runoff, artificial drainage, etc.). Because nitrogen movement is largely driven by water movement, the important water flowpath(s) should play a large role in determining loads, concentrations, and yields. This research highlights the key variables that are applicable across a wide variety of watersheds and provides a better understanding of the variables that have a disproportionately large effect on nitrogen transport, influencing nitrogen loads, yields, and concentrations to a much greater extent than other variables. One of the outcomes of the study is a series of multiple linear regression models which utilize the variables extracted by recursive partitioning and random forest regression. These models were used for estimating the annual amount of nitrogen in streams and served as a test for the application of a small number of important, highly influential variables across a wide variety of watersheds.

2. Methods

2.1   Watersheds and their characteristics

This study focused on small streams with watersheds that range in size from 5 to 585 km^2. Larger streams were excluded in an effort to reduce the heterogeneity of variables within the watersheds. Data for 85 geospatial, environmental, and watershed variables were collected for 834 watersheds across the contiguous United States for which there were estimates of annual loads. Watersheds were removed from the analyses if data were missing for any of the ninety variables. Of the 834 starting watersheds, 509 met all criteria (Figure 5.1). Although missing some data, an additional 127 watersheds had all necessary data to be used as validation watershed later in the research. A full list of watersheds and their descriptions is available in Appendix E Table E5.1.

The nitrogen data used for this study was processed in support of seven SPARROW (SPAtially Referenced Regressions On Watershed attributes) models (Brown et al., 2011, Hoos and McMahon, 2009, Moore et al., 2011, Rebich et al., 2011,

Robertson and Saad, 2011, Saleh and Domagalski, 2012, Wise and Johnson, 2011).
Many years (~30 years) of streamflow data and stream concentration data were assessed
to create nitrogen load estimates for each stream. All load estimates were then
temporally de-trended within each watershed and standardized to a single year centered
around 2002. De-trending in this manner applies the physical state of the environment
during 2002 to each watershed, allowing for equal comparison of watersheds where data
may have been collected during different time periods (Schwarz et al., 2006). This form
of standardization makes the load estimates well-suited for spatial comparisons across a
wide area (Schwarz et al., 2006). Annual yield was calculated by dividing the load by
watershed area. Annual, flow-weighted mean concentration was calculated as the annual
load divided by annual stream discharge.

Physical and hydrological watershed variables, land use information, soil and
climate data, background nitrogen concentrations, and anthropogenic chemical use data
were obtained from a variety of sources (Table 5.1 and Appendix E Table E5.1). These
variables were separated into variable groups (chemical, land use, hydrologic,
geographic, soil, flowpath, climate, and complex) based on the type of information
provided by the variable. Some variables incorporate multiple factors into a single
variable, such as nutrient ecoregion and hydrologic landscape region. These are termed
complex variables. For example, nutrient ecoregion is considered a complex variable
because it incorporates a large amount of intrinsic information (geology, physiography,
vegetation, climate, soils, land use, wildlife, and hydrology) within a single classifier.
Data for most of the 85 variables were long term-average values, which worked well with
the long-term, standardized load estimates. For some, lack of data availability prevented
the use of long-term averages for all variables, such as the application rate of agricultural
chemicals and fertilizers. Data for these variables were from a single year as close to
2002 as possible (Table 5.1 and Appendix E Table E5.1). Watersheds were assumed to
be in a steady state based on conditions in 2002, but the technique used in this study can
be applied to other, or updated data sets.

Land use in each watershed was classified according to the area of agriculture
(based on the national land cover database - NLCD 81 and 82) or developed (NLCD 21,

22, 23, and 24) land present in the watershed. Watersheds with >25% developed land were classified as developed and watersheds with >25% agricultural land were classified as agricultural. If watersheds had both >25% developed and >25% agricultural land, they were classified as the land use that had the greater percent. All watersheds that were not classified as agricultural or developed were classified as undeveloped. Classifying watersheds this way led to three groups which each contained >100 watersheds. The largest group (undeveloped) contained 221 watersheds and the smallest group (developed) contained 114 watersheds. Groups were determined to have statistically different mean loads, yields, and concentrations by the rank sum test. The specific rank sum test used was based on the White modification to the Wilcoxon rank sum test and can be found in Ambrose et al. (2002). Two other methods of grouping watersheds were also tested and are explained in Appendix E.

Baseflow index estimates were obtained by two methods: the BFI program (Wahl and Wahl, 1995) (BFI-WAHL), and the National Hydrography Dataset (USGS, 2014) (BFI-NHD). Due to differences in how each method estimated baseflow index, results from the BFI program and the NHD were often different, although were positively correlated (r = 0.83). Both estimates were used in the analysis to determine if either estimate of baseflow index is an important, influential variable when estimating nitrogen in streams.

2.2   Statistical methods

2.2.1   Important variable extraction

The "rpart" (Therneau et al., 2013) and "randomForest" (Liaw and Wiener, 2012) packages in the statistical software program R (Version 3.0.2) (R Core Team, 2009) were used in the determination of the important variables for estimating the loads, yields, and concentrations. These techniques do not require data to be normalized prior to use which makes them well suited for analysis of these watershed data. The rpart package uses recursive partitioning (henceforth referred to as RPART) to generate a single regression tree, and uses a complexity parameter to prune the tree in an effort to reduce over-fitting of the data (Therneau et al., 2013). The randomForest package performs random forest regression (henceforth referred to as RFR) by using an ensemble learning method which

generates a series of bootstrapped regression trees, with each tree generated using a random subset of variables. Results based on RFR are produced by averaging the results of all generated trees (Liaw and Wiener, 2012). RFR has the tendency to select categorical variables with a greater number of categories and also to select variables that are highly correlated (Strobl et al., 2008). The two methods chosen for variable selection, RPART and RFR, analyzed the same data and the results were compared. Each method operates differently, but each method is a valid way of selecting important variables. The comparison of the results helps to reinforce the importance of variables that appear in both analyses.

For the analysis of watersheds using RPART within R, the method was set to "anova" to produce a regression tree. The complexity parameter was set at 0.01 as suggested by the model developers. All other arguments were left as default. The variables that appeared in the regression tree were considered important. An example of one of the regression trees is available in the Appendix E Figure E5.2. For RFR within R, 1000 regression trees were generated for each analysis. To ensure reproducibility when using R, set.seed(1) was used when generating all results with RFR. Variable importance is quantified with RFR by comparing the percent increase in mean squared error (%IncMSE) of dependant variable predictions before and after randomly permuting all of the values within each independent variable. A large %IncMSE after permuting the data of an independent variable suggested that the variable was important for estimating dependant variable values. Independent variables were ordered based on %IncMSE values, and were plotted from greatest increase to least. An example of one of these figures is available in the Appendix E Figure E5.3. Important variables were selected from the greatest %IncMSE to a visual cutoff point on the plot. The cutoff point occurred where the change in %IncMSE became relatively small and fairly constant from one variable to the next on the ranked order plot, signified by a noticeable or abrupt change in slope. The remaining variables were eliminated as not important.

Separate analyses were conducted for each of the three dependent variables (load, yield, and concentration) for each group of watersheds based on dominant land use type (undeveloped, developed, and agricultural). This resulted in 9 analyses for both RPART

and RFR, for a total of 18 separate calculations (Figure 5.2). Analyses of loads, yields, and concentrations were conducted using the same set of variables. However, all area dependent variables that were used during the load analyses were converted to area normalized variables for use in the yield and concentration analyses by dividing the values by watershed area (or buffer area for variables relating to the 100 m buffer) to get values that were independent of area (per km^2) (Table 5.1 and Appendix E Table E5.1). Yields and concentrations (dependent variables) were assumed independent of watershed area due to low Pearson's correlation (r) between watershed areas and yields (-0.03) and between watershed areas and concentrations (0.07).

2.2.2   Multiple linear regression equations

The final important variables obtained by RPART and RFR from the 18 analyses were used to create a series of multiple linear regression models. These multiple linear regression models were used to test the ability of the variables for estimating annual loads, yields, and concentrations. Although RPART and RFR do not require the assumption that data are normally distributed, multiple linear regression does. Data were tested for normality by measuring the skewness of the data for each variable. Log10-transformation was used if variables had a skewness >1.0 as suggested by Bulmer (2012). If the logarithmic transformation decreased the skewness of the variable data, the transformed data were used in the creation of the multiple linear regression equations, otherwise the non-transformed data were used. The dependent variables (load, yield, and concentration) were also log10-transformed. Variables were removed from the multiple linear regression models if they had a variance inflation factor (VIF) >5 as suggested by Rogerson (2001). This eliminated major collinearity among the explanatory variables within each linear model. R^2 was calculated for each model to gain insight into how much variability in the dependent variable could be explained when the variables extracted by RPART or RFR were applied to linear regression models. Root mean squared error normalized to the mean of the measured loads, yields, or concentrations (often referred to as the coefficient of variation of the root mean squared error) were also calculated between the measured values and those estimated by the multiple linear regression equations. Models that resulted in a normalized root mean squared error

(NRMSE) <0.4 were classified as good fitting models, those models with a NRMSE between 0.4 and 0.9 were classified as satisfactory models, and models with NRMSE > 0.9 were classified as poorly fit models as suggested by Escurra et al. (2014). Plots of the residual values are presented as Figures E5.4-E5.6 in Appendix E.

After the important variables were extracted from the 509 watersheds, an additional 127 watersheds had data for all of the extracted important variables. The 509 watersheds used for the determination of important variables were then used as the calibration watersheds for the linear regression models, and the 127 additional watersheds were then able to be used for model validation. These validation watersheds were distributed proportionally among each land use group, resulting in approximately 80% of the watersheds in each group for model calibration and 20% for model validation.

3. Results

3.1 Formation of watershed groups based on land use

Prior to splitting watersheds into dominant land use groups, initial analyses for finding important variables that influence loads, yields, and concentrations were completed using all 509 calibration watersheds. One or more of the NLCD land use codes appeared as the root node of all decision trees created by RPART (for loads, yields, and concentrations) and was always the most important or second most important variable when using RFR. This made it evident that the predominant land use is very important in determining annual loads, yields, and concentrations. These preliminary results suggested that splitting the watersheds into separate groups based on dominant land use type would help to pull out the other important variables in each subset of watersheds.

The three groups of watersheds (undeveloped, developed, and agricultural) had significantly different mean values (rank sum test, $\alpha = 0.05$) from one another with respect to loads, yields, and concentrations. In general, the agricultural watersheds had the highest loads, yields, and concentrations, while the undeveloped watersheds had the lowest (Figure 5.3A-C). A median of 52% of the land area in the agricultural watersheds was devoted to agriculture, a median of 61% of the land area was developed in the

predominantly developed watersheds, and a median of 87% of the land in undeveloped watersheds was undeveloped. This suggests that developed lands and agricultural lands have a disproportionate affect on loads, yields, and concentrations, influencing nitrogen in streams to a larger extent than other variables.

3.2　Total nitrogen loads

There were some important variables for estimating loads that were extracted for undeveloped, developed, and agricultural streams. Nitrogen application and many of the predominant land use variables were extracted (Table 5.2), indicating their general importance across all streams. The mass of nitrogen applied to agricultural land within the watershed was important in agricultural and undeveloped watersheds when using either RPART or RFR. This implies that even if <25% of the land area within the watershed is agricultural; the nitrogen applied to that land has a disproportionately large effect on the loads of nitrogen in the stream. The relative area of agricultural and/or developed land present in the watershed remained an important variable for determining loads even though watersheds were grouped according to dominant land use. In particular, the amount of developed land in developed watersheds, and the amount of agricultural land in agricultural watersheds were important. Both agricultural and developed land use were also extracted with RFR as important in undeveloped watersheds, reinforcing the idea that agricultural and developed land have a disproportionate affect on loads.

3.2.1　Agricultural watersheds

The agricultural watersheds resulted in the fewest important variables extracted by RPART or RFR when compared to the developed and undeveloped watersheds (Table 5.2). Each method resulted in four important variables. All but one important variable from both the RPART and RFR methods was either a chemical variable or land use variable. Nitrogen application on agricultural land and agricultural land area were extracted as important variables by both methods. As nitrogen application and total area of agriculture increase, loads tended to increase.

3.2.2   Developed watersheds

RPART and RFR extracted five and six important variables, respectively, when analyzing the developed watersheds. Each method had multiple land use variables, but results from RPART also included three flowpath related variables as being important when estimating loads (Table 5.2). Total developed area and impervious area within the 100 meter buffer around streams were listed as important variables by both RPART and RFR. Although the watersheds were already classified as developed, the total developed area in the watershed continued to be an important variable in determining the loads in the streams. And, in these watersheds, the percent of imperviousness within the buffer is likely a surrogate or indicator of developed land within the watersheds. Developed watersheds in this study had an average of 12.3 ($\pm$9.4)% impervious buffer area, whereas undeveloped and agricultural watersheds had a combined average of 1.2 ($\pm$1.4)%.

3.2.3   Undeveloped watersheds

The undeveloped watersheds had a greater number of variables extracted by RPART and RFR, as compared to developed and agricultural watersheds (Table 5.2). Chemical, land use, and hydrologic variables were common between RPART and RFR, with RPART including two soil variables as well. Both methods suggest that nitrogen applied on the agricultural lands of the predominantly undeveloped watersheds played an important role in determining loads. Forested canopy area and average streamflow also were extracted as important variables by both RPART and RFR. Greater areas of forested canopy possibly lead to large inputs of organic nitrogen to the stream. Forested canopy cover often has a high amount of herbaceous litter (Asner et al., 2003) which has a high nitrogen concentration (Reich et al., 2001), leading to higher levels of organic nitrogen that could have been delivered to streams, as compared to non-forested, undeveloped areas. Higher streamflows likely resulted in larger loads due to the greater potential for nitrogen movement to the streams (Craig et al., 2008).

3.3   Total nitrogen yields

Many variables relating to yields were important across land use groups, such as nitrogen application rate, water yield from the watershed, and developed and impervious land area. Nitrogen applied to agricultural land was important in the agricultural

watersheds as well as the undeveloped watersheds, indicating the effect that small agricultural areas can have on yields in mainly undeveloped watersheds (Table 5.2). Water yields from watersheds were an important variable in undeveloped and developed watersheds. Larger volumes of water travelling to a stream provide an opportunity for a greater amount of nitrogen to be moved to the stream. The percent of impervious land and the percent of developed land within the 100 meter buffer surrounding streams appeared as important variables in undeveloped watersheds. Both variables have a positive relationship with the amount of developed land within the predominantly undeveloped watersheds (r = 0.62 and 0.59, respectively). Many of these important variables again indicate a disproportionate effect that developed and agricultural land can have on the amount of nitrogen in streams.

3.3.1 Agricultural watersheds

Analysis of yields in the agricultural watersheds resulted in equal or fewer variables than were extracted for the developed watersheds and consistently fewer variables than were extracted for the undeveloped watersheds (Table 5.2). RPART and RFR extracted multiple types of variables, however variables extracted by RPART included variables from three additional variable types (hydrologic, geographic, and flowpath variables). Both methods did extract at least one chemical and one land use variable. Nitrogen application rate on agricultural land was the only variable shared among the two methods in the agricultural watersheds. Higher rates of nitrogen application lead to greater amounts of nitrogen that could be moved from the landscape to streams, which can increase nitrogen yields.

3.3.2 Developed watersheds

Variables that were extracted as important for determining yields in developed watersheds ranged widely across many types of variables. Overlap between variable types extracted by RPART and RFR occurred and included land use, complex, and soil related variables (Table 5.2). Percent forested canopy cover in the watershed and the soil K-factor (soil erodibility from the revised universal soil loss equation) were two of the variables that were extracted as being important by both RPART and RFR. As opposed to load analyses, forested canopy had a negative correlation with respect to yields. This

mitigating effect could be due to nitrogen uptake by the trees, or simply because an increase in forested area results in a smaller area of developed land, which is likely a greater contributor of nitrogen to streams as evident by the greater mean yield (Figure 5.3B). K-factor, as well as the other important soil variables that were extracted, relate to the texture and permeability of the soil. These variables could be serving as surrogates for the amount of interaction between the soil and water, however this is speculation.

3.3.3 *Undeveloped watersheds*

Undeveloped watersheds had the greatest number of variables extracted by the RPART and RFR analyses, as compared to developed and agricultural watersheds when analyzing yields (Table 5.2). The important variables were distributed across many variable types. Variable types that contained important variables extracted by both RPART and RFR were chemical, land use, and hydrologic. Agricultural and developed land use variables were particularly prevalent. Nitrogen application rate, percent imperviousness in the 100 meter buffer, and percent developed land within the buffer (both related to development), and water yields were extracted as important variables by both RPART and RFR. These variables show, as with the analyses of loads, that the area of agricultural or developed land in predominantly undeveloped watersheds plays a large role in determining the amount of nitrogen in streams.

3.4 Total nitrogen concentrations

Several variables were commonly extracted as important when analyzing concentrations across the three watershed groups, such as population density, forested canopy, and ecoregion. Nitrogen application on agricultural land was important in the predominantly agricultural watersheds (Table 5.2). The increase in source strength of nitrogen caused by applications of nitrogen can cause higher concentrations in the stream. Population density was important in the undeveloped and agricultural watersheds, as it is related to the amount of developed land within the watershed. Population density had a Pearson's r of 0.85 against the percent of developed land in the undeveloped watersheds and 0.83 in the agricultural watersheds. Forested canopy area was extracted as an important variable in all watershed groups by both RPART and RFR methods. The greater amount of forested canopy area equated to less agricultural and developed land in

the three watershed groups. And, because agricultural and developed land uses seem to have a disproportionate affect on the amount of nitrogen in streams, the increased canopy cover prevented the increases in nitrogen concentrations associated with agricultural and developed land uses. Forested canopy may also act as a "cover crop", resulting in a mitigating effect on nitrogen concentration in the stream as the trees and other plants sequester nitrogen before it reaches the streams (Swank et al., 2014). The nutrient ecoregion classification was also important across all watershed groups. Watersheds within a single nutrient ecoregion share similar geographic, physical, climate, and land use factors (Omernik, 1987) which may have led to similarities in nitrogen concentrations.

3.4.1   Agricultural watersheds

The RPART and RFR methods had 6 and 5 variables, respectively, for the agricultural watersheds. Those variables covered many different variable types with overlap of chemical, land use, and complex variables (Table 5.2). However, both RPART and RFR extracted nitrogen application rate on agricultural land, predominant nutrient ecoregion, and percent forested canopy cover within the watershed as important variables. The interactions between these variables and concentrations have been explained previously.

3.4.2   Developed watersheds

In general, the fewest number of important variables were extracted by RPART and RFR for concentration estimation in the developed watersheds (Table 5.2). Both methods resulted in land use variables and soil variables being classified as important. Percent canopy cover and soil K-factor were the important variables that appear in the results from both RPART and RFR. These are two variables that were also important for determining yields in developed watersheds, and likely relate to nitrogen concentrations for similar reasons.

3.4.3   Undeveloped watersheds

The undeveloped watersheds had the greatest average number of extracted variables by RPART and by RFR, as compared to developed and agricultural watersheds (Table 5.2). This was true when examining loads and yields in undeveloped watersheds

111

as well. Land use, hydrologic, and geographic variables were extracted as important by both RPART and RFR analyses, with multiple other variable types not shared by the two methods. Percent forested canopy in the watershed, annual water yield, and population density were important variables extracted by both RPART and RFR. Canopy cover (an indicator of undeveloped land) was negatively correlated with concentrations, and population density (an indicator of developed land) was positively correlated with concentrations. Both relate to land use, which has been an important driver of nitrogen in all other analyses. This may imply that canopy cover, which was also extracted as important in agricultural and developed watersheds, may be a disproportionately important land use variable in undeveloped watersheds, acting to reduce or prevent the increase of concentration in streams.

3.5 Application of important factors

A series of multiple linear regression equations (models) were created using the variables that RPART and RFR extracted as the most important for determining loads, yields, and concentrations (Table 5.3). The nine models created from the variables extracted by the RPART and the nine models created from the variables extracted by RFR, although different, often produce similar R^2 values (Table 5.3). However, the models generated from RFR variables used an average of one fewer variable per model and had normalized root mean squared errors (NRMSE) that were nearly equivalent or smaller than those from models using RPART variables (Table 5.4).

3.5.1 Total nitrogen load estimation

The six models generated using variables extracted by RPART and RFR for estimating loads among the three watershed groups (Table 5.3) have the highest mean adjusted R^2 when compared to the models generated for estimating nitrogen yields or concentrations. Among the models for estimating loads, the models for undeveloped watersheds had the highest average R^2 values when compared to the models for the developed or agricultural watersheds.

The regression models for agricultural and undeveloped watersheds performed well when comparing model estimated loads to the loads based on actual stream

measurements. Precision of estimates was measured by the NRMSE. NRMSE for the agricultural and undeveloped calibration and validation watersheds suggest satisfactory fitting of the data (Table 5.4), although the model applied to the calibration set of agricultural watersheds using RPART variables was classified as poorly fitting due to the elevated NRMSE (Escurra et al. 2014). Figures 5.4A-C show the model estimated loads against the measured loads with a 1:1 reference line. Load estimates for the calibration watersheds in all three land use groups cluster around the 1:1 line. Strength and direction of bias was measured as the percent of overestimated values. Loads were overestimated and underestimated with roughly the same frequency (Table 5.4). For the validation analyses, the models had a greater tendency to overestimate loads. However, NRMSE remained satisfactory in the undeveloped watersheds and decreased in the agricultural watersheds. This suggests that the models for estimating loads in agricultural and undeveloped watersheds (and the important variables associated with the models) are able to be applied to a wide range of unstudied watersheds. Although both RPART and RFR based models tended toward a higher frequency of overestimating the measured values in the validation set of undeveloped watersheds (Table 5.4), the results remained satisfactory based on the classification of NRMSE suggested by Escurra et al. (2014). The models for the developed watersheds performed poorly based on the same classification. Although R^2 values were quite high for the developed watersheds models, the NRMSE suggest poorly fitting models for both the calibration and validation watersheds. Plots of the residuals for estimates of loads for calibration and validation watersheds (as well as for the yield and concentration estimates) are in Appendix E. Residual plots for load estimates show no clear bias or heteroscedasticity (Figures E5.4A-C).

3.5.2   *Total nitrogen yield estimation*

The models for estimating yields among the three watersheds groups had systematically lower R^2 values than the models for estimating loads (Table 5.3). The RPART based model for estimating yields in agricultural watersheds had a higher R^2 values than the similar RFR model.

Although the R^2 values from the RFR based models for agricultural and developed watersheds were smaller, the RFR models had equivalent or lower NRMSE than the RPART based models for the calibration watersheds (Table 5.4). For the validation watersheds, the NRMSE from the RFR based models for agricultural and undeveloped watersheds were considerably smaller than from the RPART models. These two RFR models produced satisfactorily fitting models for the validation watersheds. As with the models for estimating loads, results from applying the models for yield estimation to the calibration and validation data sets among the three watershed groups produced similar outcomes when plotted (Figures 5.5A-C), with nearly all values clustering around the 1:1 line. Estimates by RPART and RFR based models applied to the calibration data overestimated and underestimated with approximately the same frequency across all watershed groups, except the regression model that used the variables from RPART in the model for undeveloped watersheds, in which case the measured values were overestimated 71% of the time. For the validation analyses, the agricultural models had the greatest increase in frequency of overestimates (Table 5.4) compared to the developed and undeveloped watersheds, but the model which used the RFR extracted variables had a satisfactory NRMSE. The RPART based models for estimating yields, when applied to the validation watersheds in each of the land use groups, fit the data poorly based on the elevated NRMSE values, which may reduce their utility in unstudied watersheds. Residual plots for yield estimates show no clear bias or heteroscedasticity (Appendix E Figures E5.5A-C).

3.5.3   *Total nitrogen concentration estimation*

Models for estimating concentrations most often had lower R^2 values compared to models for estimating loads and yields (Table 5.3). Although the R^2 values tended to be the lowest, the concentration models for agricultural watersheds had higher R^2 values than the models for estimating yields in the agricultural watersheds.

The R^2 values from the RFR based models were approximately equivalent or smaller than those from the RPART based models. However, the RFR based models fit the data more accurately than the models which used variables extracted by RPART, based on the NRMSE. NRMSE from the RFR based models in all watershed groups

114

suggest satisfactory fitting of the models to the measured values, and were equivalent or smaller than those from the RPART based models for the calibration and validation data sets (Table 5.4). Figures 5.6A and B show modeled concentration estimates clustering around the 1:1 line when RPART and RFR based models were applied to the agricultural and developed watersheds, but show deviation from the line in the undeveloped watersheds (Figure 5.6C) when the RPART based model was applied. Calibration estimates by RPART and RFR based models overestimated and underestimated the measured concentrations with similar frequency across all watershed groups except when using the variables from RPART in the undeveloped watersheds, in which case the concentrations were overestimated 86% of the time (Table 5.4). The RPART based model applied to the undeveloped watersheds performed poorly, particularly when applied to the set of validation watersheds. This implies that the variables from RPART do not extend well to undeveloped watersheds outside of the calibration watersheds. An unidentified variable, either not extracted by RPART or not included in the list of possible variables for this research, may be the cause of this consistent bias. Or, the watersheds of the validation set were systematically different than those in the calibration set (Figure 5.1). Any difference among the two sets of watersheds did not appear to be based on location though, as the watersheds with the greatest discrepancy (>2mg N/L) were scattered across the United States. RPART models did, however, perform satisfactorily in the agricultural and developed watersheds, and the residual plots for these watersheds show no clear bias or heteroscedasticity (Appendix E Figure E5.6A-C).

4. Discussion

4.1   Importance of a few variables

A main goal of this research was to gain a better understanding of the factors that affect the amount of nitrogen in streams. This was accomplished by finding the physical variables that are strongly related to nitrogen loads, yields, and concentrations. Dominant land use was an important factor when estimating nitrogen in streams, as evidenced by the ubiquitous extraction by RPART and RFR during the initial analyses which used all 509 calibration watersheds. After watersheds were divided into three groups based on

land use, the important variables from each group of watersheds were compared to determine widely applicable important variables. The variables listed in Table 5.2, along with the results presented in Tables 5.3 and 5.4 and Figures 5.4-5.6 provide evidence that informative models can be created and applied to a wide range of watersheds with the use of relatively few variables. The linear regression models were able to explain a large portion of variability in loads, yields, and concentrations with satisfactory accuracy based on the NRMSE as suggested by Escurra et al. (2014).

A few variables had wide ranging importance across the land use groups and across load, yield, and concentration analyses. Nitrogen application (application mass for load analyses; application rate for yield and concentration analyses) was extracted from RPART and RFR in nearly two thirds of all analyses, and in every analysis involving agricultural watersheds. The use of nitrogen fertilizer, particularly in agricultural watersheds has a tremendous affect on the source strength of nitrogen within the watersheds, and thus, determines the amount of nitrogen that is able to be moved across the landscape. The positive correlation between nitrogen application and loads, yields, and concentrations implies that reductions in nitrogen use in agriculture will help decrease the amount of nitrogen that could move to a stream.

Agricultural and developed land uses were shown to have a strong influence on nitrogen in streams. The same land use variables that were used to split the watersheds into the three groups (agricultural, developed, and undeveloped) were frequently extracted by RPART and RFR as important variables within those watershed groups. Agricultural and developed land areas (or surrogate variables such as nitrogen application or impervious buffer, respectively) were important variables extracted for the undeveloped watersheds. The exaggerated influence of agricultural and developed land use was evident within the undeveloped watersheds where those land use types were limited in area (<25% of total). Within the agricultural watersheds, the land area used for agriculture remained an important variable extracted by RPART, RFR, or both when analyzing loads, yields, and concentrations. And, within the developed watersheds, the total area of developed land, or surrogates of the amount of undeveloped land (forested canopy area), were commonly important variables. Agricultural and developed lands

116

frequently require water flowpath modification (additions of roads, culverts, ditches, storm drains, subsurface drains, etc.) to a greater degree than most undeveloped watersheds. These flowpath modifications can move larger amounts of water and nitrogen to streams than in undeveloped watersheds (Craig et al., 2008). Therefore, greater agricultural or developed land is equated with greater flowpath modification, which then can lead to increases in nitrogen in streams (Craig et al., 2008, Lindsey et al., 1998).

4.2   Ease and challenges of variables applied to linear regression models

The important variables extracted by RPART and RFR were used to develop regression models to estimate loads, yields, and concentrations in unstudied areas. Linear models for estimating total nitrogen loads were often the simplest models, frequently using the fewest number of variables while producing the highest R^2 and satisfactorily low NRMSE when compared to estimates of yields or concentrations, particularly in agricultural and undeveloped watersheds. Nitrogen loads are strongly dependent on streamflow (Tu, 2009), and streamflow is often governed by the variables used in this study (Brooks et al., 2012). These relations between the load, streamflow, and the other independent variables likely link the loads strongly with the independent variables.

Agricultural watersheds were well modeled when compared to developed and undeveloped watersheds, often requiring the fewest number of variables while resulting in the highest mean R^2 values and satisfactory NRMSE for both the calibration and validation watersheds. The landscape and management modifications that agricultural areas undergo to make the land suitable for agriculture alter the importance of a select few variables (nitrogen application, for example) in the watershed, and can lead to reductions in landscape heterogeneity.

The watersheds that were classified as undeveloped consistently resulted in more extracted variables in the regression trees created by RPART and resulted in more important variables to be extracted by RFR than for the developed or agricultural watersheds. This suggests that undeveloped watersheds are more heterogeneous than agricultural or developed watersheds, and/or individual processes (controls on nitrogen movement to streams) are not as dominant as they are in the agricultural and developed

117

watersheds. Without the landscape alteration associated with agriculture or development, all factors on the landscape retain their natural level of importance, with no single variable importance being anthropogenically increased or decreased to create a greater effect on the amount of nitrogen in streams.

4.3   Relative importance of the flowpath variables

Flowpath variables relate directly to how water and nitrogen move to streams, and have been shown in previous research to have a great effect on the amount of nitrogen in streams (Bernhardt et al., 2008, Frank et al., 2000, Goodridge and Melack, 2012, Molenat et al., 2002, Schilling and Zhang, 2004, Sebestyen et al., 2000, Tesoriero et al., 2009). Flowpath variables were expected to be extracted as important variables by RPART and RFR with a higher frequency than they were. However, flowpath related variables were only extracted sporadically throughout all analyses in each of the different watershed groups, appearing in five of the 18 analyses. There are a few possible reasons why the flowpath variables were not extracted by RPART and RFR more frequently. The flowpath variables may have been obscured by other variables. For instance, land use was shown to play a large role in the estimation of loads, yields, and concentrations through the analyses of this study. Although land use is not a flowpath variable as defined by this study, land use does have a large effect on water movement through and across the landscape, and thus, is indirectly related to the flowpath variables (Bernhardt et al., 2008). The relevant information regarding flowpaths may be contained within the land use variables. Or, due to the difficulty of estimating many flowpath variables, their values may have been incorrectly estimated. The difficulty of estimating flowpath variables could also prevent those variables from being extracted by RPART or RFR. Two measures of baseflow index were included in the analysis because both are widely accepted methods, but each is based on different assumptions. There were differences in estimates of baseflow by the two different methods, highlighting the difficulty of confidently estimating or measuring some flowpath variables.

4.4   Implication of this research for modeling purposes

An additional benefit of this research, aside from identifying the important variables for estimating total nitrogen in streams, was the confirmation that these

variables can produce reasonable predictions when applied to unstudied watersheds. Using all 85 variables from this study to produce multiple linear regression models led to models with high adjusted R^2 values. However, the likelihood that data regarding those variables are available in unstudied watersheds is very small, making those models impractical. Many of the models that were created by this research fit the data from calibration and validation watersheds satisfactorily (high R^2 and low NRMSE) (Tables 5.3 and 5.4) using a comparatively small number of variables. Because the NRMSE from the application of the models to the validation watersheds was often similar to those of the calibration watersheds, these models could be quickly applied to any watershed with reasonable confidence and will be much less restricted by data availability.

The important variables extracted by RPART and RFR, and applied to multiple linear regression models produced satisfactory estimates of the amount of nitrogen in streams as quantified by the R^2 and NRMSE. Although RPART and RFR were used separately, each method often extracted the same important variables (or same variable type), which suggests that the extracted variables do play an important role in determining the amount of nitrogen in streams. This should not, however, discount the variables that were only extracted by one of the methods. Differences in how RPART and RFR operate naturally led to different results. But, the results from one method reinforced the results from the other method. RPART did often extracted more variables and frequently produced results (NRMSE) that were of poorer quality than results from RFR based models. This makes extrapolation to unstudied watersheds less certain when using the RPART based models, particularly when estimating nitrogen concentrations in undeveloped watersheds. Satisfactory application to calibration and validation watersheds, coupled with the requirement of fewer variables for model creation suggests that the combination of variables extracted by RFR provide more inherent information about the amount of nitrogen in streams than the combination of variables extracted by RPART. RFR does have the tendency to extract highly correlated variables (Strobl et al., 2008), which can be seen in Table 5.3 as a greater tendency to produce variables that result in a variance inflation factor (VIF) >5. The collinearity of these variables is a problem when creating linear regression models and must be dealt with by either

119

removing those variables, or by choosing a method which allows for collinearity among independent variables.

Although the physical movement of nitrogen to streams is exceedingly complex, a few, well selected variables can provide a good approximation of the amount of nitrogen that enters the streams. A small set of widely applicable variables that could be quickly and easily applied to, or used in the creation of, simple models will be useful in the prediction of total nitrogen loads, yields, and concentrations across a wide variety of streams and their watersheds. The predictions provided by the models from this research, or new models based on this research can inform and isolate areas where more in-depth research, the use of more complex modeling techniques, or alterations to watershed management measures may be beneficial.

References

Alexander, R.B., R.A. Smith, and G.E. Schwarz. 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. Nature, 403(6771), 758-761.

Ambrose III, H.W., K.P. Ambrose, D.J. Emlen, and K.L. Bright. 2002. A handbook of biological investigation. Hunter Textbooks. Inc. Winston-Salem, NC.

Asner, G.P., J.M. Scurlock, and J.A. Hicke. 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. Global Ecology and Biogeography, 12(3), 191-205.

Bernhardt, E.S., L.E. Band, C.J. Walsh, and P.E. Berke. 2008. Understanding, managing, and minimizing urban impacts on surface water nitrogen loading. *Annals of the New York Academy of Sciences*, *1134*(1), 61-96.

Billy, C., F. Birgand, P. Ansart, J. Peschard, M. Sebilo, and J. Tournebize. 2013. Factors controlling nitrate concentrations in surface waters of an artificially drained agricultural watershed. Landscape ecology, 28(4), 665-684.

Brooks, K.N., P.F. Ffolliott, and J.A. Magner. 2012. *Hydrology and the Management of Watersheds*. John Wiley & Sons.

Brown, J.B., L.A. Sprague, and J.A. Dupree. 2011. Nutrient Sources and Transport in the Missouri River Basin, with Emphasis on the Effects of Irrigation and Reservoirs1.

JAWRA Journal of the American Water Resources Association, 47(5), 1034-1060.

Bulmer, M.G. 2012. *Principles of statistics*. Courier Corporation.

Craig, L.S., M.A. Palmer, D.C. Richardson, S. Filoso, E.S. Bernhardt, B.P. Bledsoe, M.W. Doyle, P.M. Groffman, B.A. Hassett, S.S. Kaushal, P.M. Mayer, S.M. Smith and P.R. Wilcock. 2008. Stream restoration strategies for reducing river nitrogen loads. Frontiers in Ecology and the Environment, 6(10), 529-538.

Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.M. Gronberg, P.A. Hamilton, K.J. Hitt, D.K. Mueller, M.D. Munn, B.T. Nolan, L.J. Puckett, M.G. Rupert, T.M. Short, N.E. Spahr, L.A. Sprague, and W.G. Wilber. 2010. The quality of our Nation's waters—Nutrients in the Nation's streams and groundwater, 1992–2004, U.S. Geological Survey Circular 1350, 174 p.

Escurra, J.J., V. Vazquez, R. Cestti, E. De Nys, and R. Srinivasan. 2014. Climate change impact on countrywide water balance in Bolivia. Regional Environmental Change, 14(2), 727-742.

Frank, H., S. Patrick, W. Peter, and F. Hannes. 2000. Export of dissolved organic carbon and nitrogen from Gleysol dominated catchments–the significance of water flow paths. Biogeochemistry, 50(2), 137-161.

Goodridge, B.M. and J.M. Melack. 2012. Land use control of stream nitrate concentrations in mountainous coastal California watersheds. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, *117*(G2), doi: 10.1029/2011JG001833

Hively, W.D., C.J. Hapeman, L.L. McConnell, T.R. Fisher, C.P. Rice, G.W. McCarty, M.S. Ali, 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.H. Fetcho. 2011. Relating nutrient and herbicide fate with landscape features and characteristics of 15 subwatersheds in the Choptank River watershed. Science of the total environment, 409(19), 3866-3878.

Hoos, A.B. and G. McMahon. 2009. Spatial analysis of instream nitrogen loads and factors controlling nitrogen delivery to streams in the southeastern United States using spatially referenced regression on watershed attributes (SPARROW) and regional classification frameworks. Hydrological Processes, 23(16), 2275-2294.

Howarth, R. W., G. Billen, D. Swaney, A. Townsend, N. Jaworski, K. Lajtha, K., J.A. Downing, R. Elmgren, N. Caraco, T. Jordan, F. Berendse, J. Freney, V. Kudeyarov, P. Murdoch and Z. Zhao-Liang. 1996. Regional nitrogen budgets and riverine N and P fluxes for the drainages to the North Atlantic Ocean: Natural and

human influences. In Nitrogen cycling in the North Atlantic Ocean and its watersheds (pp. 75-139). Springer Netherlands.

Liaw, A. and M. Wiener. 2012. randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4.6-7

Lindsey B.D., K.J. Breen, M.D. Bilger, and R.A. Brightbill. 1998. Water quality in the lower Susquehanna River basin, Pennsylvania and Maryland, 1992–95. Washington, DC: US Geological Survey. Circular 1168.

Molénat, J., P. Durand, C. Gascuel-Odoux, P. Davy, and G. Gruau. 2002. Mechanisms of nitrate transfer from soils to stream in an agricultural watershed of French Brittany, Water Air Soil Pollut., 133(1-4), 161–183. doi: 10.1023/A:1012903626192

Moore, R.B., C.M. Johnston, R.A. Smith, and B. Milstead. 2011. Source and Delivery of Nutrients to Receiving Waters in the Northeastern and Mid-Atlantic Regions of the United States1. JAWRA Journal of the American Water Resources Association, 47(5), 965-990.

Mueller, D.K., B.C. Ruddy, and W.A. Battaglin. 1997. Logistic model of nitrate in streams of the upper-midwestern United States. Journal of Environmental Quality, 26(5), 1223-1230.

Oehler, F. and A.H. Elliott. 2011. Predicting stream N and P concentrations from loads and catchment characteristics at regional scale: A concentration ratio method. Science of the Total Environment, 409(24), 5392-5402.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the Association of American Geographers 77(1):118-125.

Onderka, M., S. Wrede, M. Rodný, L. Pfister, L., Hoffmann, and A. Krein. 2012. Hydrogeologic and landscape controls of dissolved inorganic nitrogen (DIN) and dissolved silica (DSi) fluxes in heterogeneous catchments. Journal of Hydrology, 450, 36-47.

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

Peterson, E.W. and C. Benning. 2013. Factors influencing nitrate within a low-gradient agricultural stream. Environmental earth sciences, 68(5), 1233-1245.

Preston, S.D., R.B. Alexander, G.E. Schwarz, and C.G. Crawford. 2011. Factors Affecting Stream Nutrient Loads: A Synthesis of Regional SPARROW Model Results for the Continental United States1. JAWRA Journal of the American Water Resources Association, 47(5), 891-915.

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org

Rebich, R.A., N.A. Houston, S.V. Mize, D.K. Pearson, P.B. Ging, and C.E. Hornig. 2011. Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico from Streams in the South-Central United States1. JAWRA Journal of the American Water Resources Association, 47(5), 1061-1086.

Reich, P.B., D.W. Peterson, D.A. Wedin, and K. Wrage. 2001. Fire and vegetation effects on productivity and nitrogen cycling across a forest-grassland continuum. *Ecology*, *82*(6), 1703-1719.

Reynolds, B. and A. Edwards. 1995. Factors influencing dissolved nitrogen concentrations and loadings in upland streams of the UK. Agricultural Water Management, 27(3), 181-202.

Ribaudoa M.O., R. Heimlich, and M. Peters. 2005. Nitrogen sources and Gulf hypoxia: potential for environmental credit trading. *Ecological Economics* **52**(2) : 159-168. DOI:10.1016/j.ecolecon.2004.07.021

Robertson, D.M. and D.A. Saad. 2011. Nutrient inputs to the laurentian great lakes by source and watershed estimated using SPARROW watershed models1. JAWRA Journal of the American Water Resources Association, 47(5), 1011-1033.

Rogerson, P. 2001. Statistical methods for geography. Sage.

Saleh, D. and J.L. Domagalski. 2012. Using SPARROW to Model Total Nitrogen Sources, and Transport in Rivers and Streams of California and Adjacent States, USA. In *AGU Fall Meeting Abstracts* (Vol. 1, p. L06).

Schilling, K. and Y.K. Zhang. 2004. Baseflow contribution to nitrate-nitrogen export from a large, agricultural watershed, USA. Journal of Hydrology, 295(1), 305-316.

Schilling, K.E. and R.D. Libra. 2000. The relationship of nitrate concentrations in streams to row crop land use in Iowa. Journal of Environmental Quality, 29(6), 1846-1851.

Schwarz, G.E., A.B. Hoos, R.B. Alexander, and R.A. Smith. 2006. The SPARROW surface water-quality model: Theory, application, and user documentation, U.S. Geological Survey Techniques and Methods Report, Book 6, Chapter B3.

Sebestyen, S.D., E.W. Boyer, J.B. Shanley, C. Kendall, D.H. Doctor, G.R Aiken, and N. Ohte. 2008. Sources, transformations, and hydrological processes that control stream nitrate and dissolved organic matter concentrations during snowmelt in an upland forest. *Water Resources Research*, *44*(12).

Stålnacke, P., A. Pengerud, M. Bechmann, J. Garnier, C. Humborg, and V. Novotny. 2009. Nitrogen driving force and pressure relationships at contrasting scales: implications for catchment management. International Journal of River Basin Management, 7(3), 221-232.

Strobl, C., A.L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis. 2008. Conditional variable importance for random forests. *BMC bioinformatics*, *9*(1), 307.

Swank, W.T., J.D. Knoepp, J.M. Vose, S.N. Laseter, and J.R. Webster. 2014. Response and Recovery of Water Yield and Timing, Stream Sediment, Abiotic Parameters, and Stream Chemistry Following Logging. *Long-Term Response of a Forest Watershed Ecosystem: Clearcutting in the Southern Appalachians*, 36.

Tesoriero, A.J., J.H. Duff, D.M. Wolock, N.E. Spahr, and J.E. Almendinger. 2009. Identifying pathways and processes affecting nitrate and orthophosphate inputs to streams in agricultural watersheds. *Journal of environmental quality*, *38*(5), 1892-1900.

Therneau, T.M., B. Atkinson, and B. Ripley. 2013. rpart: Recursive partitioning. R package version, 4.1-3

Tu, J. 2009. Combined impact of climate and land use changes on streamflow and water quality in eastern Massachusetts, USA. *Journal of Hydrology*, *379*(3), 268-283.

United States Geological Survey (USGS) 2014. National Hydrography Dataset (NHD). United States Geological Survey, Reston, VA.

Vitousek, P.M., J.D. Aber, R.W. Howarth, G.E. Likens, P.A. Matson, D.W. Schindler, W.H. Schlesinger, and D.G. Tilman. 1997. Human alteration of the global nitrogen cycle: sources and consequences. *Ecological applications*, *7*(3), 737-750.

Wahl, K.L. and Wahl, T.L., 1995. Determining the Flow of Comal Springs at New Braunfels, Texas, Texas Water '95, American Society of Civil Engineers, August 16-17, 1995, San Antonio, Texas, pp. 77-86.

Wise, D.R., and H.M. Johnson. 2011. Surface-Water Nutrient Conditions and Sources in the United States Pacific Northwest1. JAWRA Journal of the American Water Resources Association, 47(5), 1110-1135.

Table 5.1: List of important variables that were extracted by recursive partition (RPART) and random forest regression (RFR) analyses. Variables are long term average values unless otherwise noted in the description. A full list of all variables from this research is located in Appendix E Table E5.2.

Dependent variable name	Variable description	Data source
TNLoad	Total nitrogen load (kg N/yr)	1
TNYield	Total nitrogen yield (kg N/km^2/yr)	Calculated
TNConc	Flow-weighted mean nitrogen concentration (mg N/L)	Calculated

Independent variable name (area-normalized, if applicable)	Variable description	Data source
BGTNLoad	Total nitrogen background load (kg N/yr)	2
NMass (NRate[‡])	Mean nitrogen applied to cultivated and pasture agricultural land in 1997 (kg N)	4
Canopy (CanopyPer[#])	Mean canopy cover within the watershed (km^2)	5
ImpBuff (ImpBuffPer[+])	Mean impervious land area within 100m Buffer (km^2)	5
NLCDAg (NLCDAgPer[#])	Total area of 2006 national land cover dataset (NLCD) agricultural land use classes 81 and 82 (km^2)	5
NLCDDev (NLCDDevPer[#])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 21, 22, 23, and 24 (km^2)	5
NLCDDevI (NLCDDevIPer[#])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 22, 23, and 24 (km^2)	5
NLCDAgBuff (NLCDAgBuffPer[+])	Total area of 2006 national land cover dataset (NLCD) agricultural land use classes 81 and 82 in the 100m buffer (km^2)	5
NLCDDevBuff (NLCDDevBuffPer[+])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 21, 22, 23, and 24 in the 100m buffer (km^2)	5
NLCDDevIBuff (NLCDDevIBuffPer[+])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 22, 23, and 24 in the 100m buffer (km^2)	5
HLR	Predominant hydrologic landscape region (HLR) category	6

Independent variable name (area-normalized, if applicable)	Variable description	Data source
MaxEcoReg	Predominant level III nutrient ecoregion category within the watershed	5
Streamflow	Mean annual streamflow (m^3/s)	1
H2OVolume (H2OYield[#])	Annual stream volume (m^3)	Calculated
Irrigated (IrrigatedPer)	Irrigated land within the watershed (km^2)	5
Area	Total watershed area (km^2)	4
MRB	SPARROW major river basin identifier	1
Pop2000	2000 Population density (people/km^2)	7
AWC	Mean soil available water capacity (cm/cm) within the watershed	6
HSGBD	Mean percent of soil within the watershed as hydrologic soil group HGBD	6
HSGC	Mean percent of soil within the watershed as hydrologic soil group HGC	6
Kfactor	Mean soil K-factor within the watershed	6
PercOM	Mean percent organic matter in soil within the watershed	6
PercSilt	Mean percent silt in soil within the watershed	6
SRL55	Percent of soil restrictive layer above 55cm within the watershed	4
SubDrain	Land subject to subsurface drainage within the watershed in 1992 (km^2)	4
ArtDrain	Artificially drained land within the watershed in 1992 (km^2)	4
BFI-WAHL	Baseflow index: from Wahl and Wahl (1995) BFI program	8
BFI-NHD	Baseflow index: from national hydrology dataset	6
Runoff	Mean estimated runoff (mm/yr)	9
SatOF	Mean estimated saturation overland flow (% of total streamflow)	6
AnnualPdays	Mean number of days with precipitation per year (days)	10
MayPdays	Mean number of days of precipitation in May (days)	10

Independent variable name (area-normalized, if applicable)	Variable description	Data source
JulPdays	Mean number of days of precipitation in July (days)	10
NovPdays	Mean number of days of precipitation in November (days)	10

‡ Application rate (kg/km^2)

Percent of watershed area

+ Percent of 100m buffer area

1 Moore et al., 2011, Hoos and McMahon, 2009, Robertson and Saad, 2011, Brown et al., 2011, Rebich et al., 2011, Wise and Johnson, 2011, Saleh and Domagalski, 2012

2 Smith et al. 2003

3 Personal communications with Mike Wieczorek and Naomi Nakagaki, January 15, 2015

4 Personal communications with Mike Wieczorek, January 15, 2015

5 Personal communications with Mike Wieczorek and Andrew LaMotte, January 15, 2015

6 Personal communications with Mike Wieczorek and Dave Wolock, January 15, 2015

7 Personal communications with Mike Wieczorek and Curtis Price, January 15, 2015

8 Wahl and Wahl, 1995

9 Personal communications with Mike Wieczorek, Dave Wolock, and Gregory McCabe, January 15, 2015

10 Personal communications with Mike Wieczorek, James Falcone, and Ryan Hill, January 15, 2015

Table 5.2: Important variables from recursive partitioning (RPART) and random forest regression (RFR) for total nitrogen load, yield, and concentration analyses. Shaded cells represent important variables extracted from RPART or RFR and correspond to the listed Pearson correlation coefficient (r) between that variable and the dependent variable (total nitrogen load, yield, or concentration). Complex variables are those variables which contain intrinsic information relating to geology, physiography, vegetation, climate, soils, land use, wildlife, or hydrology, but are presented as a single classification. See Table 5.1 for variable descriptions.

Variable group	Variable name	Load Ag r	RPART	RFR	Load Dev r	RPART	RFR	Load Undev r	RPART	RFR	Yield Ag r	RPART	RFR	Yield Dev r	RPART	RFR	Yield Undev r	RPART	RFR	Conc Ag r	RPART	RFR	Conc Dev r	RPART	RFR	Conc Undev r	RPART	RFR
Chemical	BGTNLoad [#]							0.69																				
	NMass* [#]	0.82			0.58			0.54			0.63						0.36			0.68								
Land use	Canopy*							0.70						-0.43			0.26			-0.64			-0.48			-0.38		
	ImpBuff* [#]	0.56			0.72									0.38									0.21			0.37		
	NLCDAg* [#]	0.80						0.59			0.63			0.38						0.67								
	NLCDDev* [#]				0.77			0.69						0.35														
	NLCDDevI* [#]				0.72									0.33												0.35		
	NLCDAgBuff*										0.52									0.66								
	NLCDDevBuff* [#]				0.73						0.29			0.39														
	NLCDDevIBuff* [#]				0.71			0.35						0.36												0.31		
Complex	HLR										NA						NA									NA		
	MaxEcoReg										NA									NA			NA			NA		
Hydrologic	Streamflow [#]				0.81																							
	H2OVolume* [#]				0.81									0.39			0.59									-0.40		
	Irrigated* [#]										-0.17																	
Geographic	Area																									0.20		
	MRB																									NA		
	Pop2000* [#]										0.12									-0.07						0.27		
Soil	AWC										0.46																	
	HSGBD [#]													-0.05														
	HSGC																						0.05					
	Kfactor										0.36												0.47					
	PercOM [#]																						-0.21					
	PercSilt							0.07			0.26																	
	SRL55 [#]							0.06			0.32																	
Flowpath	SubDrain [#]	0.40																										
	ArtDrain* [#]				0.21																							
	BFI-WAHL				-0.08																							
	BFI-NHD				-0.28																					-0.12		
	Runoff										0.14						0.29									-0.01		
	SatOF																						-0.27					
Climate	AnnualPdays [#]										-0.01									-0.27								
	JulPdays																											
	MayPdays																0.36											
	NovPdays	0.12																										

* Total value when analyzing total nitrogen load, area-normalized value when analyzing total nitrogen yield and concentration
[#] Log10-transformed for use in linear regression equations   NA – Not applicable

Table 5.3: Multiple linear regression equations produced with variables extracted by recursive partitioning (RPART) and random forest regression (RFR) analyses with corresponding R^2 for equations. Variables with a variance inflation factor (VIF) >5 were removed from multiple linear regression equations (*italicized*).

Dependent variable	Watershed group	Variable extraction method	Multiple linear regression model equation	R^2
Total nitrogen load[#] (kg/yr)	Agricultural watersheds	RPART	0.862 + 0.748 (NMass[#]) + *NLCD8X* + 0.197 (ImpBuff[#]) + 0.0394 (NovPdays)	0.76
		RFR	0.780 + 0.797 (NMass[#]) + *NLCD8X* + 0.022 (SubDrain[#])	0.68
	Developed watersheds	RPART	3.72 + 0.680 (NLCDDev[#]) + 0.423 (ImpBuff[#]) - 0.0209 (ArtDrain[#]) + 0.0127 (BFI-WAHL) - 0.0134 (BFI-NHD)	0.69
		RFR	3.19 + *ImpBuff* + *NLCDDevBuff* + 0.726 (NLCDDev[#]) + 0.297 (NLCDDevIBuff[#]) + *NLCDDevI* + 0.093 (Nmass[#])	0.69
	Undeveloped watersheds	RPART	1.80 + 0.703 (Canopy[#]) + 0.093 (NLCDDevIBuff[#]) + 0.079 (NMass[#]) + *Streamflow* + 0.242 (BGTNLoad[#]) - 0.006 (SRL55[#]) + 0.006 (PercSilt)	0.76
		RFR	3.53 + *NLCDAg* + 0.049 (NLCDDev[#]) + 0.050 (NMass[#]) + *H2OVolume* + 0.428 (Streamflow[#]) + 0.380 (Canopy[#]) + 0.044 (NLCDDevIBuff[#]) + 0.149 (DevBuff[#])	0.78
Total nitrogen yield[#] (kg/km^2/yr)	Agricultural watersheds	RPART	-0.285 + 0.756 (NRate[#]) - 0.045 (Pop2000[#]) + 0.001 (Runoff) + 0.336 (NLCDDevBuffPer[#]) + 0.008 (Irrigated[#]) + 0.004 (PercSilt)	0.57
		RFR	-0.246 + 539 (NRate[#]) + 0.772 (NLCDAgPer[#]) + 0.006 (PercSilt) - 0.004 (NLCDAgBuffPer)	0.48
	Developed watersheds	RPART	3.16 - 0.009 (CanopyPer) + 0.371 (H2OYield[#]) + HLR[‡] + 0.642 (Kfactor) - 1.08 (AnnualPdays[#]) + 0.065 (SRL55[#])	0.51
		RFR	2.57 + 0.229 (Kfactor) + 1.64 (AWC) - 0.011 (CanopyPer) + MaxEcoReg[‡] + HLR[‡]	0.50
	Undeveloped watersheds	RPART	-0.607 - 0.081 (NLCDDevBuffPer[#]) + 0.077 (NRate[#]) + 0.024 (MayPdays) + 0.487 (H2OYield[#]) + 0.0002 (CanopyPer) + 0.0004 (Runoff) + 0.219 (ImpBuffPer[#]) + 0.009 (HSGBD[#])	0.51
		RFR	-0.831 + *ImpBuffPer* + 0.075 (NRate[#]) + 0.170 (NLCDDevBuffPer[#]) + *NLCDAgPer* + 0.052 (NLCDDevIBuffPer[#]) + 0.125 (NLCDDevIPer[#]) + 0.566 (H2OYield[#]) - 0.177 (NLCDDevPer[#])	0.52
Total nitrogen concentration[#] (mg N/L)	Agricultural watersheds	RPART	-1.60 + MaxEcoReg[‡] + 0.671 (NRate[#]) + 0.287 (ImpBuffPer[#]) - 0.044 (JulPdays) - 0.096 (Pop2000[#]) - 0.001 (CanopyPer)	0.63
		RFR	-1.78 + 0.428 (NRate[#]) + MaxEcoReg[‡] + 0.534 (NLCDAgPer[#]) + *NLCDAgBuffPer* - 0.003 (CanopyPer)	0.60
	Developed watersheds	RPART	0.493 - 0.009 (CanopyPer) + 1.36 (Kfactor) - 0.002 (HSGC) - 0.195 (PercOM[#]) - 0.029 (SatOF)	0.34
		RFR	-0.076 - 0.009 (CanopyPer) + 1.51 (Kfactor) + MaxEcoReg[‡]	0.35
	Undeveloped watersheds	RPART	3.50 + 0.095 (Pop2000[#]) - 0.005 (CanopyPer) - 0.532 (H2OYield[#]) + MRB[‡] + 0.089 (ImpBuffPer[#]) + HLR[‡] - 0.002 (BFI-WAHL) + 0.001 (Runoff) - 0.00001 (Area) + MaxEcoReg[‡]	0.44
		RFR	1.93 - 0.370 (H2OYield[#]) - 0.002 (CanopyPer) + 0.063 (Pop2000[#]) + 0.052 (NLCDDevIPerc[#]) + 0.085 (NLCDDevIBuffPer[#])	0.31

\# Variable is Log10-transformed

‡ Categorical variable; coefficients values for insertion into multiple linear regression equations are in Table E5.3 of Appendix E.

Table 5.4: Percentage of overestimated values from the application of each model to the various sets of watersheds. Root mean squared error of estimates normalized to the mean measured value (NRMSE) is presented for the calibration and validation data sets.

Dependent variable	Watershed group	Variable extraction method	Calibration		Validation	
			NRMSE	% overestimate	NRMSE	% overestimate
Total nitrogen load	Agricultural	RPART	0.92	50	0.62	71
		RFR	0.88	49	0.59	64
	Developed	RPART	1.0	48	1.6	62
		RFR	1.1	53	1.5	65
	Undeveloped	RPART	0.66	47	0.79	36
		RFR	0.67	52	0.84	53
Total nitrogen yield	Agricultural	RPART	0.65	51	0.98	67
		RFR	0.65	51	0.78	67
	Developed	RPART	0.63	51	1.1	46
		RFR	0.58	49	1.8	58
	Undeveloped	RPART	1.0	71	1.7	69
		RFR	0.67	48	0.65	51
Total nitrogen concentration	Agricultural	RPART	0.61	45	0.87	52
		RFR	0.61	47	0.91	52
	Developed	RPART	0.85	53	0.67	65
		RFR	0.73	48	0.60	35
	Undeveloped	RPART	0.75	86	6.3	93
		RFR	0.65	49	0.54	51

Figure 5.1: Distribution and land use classification of study watersheds
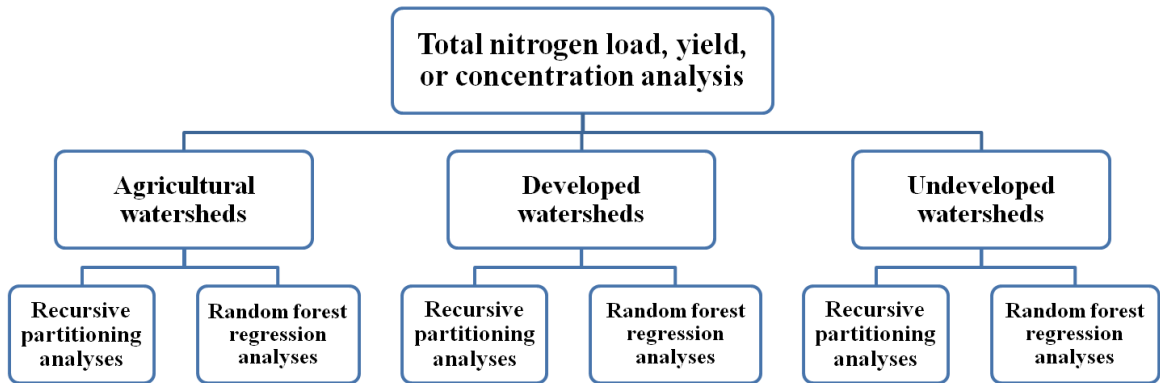
Figure 5.2: Breakdown of important variable extraction analyses conducted for total nitrogen loads, yields, and concentrations.
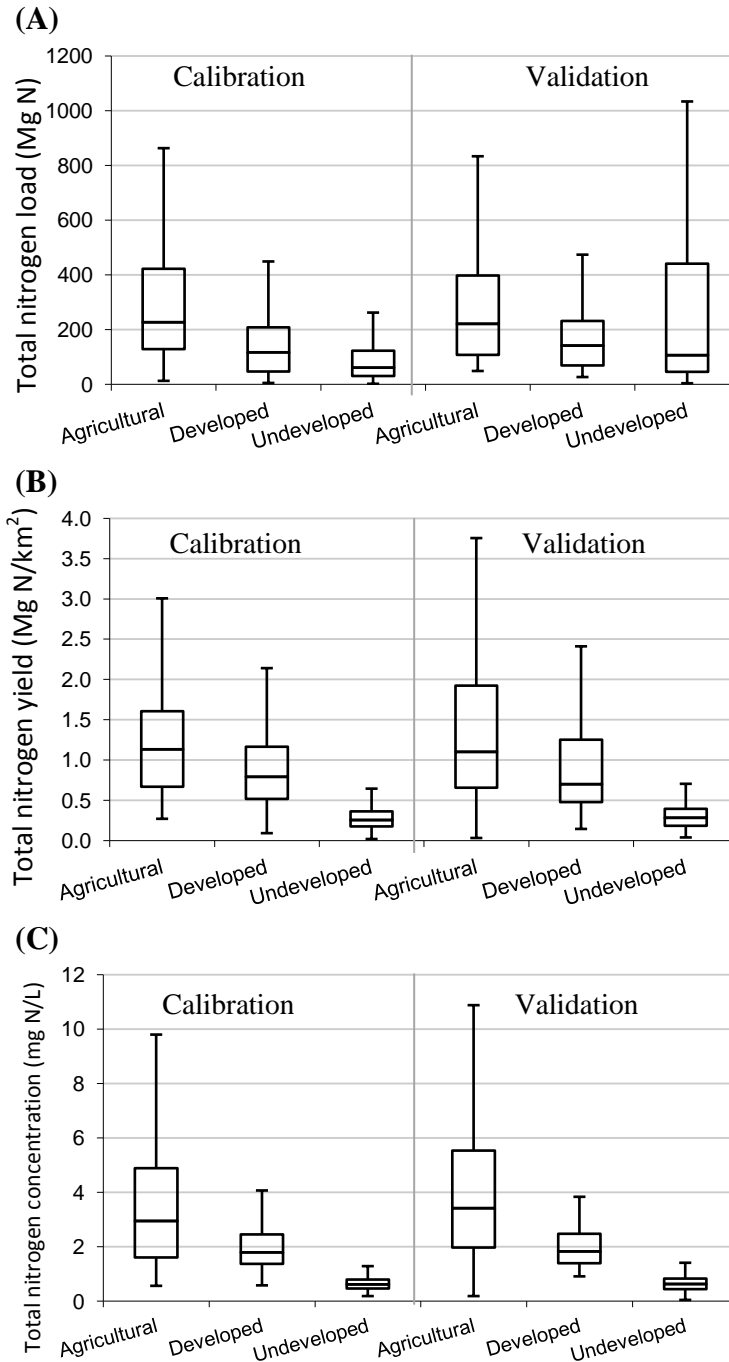
Figure 5.3: Boxplots of annual total nitrogen **(A)** loads, **(B)** yields, and **(C)** concentrations for calibration and validation streams and their watersheds. Outliers were removed from the figures to increase detail. Outliers were defined as values greater than the 75th percentile plus 1.5 times the interquartile range, or values less than the 25th percentile minus 1.5 times the interquartile range. For calibration, n = 509. For validation, n = 127.

(A)

(B)

(C)
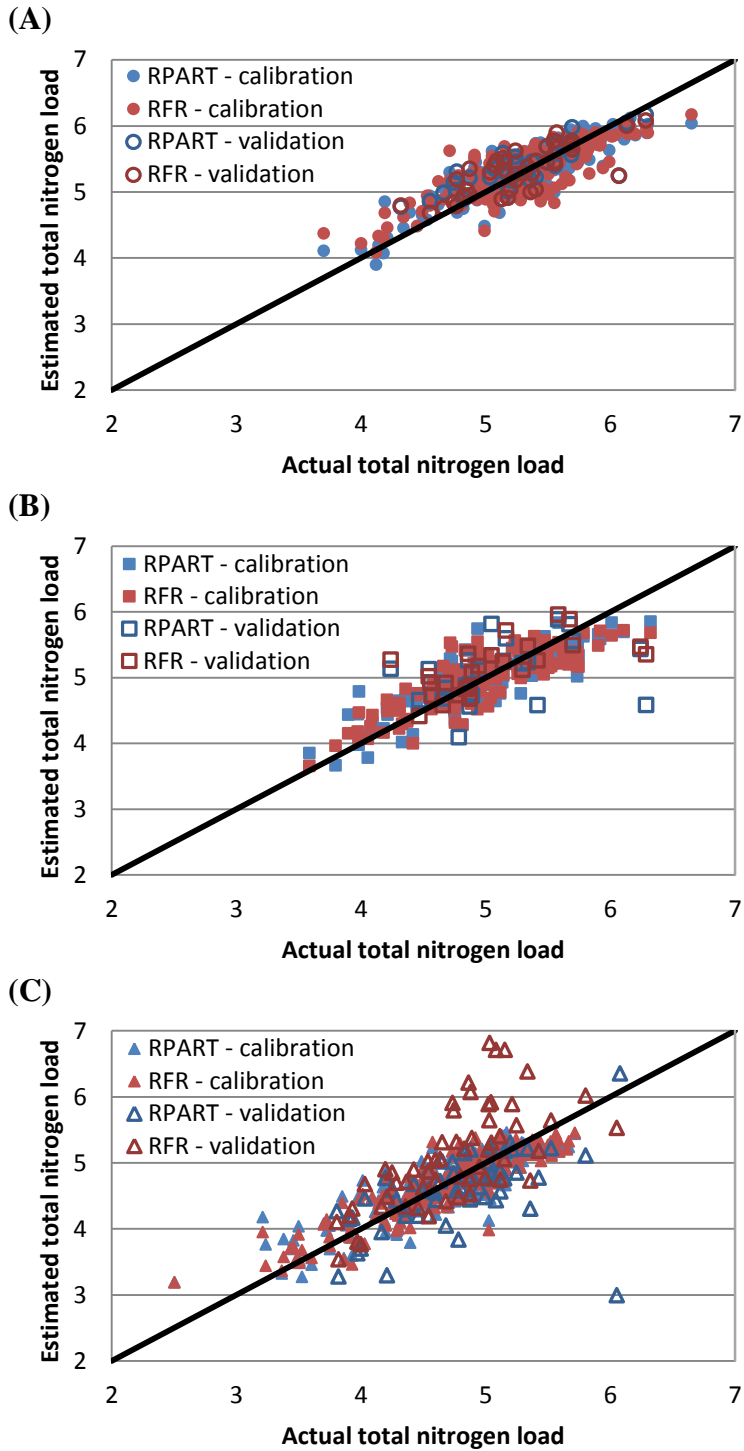
Figure 5.4: Total annual nitrogen load (log10-transformed) estimates from multiple linear regression equations for calibration and validation data sets compared to stream loads based on measurements for **(A)** agricultural, **(B)** developed, and **(C)** undeveloped watersheds using important variables from recursive partitioning and random forest regression. Solid black line represents a 1:1 relationship.

(A)

(B)

(C)
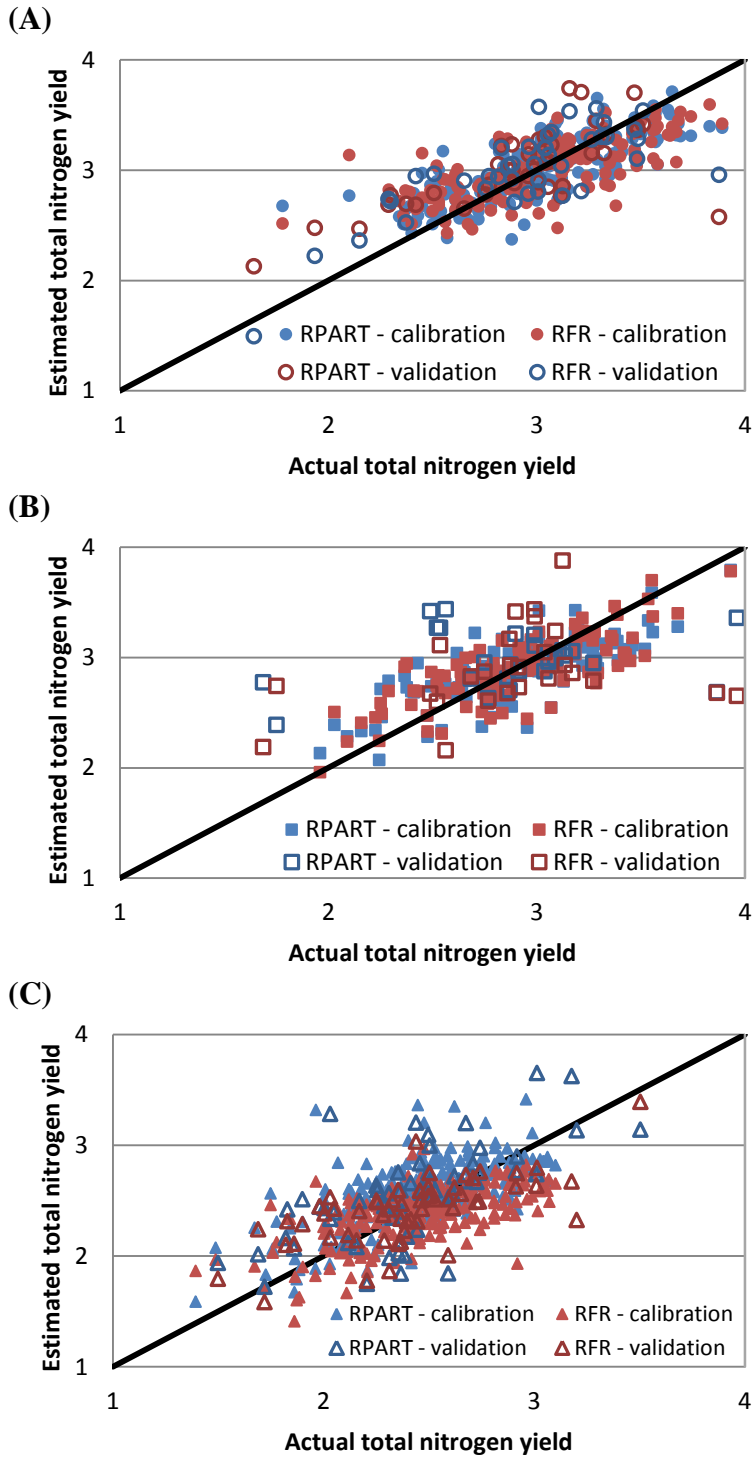
Figure 5.5: Total annual nitrogen yield (log10-transformed) estimates from multiple linear regression equations for calibration and validation data sets compared to watershed yields based on measurements for **(A)** agricultural, **(B)** developed, and **(C)** undeveloped watersheds using important variables from recursive partitioning and random forest regression. Solid black line represents a 1:1 relationship.
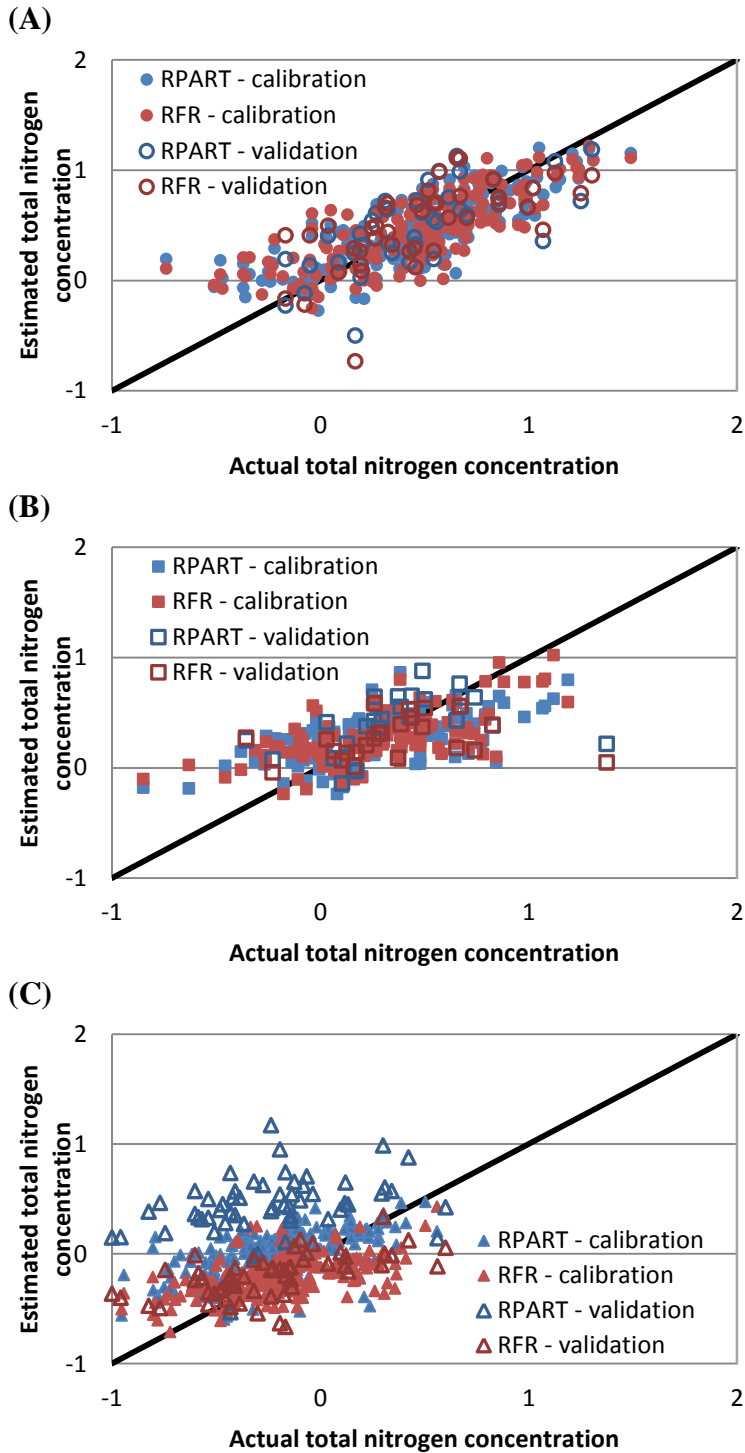
Figure 5.6: Total annual nitrogen concentration (log10-transformed) estimates from multiple linear regression equations for calibration and validation data sets compared to stream concentrations based on measurements for **(A)** agricultural, **(B)** developed, and **(C)** undeveloped watersheds using important variables from recursive partitioning and random forest regression. Solid black line represents a 1:1 relationship.

Chapter 6: Summary

1. <u>Benefits of this research</u>

This research advances the understanding of hydrograph separation techniques and improves upon a widely used method of hydrograph separation. Understanding the strengths and shortcomings of each method of hydrograph separation allows for a more informed choice of method when applied to a stream. This study also reinforces the understanding of the effects that flowpath can have on nitrogen movement to streams, by confirming expected results of nitrogen concentrations in streams of highly modified, agricultural watersheds. Finally, this research provides additional evidence which suggests that a small number of carefully selected variables can provide significant information about the long term annual total nitrogen loads and concentrations in streams, and the yields from their watersheds.

Both graphical and chemical tracer-based hydrograph separation (BFI program from Wahl and Wahl (1995) and end-member mixing analysis, respectively) can be used for hydrograph separation, if the data are available. Due to current data availability, graphical methods of hydrograph separation are more widely applicable because they require less data. However, as technology advances and the number and type of high temporal resolution water quality sensors increase, the number of streams where ratio-EMMA can be applied will increase and become more diverse. The more process oriented ratio-EMMA could be used to identify the conditions under which graphical methods may falter. This will lead to a great increase in our understanding of a process that affects the amount of nitrogen in streams.

Distillation of a large number of watershed variables into a small subset which explain a large amount of variability in total nitrogen levels in streams will be a benefit to watershed management and future research. Prioritization of limited resources is a constant concern for watershed managers and researchers alike. The application of the important variables extracted from this research could inform decisions regarding the allocation of resources, and provide a defensible course of action.

2. **Future research possibilities**

The results from this research have provided insight into the processes that move water and nitrogen from the landscape to streams. However, many questions have been left unanswered, and many new questions have arisen. The streams and watersheds examined in this research were all relatively small, which is a positive characteristic for a watershed study, as the heterogeneity within the watershed can be limited. However, at larger stream sizes, heterogeneity will likely be greater.

A research idea that came from this research was to test the effectiveness of graphical hydrograph separation at larger stream sizes. It is theorized that at some point within a river network, the hydrograph will become attenuated to a level that would hinder the graphical methods from accurately estimating slowflow. It will be beneficial to quantify the accuracy of multiple methods of hydrograph separation at various spatial scales to determine thresholds of reliability for each method.

Although the ratio-EMMA method tested in this research succeeded as a proof of concept, improvements to the current method can still be made. Nitrate concentration, as used alongside specific conductance in this ratio-EMMA research, is not ideally suited for use with EMMA because it can undergo substantial transformations as it is transported across the landscape, and after it enters the stream. However, as new continuous sensors are created which are able to measure additional water parameters that are well suited for EMMA, such as oxygen-18 (Munksgaard et al., 2012), the data can be applied to the ratio-EMMA. Furthermore, an additional assumption that remains within the ratio-EMMA is that of a constant chemical concentration in the slowflow end-member. Information from new field studies and/or a literature review to determine the magnitude and timescale of changes in slowflow chemical concentration will further increase the accuracy of the ratio-EMMA method.

Additional variables could also be examined for their importance in estimating total nitrogen in streams. Although 90 variables were examined for their wide ranging importance in determining nitrogen levels in streams, the selected variables do not represent an exhaustive list. Studies have shown that variables such as topography (Onderka et al., 2012) and stream sinuosity (Peterson and Benning, 2013) may be

important drivers of nitrogen levels in streams. These data were not available to be collected for all watersheds at the time of this research, but could prove useful for estimating nitrogen in streams. Along the same lines, determining how to best apply the important variables to models could allow for higher prediction accuracy when estimating total nitrogen loads, yields, and concentrations. When using multiple linear regression, important variables were occasionally removed from an equation due to an elevated variance inflation factor. The removal of variables from the multiple linear regression models represents a loss of important information. Retention of all important variables through the use of an alternative method, such as partial least squares regression, could allow for more accurate estimates of total nitrogen loads, yields, and concentrations in unstudied streams.

Although this study provides a small number of variables that are highly correlated with nitrogen in streams, establishing causal links between each variable and nitrogen movement to streams is incomplete. Much research is still needed to determine these links in an effort to reduce nitrogen leaching to groundwater, runoff to streams, and export to coastal areas.

References

Munksgaard, N.C., C.M. Wurster, A. Bass, I. Zagorskis, and M.I. Bird. 2012. First continuous shipboard $\delta18O$ and δD measurements in sea water by diffusion sampling—cavity ring-down spectrometry. Environmental Chemistry Letters, 10(3), 301-307.

Onderka, M., S. Wrede, M. Rodný, L. Pfister, L., Hoffmann, and A. Krein. 2012. Hydrogeologic and landscape controls of dissolved inorganic nitrogen (DIN) and dissolved silica (DSi) fluxes in heterogeneous catchments. Journal of Hydrology, 450, 36-47.

Peterson, E.W. and C. Benning. 2013. Factors influencing nitrate within a low-gradient agricultural stream. Environmental earth sciences, 68(5), 1233-1245.

Wahl, K.L. and Wahl, T.L. 1995. Determining the Flow of Comal Springs at New Braunfels, Texas,*Texas Water '95*, American Society of Civil Engineers, August 16-17, 1995, San Antonio, Texas, pp. 77-86

Bibliography

Alexander, R.B., R.A. Smith, and G.E. Schwarz. 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. Nature, 403(6771), 758-761.

Ali, G.A., A.G. Roy, M.C. Turmel, and F. Courchesne. 2010. Source-to-stream connectivity assessment through end-member mixing analysis. *Journal of Hydrology* **392**(3) : 119-135. doi: 10.1016/j.jhydrol.2010.07.049

Ambrose III, H.W., K.P. Ambrose, D.J. Emlen, and K.L. Bright. 2002. A handbook of biological investigation. Hunter Textbooks. Inc. Winston-Salem, NC.

Anderson, D.M., P.M. Glibert, and J.M. Burkholder. 2002. Harmful Algal Blooms and Eutrophication Nutrient Sources, Composition, and Consequences. Estuaries, 25(4b):704-726.

Asner, G.P., J.M. Scurlock, and J.A. Hicke. 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. Global Ecology and Biogeography, 12(3), 191-205.

Ator, S.W., J.M. Denver, and M.J. Brayton. 2005. Hydrologic and geochemical controls on pesticide and nutrient transport to two streams on the Delmarva Peninsula. US Department of the Interior, US Geological Survey.

Bachman, L.J., D.E. Krantz, and J.K. Böhlke. 2002. Hydrogeologic framework, ground-water geochemistry, and assessment of nitrogen yield from base flow in two agricultural watersheds, Kent County, Maryland. U.S. Environmental Protection Agency Report EPA/600/R-02/008, 79 p.

Baron, J.S., E.K. Hall, B.T. Nolan, J.C. Finlay, E.S. Bernhardt, J.A. Harrison, F. Chan, and E.W. Boyer. 2013. The interactive effects of excess reactive nitrogen and climate change on aquatic ecosystems and water resources of the United States. *Biogeochemistry*, *114*(1-3), 71-92, doi: 10.1007/s10533-012-9788-y

Bernhardt, E.S., L.E. Band, C.J. Walsh, and P.E. Berke. 2008. Understanding, managing, and minimizing urban impacts on surface water nitrogen loading. *Annals of the New York Academy of Sciences*, *1134*(1), 61-96.

Bijeriego, M, R.D. Hauch, and R.A. Olson. 1979. Uptake, translocation, and utilization of ^{15}N-depleted fertilizer in irrigated corn. *Soil Science. Society of America Journal* **43**(3) : 528-533. doi: 10.2136/sssaj1979.03615995004300030020x

Billy, C., F. Birgand, P. Ansart, J. Peschard, M. Sebilo, and J. Tournebize. 2013. Factors controlling nitrate concentrations in surface waters of an artificially drained agricultural watershed. Landscape ecology, 28(4), 665-684.

Blann, K.L., J.L. Anderson, G.R. Sands, and B. Vondracek. 2009. Effects of Agricultural Drainage on Aquatic Ecosystems: A Review. Critical Reviews in Environmental Science and Technology 39(11), 909-1001. doi: 10.1080/10643380801977966

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 Resources Research, 31(9), 2,319-2,339. doi:10.1029/95WR01584

Botkin, D.B. and E.A. Keller. 2007. Environmental Science: Earth as a living planet. Sixth edition. John Wiley and Sons Inc. Hoboken, New Jersey.

Brooks, K.N., P.F. Ffolliott, and J.A. Magner. 2012. *Hydrology and the Management of Watersheds*. John Wiley & Sons.

Brown, J.B., L.A. Sprague, and J.A. Dupree. 2011. Nutrient Sources and Transport in the Missouri River Basin, with Emphasis on the Effects of Irrigation and Reservoirs1. JAWRA Journal of the American Water Resources Association, 47(5), 1034-1060.

Bunce, N.J. 1993. Introduction to Environmental Chemistry. Wuerz Publishing Ltd. Winnipeg, Canada.

Bulmer, M.G. 2012. *Principles of statistics*. Courier Corporation.

Camargo, J.A. and A. Alonso. 2006. Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment. Environment International. 32(6):831–849. doi:10.1016/j.envint.2006.05.002

Canfield, D.E., A.N. Glazer, and P.G. Falkowski. 2010. The evolution and future of earth's nitrogen cycle. Science. 330(6001):192-196. DOI: 10.1126/science.1186120

Capel, P.D., K.A. McCarthy, and J.E. Barbash. 2008. National, holistic, watershed-scale approach to understand the sources, transport, and fate of agricultural chemicals: Supplemental material section. *Journal of Environmental Quality* **37**(3) : 983–993. doi: 10.2134/jeq2007.0226

Cassell, E.A. and J.C. Clausen. 1993. Dynamic Simulation Modeling for Evaluating Water Quality Response to Agricultural BMP Implementation, Wat. Sci. Tech. 28(3-5), 635-648.

Christophersen, N. and R.P. Hooper. 1992. Multivariate analysis of stream water chemical data: The use of principal components analysis for the end-member mixing problem. *Water Resources Research* **28**(1) : 99-107. doi: 10.1029/91WR02518

Christophersen, N., C. Neal, R.P. Hooper, R.D. Vogt, and S. Andersen. 1990. Modeling streamwater chemistry as a mixture of soilwater end-members—a step towards second-generation acidification models. Journal of Hydrology, 116(1), 307-320, doi:10.1016/0022-1694(90)90130-P

Craig, L.S., M.A. Palmer, D.C. Richardson, S. Filoso, E.S. Bernhardt, B.P. Bledsoe, M.W. Doyle, P.M. Groffman, B.A. Hassett, S.S. Kaushal, P.M. Mayer, S.M. Smith and P.R. Wilcock. 2008. Stream restoration strategies for reducing river nitrogen loads. Frontiers in Ecology and the Environment, 6(10), 529-538.

Cunningham, W.P., M.A. Cunningham, and B. Saigo. 2003. Environmental Science: A global concern. 7th edition. McGraw Hill. New York. ISBN-13: 978-0073532547

David, M.B., L.E. Drinkwater, and G.F. McIsaac. 2010. Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, *39*(5), 1657-1667, doi:10.2134/jeq2010.0115

Domagalski, J.L, S. Ator, R. Coupe, K. McCarthy, D. Lampe, M. Sandstrom, and N. Baker. 2008. Comparative Study of Transport Processes on Nitrogen, Phosphorus, and Herbicides to Streams in Five Agricultural Basins, USA, J. Environ. Qual., 37(3), 1158-1169. doi: 10.2134/jeq2007.0408

Dubrovsky, N.M. and P.A. Hamilton. 2010. Nutrients in the Nation's streams and groundwater: National Findings and Implications, U.S. Geological Survey Fact Sheet 2010-3078, 6 p.

Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.M. Gronberg, P.A. Hamilton, K.J. Hitt, D.K. Mueller, M.D. Munn, B.T. Nolan, L.J. Puckett, M.G. Rupert, T.M. Short, N.E. Spahr, L.A. Sprague, and W.G. Wilber. 2010. *The quality of our Nation's waters—Nutrients in the Nation's streams and groundwater, 1992–2004*. U.S. Geological Survey Circular 1350, Reston VA.

Eckhardt, K. 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes, 19(2), 507-515, doi: 10.1002/hyp.5675

Escurra, J.J., V. Vazquez, R. Cestti, E. De Nys, and R. Srinivasan. 2014. Climate change impact on countrywide water balance in Bolivia. Regional Environmental Change, 14(2), 727-742.

Frank, H., S. Patrick, W. Peter, and F. Hannes. 2000. Export of dissolved organic carbon and nitrogen from Gleysol dominated catchments–the significance of water flow paths. Biogeochemistry, 50(2), 137-161.

Goodridge, B.M., and J.M. Melack. 2012. Land use control of stream nitrate concentrations in mountainous coastal California watersheds. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, *117*(G2), doi: 10.1029/2011JG001833

Green, M.B, 2007. Hydrologic Control of Stream Water N:P Ratios, PhD Thesis, Water Resources Science, University of Minnesota, St. Paul, Minnesota

Green, C.T., L.J. Puckett, J.K. Böhlke, B.A. Bekins, S.P. Phillips, L.J. Kauffman, J.M. Denver, and H.M. Johnson. 2008. Limited Occurrence of Denitrification in Four Shallow Aquifers in Agricultural Areas of the United States, J. Environ. Qual., 37(3), 994-1009. doi: 10.2134/jeq2006.0419

Gronberg, J.M. and C.R. Kratzer. 2006. Environmental setting of the Lower Merced River Basin, California, U.S. Geological Survey Scientific Investigations Rep. 2006–5152, Reston, VA.

Hancock, T.C. and M.J. Brayton. 2006. *Environmental setting of the Morgan Creek Basin, Maryland, 2002–04*. U.S. Geological Survey Open-File Report 2006–1151, Reston VA.

Hem, J.D. 1985. *Study and interpretation of the chemical characteristics of natural water (Vol. 2254)*. Department of the Interior, US Geological Survey, Reston VA.

Hively, W.D., C.J. Hapeman, L.L. McConnell, T.R. Fisher, C.P. Rice, G.W. McCarty, M.S. Ali, 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.H. Fetcho. 2011. Relating nutrient and herbicide fate with landscape features and characteristics of 15 subwatersheds in the Choptank River watershed. Science of the total environment, 409(19), 3866-3878.

Holman, I.P., T.M. Hess, and S.C. Rose. 2011. A broad-scale assessment of the effect of improved soil management on catchment baseflow index. Hydrological Processes, 25(16), 2563-2572. doi: 10.1002/hyp.8131

Hooper, A.B., T. Vannelli, D.J. Bergmann, and D.M. Arciero. 1997. Enzymology of the oxidation of ammonia to nitrite by bacteria. Antonie van Leeuwenhoek. 71(1-2):59-67. doi: 10.1023/A:1000133919203.

Hooper, R.P. 2003. Diagnostic tools for mixing models of stream water chemistry. *Water Resources Research*, *39*(3), doi: 10.1029/2002WR001528

Hooper, R.P. and C.A. Shoemaker. 1986. A comparison of chemical and isotopic hydrograph separation. *Water Resources Research* **22**(10) : 1444-1454. doi: 10.1029/WR022i010p01444

Hooper, R.P., N. Christophersen, and N.E. Peters. 1990. Modeling streamwater chemistry as a mixture of soil water end-members—an application to the Panola Mountain Catchment, Georgia, USA, J. Hydrol., 116(1-4), 321-343. doi:10.1016/0022-1694(90)90131-G

Hoos, A.B. and G. McMahon. 2009. Spatial analysis of instream nitrogen loads and factors controlling nitrogen delivery to streams in the southeastern United States using spatially referenced regression on watershed attributes (SPARROW) and regional classification frameworks. Hydrological Processes, 23(16), 2275-2294.

Howarth, R. W., G. Billen, D. Swaney, A. Townsend, N. Jaworski, K. Lajtha, K., J.A. Downing, R. Elmgren, N. Caraco, T. Jordan, F. Berendse, J. Freney, V. Kudeyarov, P. Murdoch and Z. Zhao-Liang. 1996. Regional nitrogen budgets and riverine N and P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. In Nitrogen cycling in the North Atlantic Ocean and its watersheds (pp. 75-139). Springer Netherlands.

Hubbard, L., D.W. Kolpin, S.J. Kalkhoff, and D.M. Robertson. 2011. Nutrient and sediment concentrations and corresponding loads during the historic June 2008 flooding in eastern Iowa. J. Environ. Qual., *40*(1), 166-175. doi:10.2134/jeq2010.0257

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(1), 241-258, doi:10.1016/j.jhydrol.2006.07.006

Kadlec, R. and R.L. Knight. 1996. Treatment wetlands. Lewis Publishers. Boca Raton, Florida.

Kapoor, A., and T. Viraraghavan. 1997. Nitrate removal from drinking water—review. J. Environ. Eng. 123:371–380.

Kladivko, E. J., G.E. Van Scoyoc, E.J. Monke, K.M. Oates, and W. Pask. 1991. Pesticide and Nutrient Movement into Subsurface Tile Drains on a Silt Loam Soil in Indiana, J. Environ. Qual., 20(1), 264–270. doi: 10.2134/jeq1991.00472425002000010043x

Kramer, D.A. 2004. Nitrogen. U.S. Geological Survey Minerals Yearbook. Accessed March 17, 2015. minerals.usgs.gov/minerals/pubs/commodity/nitrogen/nitromyb04.pdf

Kronholm, S.C., and P.D. Capel. 2014. A comparison of high-resolution specific conductance-based end-member mixing analysis and a graphical method for baseflow separation of four streams in hydrologically challenging agricultural watersheds. Hydrological Processes, doi: 10.1002/hyp.10378

Lathrop, T.R. 2006. Environmental setting of the Sugar Creek and Leary Weber Ditch Basins, Indiana, 2002–04, U.S. Geological Survey Scientific Investigations Report 2006–5170, Reston, VA.

Lee, C.J., D.P. Mau, and T.J. Rasmussen. 2005. Effects of nonpoint and selected point contaminant sources on stream-water quality and relation to land use in Johnson county, northeastern Kansas, October 2002 through June 2004: U.S. Geological Survey Scientific Investigations Report 2005-5144, 104 p.

Liaw, A. and M. Wiener. 2012. randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4.6-7

Lindsey B.D., K.J. Breen, M.D. Bilger, and R.A. Brightbill. 1998. Water quality in the lower Susquehanna River basin, Pennsylvania and Maryland, 1992–95. Washington, DC: US Geological Survey. Circular 1168.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, M. Studer,P. Roudier. 2015. Cluster Analysis Extended Rousseeuw et al. R package version, 2.0-1

Marsh, K.L., G.K. Sims, and R.L. Mulvaney. 2005. Availability of urea to autotrophic ammonia-oxidizing bacteria as related to the fate of 14C- and 15N-labeled urea added to soil. Biology and Fertility of Soils, 42(2):137-145. doi: 10.1007/s00374-005-0004-2

McCarthy, K.A. and H.M. Johnson. 2009. *Effect of agricultural practices on hydrology and water chemistry in a small irrigated catchment, Yakima River basin, Washington*. U.S. Geological Survey Scientific Investigations Rep. 2009–5030, Reston, VA.

McCarthy, K.A., C.E. Rose, and S.J. Kalkhoff. 2012. *Environmental settings of the South Fork Iowa River Basin, Iowa, and the Bogue Phalia Basin, Mississippi, 2006–10*. U.S. Geological Survey Scientific Investigations Rep. 2012–5021, Reston, VA.

McGee, E. 1997. Acid Rain and Our Nation's Capital. U.S. Geological Survey. Accessed March 17, 2015. http://pubs.usgs.gov/gip/acidrain/5.html

McNamara, J.P., D.L. Kane, and L.D. Hinzman. 1997. Hydrograph separations in an Arctic watershed using mixing model and graphical techniques. *Water Resources Research* **33**(7) : 1707-1719. doi: 10.1029/97WR01033

Meessen, J.H. 2010. Urea in Ullmann's Encyclopedia of Industrial Chemistry, Wiley-VCH, Weinheim. Accessed March 17, 2015. DOI: 10.1002/14356007.a27_333.pub2

Meisinger, JJ, Bandel VA, Stanford G, and Legg JO. 1985. Nitrogen utilization of corn under minimal tillage and moldboard plow tillage: I. Four year results using labeled fertilizer on an Atlantic Coastal Plain soil. *Agron. Journal* **77**(4) : 602–611. doi:10.2134/agronj1985.00021962007700040022x

Miller, M.P., D.D. Susong, C.L. Shope, B.J. Stolp, V.M. Heilweil, and T. Marston. 2013. Continuous estimation of baseflow contribution to snowmelt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach. *Water Resources Research*

Molénat, J., P. Durand, C. Gascuel-Odoux, P. Davy, and G. Gruau. 2002. Mechanisms of nitrate transfer from soils to stream in an agricultural watershed of French Brittany, Water Air Soil Pollut., 133(1-4), 161–183. doi: 10.1023/A:1012903626192

Moore, R.B., C.M. Johnston, R.A. Smith, and B. Milstead. 2011. Source and Delivery of Nutrients to Receiving Waters in the Northeastern and Mid-Atlantic Regions of the United States1. JAWRA Journal of the American Water Resources Association, 47(5), 965-990.

Mueller, D.K. and D.R. Helsel. 1996. Nutrients in the Nation's Waters--Too Much of a Good Thing? U.S. Geological Survey Circular 1136, 24 p. Accessed March 17, 2015. http://pubs.usgs.gov/circ/circ1136/

Mueller, D.K., B.C. Ruddy, and W.A. Battaglin. 1997. Logistic model of nitrate in streams of the upper-midwestern United States. Journal of Environmental Quality, 26(5), 1223-1230.

Munksgaard, N.C., C.M. Wurster, A. Bass, I. Zagorskis, and M.I. Bird. 2012. First continuous shipboard δ18O and δD measurements in sea water by diffusion sampling—cavity ring-down spectrometry. Environmental Chemistry Letters, 10(3), 301-307.

Murakami K. 1984. Dredging for controlling eutrophication of lake Kasumigaura, Japan. Lake and Reservoir Management 1(1): 592-598.

National Atmospheric Deposition Program (NADP) (NRSP-3). 2013. NADP Program Office, Illinois State Water Survey, 2204 Griffith Dr., Champaign, IL 61820. Accessed from: http://nadp.sws.uiuc.edu/data/ntndata.aspx

National Oceanic and Atmospheric Administration. 2009. NOAA Study Shows Nitrous Oxide Now Top Ozone-Depleting Emission. Accessed March 17, 2015. http://www.noaanews.noaa.gov/stories2009/20090827_ozone.html

National Oceanic and Atmospheric Administration. 2015. Greenhouse gases. Accessed March 17, 2015. http://lwf.ncdc.noaa.gov/oa/climate/gases.html

Oehler, F. and A.H. Elliott. 2011. Predicting stream N and P concentrations from loads and catchment characteristics at regional scale: A concentration ratio method. Science of the Total Environment, 409(24), 5392-5402.

Ogunkoya, O.O. and A. Jenkins. 1993. Analysis of storm hydrograph and flow pathways using a three-component hydrograph separation model. *Journal of Hydrology* **142**(1) : 71-88. doi: 10.1016/0022-1694(93)90005-T

Olarewaju, O., M. Adetunji, C. Adeofun, and I. Adekunle. 2009. Nitrate and phosphorus loss from agricultural land: implications for nonpoint pollution, Nutrient Cycling in Agroecosystems. 85(1), 79-85. doi: 10.1007/s10705-009-9249-8.

Olson, R.V. 1980. Fate of tagged nitrogen fertilizer applied to irrigated corn. *Soil Sci. Soc. Am. J.* **44**(3) : 514-517. doi: 10.2136/sssaj1980.03615995004400030015x

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the Association of American Geographers 77(1):118-125.

Onderka, M., S. Wrede, M. Rodný, L. Pfister, L., Hoffmann, and A. Krein. 2012. Hydrogeologic and landscape controls of dissolved inorganic nitrogen (DIN) and dissolved silica (DSi) fluxes in heterogeneous catchments. Journal of Hydrology, 450, 36-47.

Payne, K.L., H.M. Johnson, and R.W. Black. 2007. *Environmental setting of the Granger Drain and DR2 Basins, Washington 2003–2004*. U.S. Geological Survey Scientific Investigations Rep. 2007–5102, Reston, VA.

Pellerin, B.A., W.M. Wollheim, X. Feng, and C.J. Vörösmarty. 2008. The application of electrical conductivity as a tracer for hydrograph separation in urban catchments. *Hydrological Processes* **22**(12): 1810-1818. doi: 10.1002/hyp.6786

Peters, N.E, 1994. Water-quality variations in a forested Piedmont catchment, Georgia, USA. J. Hydrol. 156(1-4), 73-90. doi:10.1016/0022-1694(94)90072-8

Peterson, E.W. and C. Benning. 2013. Factors influencing nitrate within a low-gradient agricultural stream. Environmental earth sciences, 68(5), 1233-1245.

Pilgrim, D.H., D.D. Huff, and T.D. Steele. 1979. Use of specific conductance and contact time relations for separating flow components in storm runoff. *Water Resources Research* **15**(2) : 329-339. doi: 10.1029/WR015i002p00329

Pinder, G.F. and J.F. Jones. 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. *Water Resources Research* **5**(2) : 438-445. doi: 10.1029/WR005i002p00438

Poor, C.J. and J.J. McDonnell. 2007. The effects of land use on stream nitrate dynamics. *Journal of Hydrology*, *332*(1), 54-68, doi:10.1016/j.jhydrol.2006.06.022

Preston, S.D., R.B. Alexander, G.E. Schwarz, and C.G. Crawford. 2011. Factors Affecting Stream Nutrient Loads: A Synthesis of Regional SPARROW Model Results for the Continental United States1. JAWRA Journal of the American Water Resources Association, 47(5), 891-915.

PRISM Climate Group. 2008. PRISM Climate Data, Northwest Alliance for Computational Science & Engineering, Oregon State University, http://prism.oregonstate.edu

Puckett, L.J., C. Zamora, H. Essaid, J.T. Wilson, H.M. Johnson, M.J. Brayton, and J.R. Vogel. 2008. Transport and Fate of Nitrate at the Ground-water/Surface-water Interface. J. Environ. Qual. 37(3), 1034-1050. doi: 10.2134/jeq2006.0550

Rasmussen, T.J., C.J. Lee, and A.C. Ziegler. 2008. Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006: U.S. Geological Survey Scientific Investigations Report 2008–5014, 103 p.

Raymond, P.A., M.B. David, and J.E. Saiers. 2012. The impact of fertilization and hydrology on nitrate fluxes from Mississippi watersheds. *Current Opinion in Environmental Sustainability*, *4*(2), 212-218, doi: 10.1016/j.cosust.2012.04.001

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org

Rebich, R.A., N.A. Houston, S.V. Mize, D.K. Pearson, P.B. Ging, and C.E. Hornig. 2011. Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico

from Streams in the South-Central United States1. JAWRA Journal of the American Water Resources Association, 47(5), 1061-1086.

Reddy, G.B. and K.R. Reddy. 1993. Fate of Nitrogen-15 Enriched Ammonium Nitrate Applied to Corn, *Soil Sci. Soc. Am. J.* **57**(1) : 111-115. doi:10.2136/sssaj1993.03615995005700010021x

Reich, P.B., D.W. Peterson, D.A. Wedin, and K. Wrage. 2001. Fire and vegetation effects on productivity and nitrogen cycling across a forest-grassland continuum. *Ecology*, *82*(6), 1703-1719.

Reynolds, B. and A. Edwards. 1995. Factors influencing dissolved nitrogen concentrations and loadings in upland streams of the UK. Agricultural Water Management, 27(3), 181-202.

Ribaudoa, M.O., R. Heimlich, and M. Peters. 2005. Nitrogen sources and Gulf hypoxia: potential for environmental credit trading. *Ecological Economics* **52**(2) : 159-168. doi:10.1016/j.ecolecon.2004.07.021

Rice, K.C. and O.P. Bricker. 1995. Seasonal cycles of dissolved constituents in streamwater in two forested catchments in the mid-Atlantic region of the eastern USA, J. of Hydrol. 170(1-4), 137-158. doi:10.1016/0022-1694(95)92713-N

Ricklefs, R.E. 2001. The Economy of Nature. 5th edition. W. H Freeman and Company. New York. ASIN: B003Q6EIBM

Rimmer, A. and A. Hartmann. 2014. Optimal hydrograph separation filter to evaluate transport routines of hydrological models. Journal of Hydrology, 514, 249-257, doi:10.1016/j.jhydrol.2014.04.033

Robertson, D.M. and D.A. Saad. 2011. Nutrient inputs to the laurentian great lakes by source and watershed estimated using SPARROW watershed models1. JAWRA Journal of the American Water Resources Association, 47(5), 1011-1033.

Robson, A. and C. Neal. 1990. Hydrograph separation using chemical techniques: an application to catchments in mid-Wales. *Journal of Hydrology* **116**(1) : 345-363. doi: 10.1016/0022-1694(90)90132-H

Rogerson, P. 2001. Statistical methods for geography. Sage.

Ross, D.S., R.J. Bartlett, F.R. Magdoff, and G.J. Walsh. 1994. Flow path studies in forested watersheds of headwater tributaries of Brush Brook, Vermont. Water Resour. Res. 30(9), 2611-2618. doi:10.1029/94WR01490

Runkel, R.L., C.G. Crawford, and T.A. Cohn. 2004. Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. U.S. Geological Survey Techniques and Methods book 4, chapter A5, Reston, VA.

Saleh, D. and J.L. Domagalski. 2012. Using SPARROW to Model Total Nitrogen Sources, and Transport in Rivers and Streams of California and Adjacent States, USA. In *AGU Fall Meeting Abstracts* (Vol. 1, p. L06).

Sanford, W.E. and J.P. Pope. 2013. The Role of Groundwater in Delaying Chesapeake Bay Restoration, Environmental Science and Technology, in press.

Sanford, W.E., D.L. Nelms, J.P. Pope, and D.L. Selnick. 2012. *Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis*. U.S. Geological Survey Scientific Investigations Report 2011–5198, Reston, VA.

Sauer, T. J., R.B. Alexander, J.V. Brahana, and R.A. Smith. 2008. The Importance and Role of Watersheds in the Transport of Nitrogen. Nitrogen in the Environment, 203.

Schilling, K. and Y.K. Zhang. 2004. Baseflow contribution to nitrate-nitrogen export from a large, agricultural watershed, USA. Journal of Hydrology, 295(1), 305-316.

Schilling, K.E. and R.D. Libra. 2000. The relationship of nitrate concentrations in streams to row crop land use in Iowa. Journal of Environmental Quality, 29(6), 1846-1851.

Schwartz, S.S. 2007. Automated Algorithms for Heuristic Base-Flow Separation. *JAWRA Journal of the American Water Resources Association*, *43*(6), 1583-1594, doi: 10.1111/j.1752-1688.2007.00130.x

Schwarz, G.E., A.B. Hoos, R.B. Alexander, and R.A. Smith. 2006. The SPARROW surface water-quality model: Theory, application, and user documentation, U.S. Geological Survey Techniques and Methods Report, Book 6, Chapter B3.

Sebestyen, S.D., E.W. Boyer, J.B. Shanley, C. Kendall, D.H. Doctor, G.R Aiken, and N. Ohte. 2008. Sources, transformations, and hydrological processes that control stream nitrate and dissolved organic matter concentrations during snowmelt in an upland forest. *Water Resources Research*, *44*(12).

Senus, M.P., M.J. Langland, and D.L. Moyer. 2005. Nutrient and sediment concentrations, loads, and trends for four nontidal tributaries in the Chesapeake

Bay watershed, 1997-2001. US Department of the Interior, US Geological Survey.

Sharpley, A.N., T. Daniel, T. Sims, J. Lemunyon, R. Stevens, and R. Parry. 2003. Agricultural Phosphorus and Eutrophication, Second Edition, Agricultural Research Service ARS–149, 38 p. Accessed March 17, 2015. http://www.ars.usda.gov/is/np/Phos&Eutro2/agphoseutro2ed.pdf

Shelton, L.R. 1994. Field Guide for Collecting and Processing Stream-Water Samples for the National Water-Quality Assessment Program, U.S. Geological Survey Open-File Report 94-455, 50 p., Reston, VA.

Sklash, M.G. and R.N. Farvolden. 1979. The role of groundwater in storm runoff. *Journal of Hydrology* **43**(1) : 45-65. doi: 10.1016/0022-1694(79)90164-1

Sloto, RA and Crouse MY. 1996. HYSEP: A computer program for streamflow hydrograph separation and analysis: U.S. Geological Survey Water-Resources Investigations Report 96-4040, Reston, VA.

Smith, E.A. 2012. *Spatial and temporal variability of preferential flow in a subsurface-drained landscape in north-central Iowa*. Doctoral dissertation, Water Resources Science. University of Minnesota, St. Paul, MN

Smith, E.L. and L.M. Kellman. 2011. Nitrate loading and isotopic signatures in subsurface agricultural drainage systems. J. Environ. Qual., 40(4), 1257-1265. doi:10.2134/jeq2010.0489.

Smith, R.A., R.B. Alexander, and G.E. Schwarz. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. *Environmental Science & Technology*, *37*(14), 3039-3047.

Stålnacke, P., A. Pengerud, M. Bechmann, J. Garnier, C. Humborg, and V. Novotny. 2009. Nitrogen driving force and pressure relationships at contrasting scales: implications for catchment management. International Journal of River Basin Management, 7(3), 221-232.

Starr, C., R. Taggar, C. Evers, and L. Starr. 2015. Biology: The unity and Diversity of Life. 14th edition. Cengage Learning. 1024 pp. ISBN-13: 978-1305073951

Stewart, M., J. Cimino, and M. Ross. 2007. Calibration of base flow separation methods with streamflow conductivity. *Ground Water* **45**(1) : 17-27. doi: 10.1111/j.1745-6584.2006.00263.x

Strobl, C., A.L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis. 2008. Conditional variable importance for random forests. *BMC bioinformatics*, *9*(1), 307.

Swank, W.T., J.D. Knoepp, J.M. Vose, S.N. Laseter, and J.R. Webster. 2014. Response and Recovery of Water Yield and Timing, Stream Sediment, Abiotic Parameters, and Stream Chemistry Following Logging. *Long-Term Response of a Forest Watershed Ecosystem: Clearcutting in the Southern Appalachians*, 36.

Tesoriero, A.J., J.H. Duff, D.M. Wolock, N.E. Spahr, and J.E. Almendinger. 2009. Identifying Pathways and Processes Affecting Nitrate and Othophosphate Inputs to Streams in Agricultural Watersheds, J. Environ. Qual. 38(5), 1892-1900. doi: 10.2134/jeq2008.0484

Therneau, T.M., B. Atkinson, and B. Ripley. 2013. rpart: Recursive partitioning. R package version, 4.1-3

Thornburg, J. 2009. *Temporal and spatial variability in nitrate and water quality parameters of subsurface drains in an agricultural stream*. MS Thesis. Water Resources Science. University of Minnesota, St. Paul, MN

Townsend, A.R. and R.W. Howarth. 2010. Fixing the global nitrogen problem. *Scientific American* **302**(2) : 64-71. doi:10.1038/scientificamerican0210-64

Tu, J. 2009. Combined impact of climate and land use changes on streamflow and water quality in eastern Massachusetts, USA. *Journal of Hydrology*, *379*(3), 268-283.

Tufekcioglu, A., J.W. Raich, T.M. Isenhart, and R.C. Schultz. 2003. Biomass, carbon and nitrogen dynamics of multi-species riparian buffers within an agricultural watershed in Iowa, USA. *Agroforestry Systems*, *57*(3), 187-198. doi: 10.1023/A:1024898615284

United States Census Bureau (USCB). 2015. Total Midyear Population for the World: 1950-2050. Washington, D.C. Online at: http://www.census.gov/population/international/data/worldpop/table_population.php. Accessed 4-25-15.

United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS). 2015. Quick Stats. Washington, D.C. Online at: http://quickstats.nass.usda.gov/?long_desc__LIKE=Corn&x=20&y=6#24D5A0FE-5F7C-3A4B-9523-8FCC678E6BA9. Accessed: 4-25-15.

United States Environmental Protection Agency (US EPA) 2004. Photochemical smog—what it means for us. USEPA Publication90/04 (March 2004). Accessed March 17, 2015. http://www.epa.sa.gov.au/xstd_files/Air/Information%20sheet/info_photosmog.pdf

United States Environmental Protection Agency (US EPA) 2014. Nitrogen Dioxide. Accessed March 17, 2015. http://epa.gov/airquality/nitrogenoxides/

United States Environmental Protection Agency (US EPA) 2015a. Effects of Acid Rain - Surface Waters and Aquatic Animals. Accessed March 17, 2015. http://epa.gov/acidrain/effects/surface_water.html

United States Environmental Protection Agency (US EPA) 2015b. Hypoxia: What is hypoxia and what causes it? Accessed March 17, 2015. http://water.epa.gov/type/watersheds/named/msbasin/hypoxia101.cfm

United States Environmental Protection Agency (US EPA) 2015c. Dissolved Oxygen. Accessed March 17, 2015. http://www.epa.gov/caddis/ssr_do_wtl.html

United States Environmental Protection Agency (US EPA) 2015d. Overview of Greenhouse Gases. Accessed March 17, 2015. http://www.epa.gov/nitrousoxide/scientific.html

United States Geological Survey (USGS) 2006. Collection of water samples (ver. 2.0): U.S. Geological Survey Techniques of Water-Resources Investigations. Book 9, chap. A4. Accessed January 21, 2011 at http://water.usgs.gov/owq/FieldManual/chapter4/html/Ch4_contents.html.

United States Geological Survey (USGS) 2011. Eutrophication. Accessed March 17, 2015. http://toxics.usgs.gov/definitions/eutrophication.html.

United States Geological Survey (USGS) 2013a. Continuous Nitrate Monitoring in Rivers. Accessed October 24, 2013 at http://water.usgs.gov/nasqan/docs/RT-BriefingSheet.pdf

United States Geological Survey (USGS) 2013b. *USGS Current Conditions for the Nation*. http://waterdata.usgs.gov/usa/nwis/uv?site_no=01493112

United States Geological Survey (USGS) 2013c. *USGS Current Conditions for the Nation*. Retrieved from http://waterdata.usgs.gov/usa/nwis/uv?06893390

United States Geological Survey (USGS) 2013d. USGS Water Data for the Nation. Accessed October 24, 2013 at http://waterdata.usgs.gov/usa/nwis/.

United States Geological Survey (USGS) 2014. National Hydrography Dataset (NHD). United States Geological Survey, Reston, VA.

United States Geological Survey (USGS) 2015a. The Gulf of Mexico Hypoxic Zone . Accessed March 17, 2015. http://toxics.usgs.gov/hypoxia/hypoxic_zone.html

United States Geological Survey (USGS) 2015b. Water properties: Dissolved oxygen. Accessed March 17, 2015. http://ga.water.usgs.gov/edu/dissolvedoxygen.html

United States Geological Survey (USGS) 2015c. Groundwater Quality. Accessed March 17, 2015.   http://ga.water.usgs.gov/edu/earthgwquality.html

Vidon, P., H. Hubbard, P. Cuadra, and M. Hennessy. 2012. Storm NO_3^- and NH_4^+ Exports in Stream, Overland Flow, and Tile Drains of the US Midwest. Annals of Environmental Science, 6(1), 5.

Vitousek, P.M., J.D. Aber, R.W. Howarth, G.E. Likens, P.A. Matson, D.W. Schindler, W.H. Schlesinger, and D.G. Tilman. 1997. Human alteration of the global nitrogen cycle: sources and consequences. *Ecological applications*, *7*(3), 737-750.

Wahl, K.L. and T.L. Wahl. 1995. Determining the Flow of Comal Springs at New Braunfels, Texas. *Texas Water '95*. American Society of Civil Engineers, August 16-17, 1995, San Antonio, Texas, pp. 77-86.

Wahl, T.L. and K.L. Wahl. 2007. BFI: A Computer Program for Determining an Index to Base Flow (Version 4.15), [Software]. http://www.usbr.gov/pmts/hydraulics_lab/twahl/bfi/ Accessed on 4 November 2013.

Wang, T. and B. Zhu. 2011. Nitrate loss via overland flow and interflow from a sloped farmland in the hilly area of purple soil, China, Nut. Cyc. In Agroecosystems, 90(3), 309-319. doi: 10.1007/s10705-011-9431-7

Wilde, F.D., D.B. Radtke, J. Gibs, and R.T. Iwatsubo. 2004 with updates through 2009. Processing of water samples (version 2.2). U.S. Geological Survey Techniques of Water-Resources Investigations. Book 9, chap. A5, Accessed February 13, 2011, at http://pubs.water.usgs.gov/twri9A5/.

Wise, D.R., and H.M. Johnson. 2011. Surface-Water Nutrient Conditions and Sources in the United States Pacific Northwest1. JAWRA Journal of the American Water Resources Association, 47(5), 1110-1135.

Woodward, S.J., R. Stenger, and V.J. Bidwell. 2013. Dynamic analysis of stream flow and water chemistry to infer subsurface water and nitrate fluxes in a lowland dairying catchment. *Journal of Hydrology*, *505*, 299-311, doi:10.1016/j.jhydrol.2013.07.044

Write, M.J. and K.L. Davidson. 1964. Nitrate accumulation in crops and nitrate poisoning in animals. Advances in Agronomy. 16:197-247.

Appendix A: Nitrogen: a summary of environmental transformations and environmental concerns

This appendix contains a summary of the environmental transformation of nitrogen between many of its chemical forms, as well as the environmental concerns relating to each form of nitrogen presented.

<u>Nitrogen: a summary of environmental transformations and environmental concerns</u>

Nitrogen is an essential nutrient for plant growth and development. In modern agriculture, nitrogen is often applied, oftentimes in the form of ammonia, urea, nitrate, or organic nitrogen (manure), to maintain high-yielding crops and soil fertility. The applied forms are transformed through natural processes into other forms with changes in the environmental conditions and the passage of time. Each of these forms of nitrogen is a different chemical, but the nitrogen element is always the same. Together, these processes allow nitrogen to cycle among its various forms and guarantees that no single form becomes a "dead end" product that accumulates in the environment.

Figure A1.1 illustrates the cycle of nitrogen, showing the transformations between its various chemical forms. The transformations are induced by chemical reactions, microbiological reactions, plant metabolism, combustion, and/or industrial reactions. Industrial processes are used to accelerate natural transformation processes to produce the desired nitrogen forms. Ammonia is produced in large quantities from atmospheric nitrogen gas for use as fertilizer through the Haber-Bosch process (Kramer, 2000). Some forms of nitrogen are temporarily formed during a transformation reaction, while others have a longer lifetime in the environment. These non-transient forms are distributed between the air, water, and biomass, depending upon their chemical properties and the environmental conditions. Figure A1.1 indicates the environmental compartment that each form of nitrogen is commonly observed in its highest concentrations.

Not all of the nitrogen that is applied to an agricultural field as fertilizer is used by the crop. The unused nitrogen, in one of its many forms, can leave the agricultural field and can cause environmental concerns when it enters areas of the environment where it

was not initially applied. Figure A1.2 connects the different forms of nitrogen to a wide range of environmental concerns, with each stable form of nitrogen having its own set of environmental concerns. Most forms of nitrogen are connected to many different environmental concerns. Some of these are direct connections. For example, nitrate is toxic to infants younger than six months of age, in which it can cause methemoglobinemia, or "blue-baby" syndrome (USGS, 2012), and to some livestock — since they are not able to metabolize it. Other connections between nitrogen forms and environmental concerns are indirect. Nitrate, when introduced to a stream at concentrations exceeding the ability of the stream to assimilate it, acts as a fertilizer and increases the growth of aquatic plants (USGS, 2011). This process, known as eutrophication, creates a number of negative cascading effects, some representing permanent changes to the ecosystem and species composition. Because many streams and rivers flow into lakes, estuaries, and oceans, the excess nitrogen carried into these water bodies may cause them to undergo eutrophication far from the original source of the nutrient.

Bibliography

Anderson, D.M., P.M. Glibert, and J.M. Burkholder. 2002. Harmful Algal Blooms and Eutrophication Nutrient Sources, Composition, and Consequences. Estuaries, 25(4b):704-726.

Botkin, D.B. and E.A. Keller. 2007. Environmental Science: Earth as a living planet. Sixth edition. John Wiley and Sons Inc. Hoboken, New Jersey.

Bunce, N.J. 1993. Introduction to Environmental Chemistry. Wuerz Publishing Ltd. Winnipeg, Canada.

Camargo, J.A. and A. Alonso. 2006. Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment. Environment International. 32(6):831–849. doi:10.1016/j.envint.2006.05.002

Canfield, D.E., A.N. Glazer, and P.G. Falkowski. 2010. The evolution and future of earth's nitrogen cycle. Science. 330(6001):192-196. DOI: 10.1126/science.1186120

Cunningham, W.P., M.A. Cunningham, and B. Saigo. 2003. Environmental Science: A
global concern. 7th edition. McGraw Hill. New York. ISBN-13: 978-
0073532547

Hooper, A.B., T. Vannelli, D.J. Bergmann, and D.M. Arciero. 1997. Enzymology of the
oxidation of ammonia to nitrite by bacteria. Antonie van Leeuwenhoek. 71(1-
2):59-67. doi: 10.1023/A:1000133919203.

Kadlec, R. and R.L. Knight. 1996. Treatment wetlands. Lewis Publishers. Boca Raton,
Florida.

Kapoor, A., and T. Viraraghavan. 1997. Nitrate removal from drinking water—review. J.
Environ. Eng. 123:371–380.

Kramer, D.A. 2004. Nitrogen. U.S. Geological Survey Minerals Yearbook. Accessed
March 17, 2015.
minerals.usgs.gov/minerals/pubs/commodity/nitrogen/nitromyb04.pdf

Marsh, K.L., G.K. Sims, and R.L. Mulvaney. 2005. Availability of urea to autotrophic
ammonia-oxidizing bacteria as related to the fate of 14C- and 15N-labeled urea
added to soil. Biology and Fertility of Soils, 42(2):137-145. doi: 10.1007/s00374-
005-0004-2

McGee, E. 1997. Acid Rain and Our Nation's Capital. U.S. Geological Survey. Accessed
March 17, 2015. http://pubs.usgs.gov/gip/acidrain/5.html

Meessen, J.H. 2010. Urea in Ullmann's Encyclopedia of Industrial Chemistry, Wiley-
VCH, Weinheim. Accessed March 17, 2015. DOI:
10.1002/14356007.a27_333.pub2

Mueller, D.K. and D.R. Helsel. 1996. Nutrients in the Nation's Waters--Too Much of a
Good Thing? U.S. Geological Survey Circular 1136, 24 p. Accessed March 17,
2015. http://pubs.usgs.gov/circ/circ1136/

Murakami K. 1984. Dredging for controlling eutrophication of lake Kasumigaura, Japan.
Lake and Reservoir Management 1(1): 592-598.

National Oceanic and Atmospheric Administration. 2009. NOAA Study Shows Nitrous
Oxide Now Top Ozone-Depleting Emission. Accessed March 17, 2015.
http://www.noaanews.noaa.gov/stories2009/20090827_ozone.html

National Oceanic and Atmospheric Administration. 2015. Greenhouse gases. Accessed
March 17, 2015. http://lwf.ncdc.noaa.gov/oa/climate/gases.html

Ricklefs, R.E. 2001. The Economy of Nature. 5th edition. W. H Freeman and Company. New York. ASIN: B003Q6EIBM

Sharpley, A.N., T. Daniel, T. Sims, J. Lemunyon, R. Stevens, and R. Parry. 2003. Agricultural Phosphorus and Eutrophication, Second Edition, Agricultural Research Service ARS–149, 38 p. Accessed March 17, 2015. http://www.ars.usda.gov/is/np/Phos&Eutro2/agphoseutro2ed.pdf

Starr, C., R. Taggar, C. Evers, and L. Starr. 2015. Biology: The unity and Diversity of Life. 14th edition. Cengage Learning. 1024 pp. ISBN-13: 978-1305073951

U.S. Environmental Protection Agency. 2004. Photochemical smog—what it means for us. USEPA Publication90/04 (March 2004). Accessed March 17, 2015. http://www.epa.sa.gov.au/xstd_files/Air/Information%20sheet/info_photosmog.pdf

U.S. Environmental Protection Agency. 2014. Nitrogen Dioxide. Accessed March 17, 2015. http://epa.gov/airquality/nitrogenoxides/

U.S. Environmental Protection Agency. 2015. Effects of Acid Rain - Surface Waters and Aquatic Animals. Accessed March 17, 2015. http://epa.gov/acidrain/effects/surface_water.html

U.S. Environmental Protection Agency. 2015. Hypoxia: What is hypoxia and what causes it? Accessed March 17, 2015. http://water.epa.gov/type/watersheds/named/msbasin/hypoxia101.cfm

U.S. Environmental Protection Agency. 2015. Dissolved Oxygen. Accessed March 17, 2015. http://www.epa.gov/caddis/ssr_do_wtl.html

U.S. Environmental Protection Agency. 2015. Overview of Greenhouse Gases. Accessed March 17, 2015. http://www.epa.gov/nitrousoxide/scientific.html

U.S. Geological Survey. 2011. Eutrophication. Accessed March 17, 2015. http://toxics.usgs.gov/definitions/eutrophication.html.

U.S. Geological Survey. 2015. The Gulf of Mexico Hypoxic Zone . Accessed March 17, 2015. http://toxics.usgs.gov/hypoxia/hypoxic_zone.html

U.S. Geological Survey. 2015. Water properties: Dissolved oxygen. Accessed March 17, 2015. http://ga.water.usgs.gov/edu/dissolvedoxygen.html

U.S. Geological Survey. 2015. Groundwater Quality. Accessed March 17, 2015. http://ga.water.usgs.gov/edu/earthgwquality.html

Write, M.J. and K.L. Davidson. 1964. Nitrate accumulation in crops and nitrate poisoning in animals. Advances in Agronomy. 16:197-247.
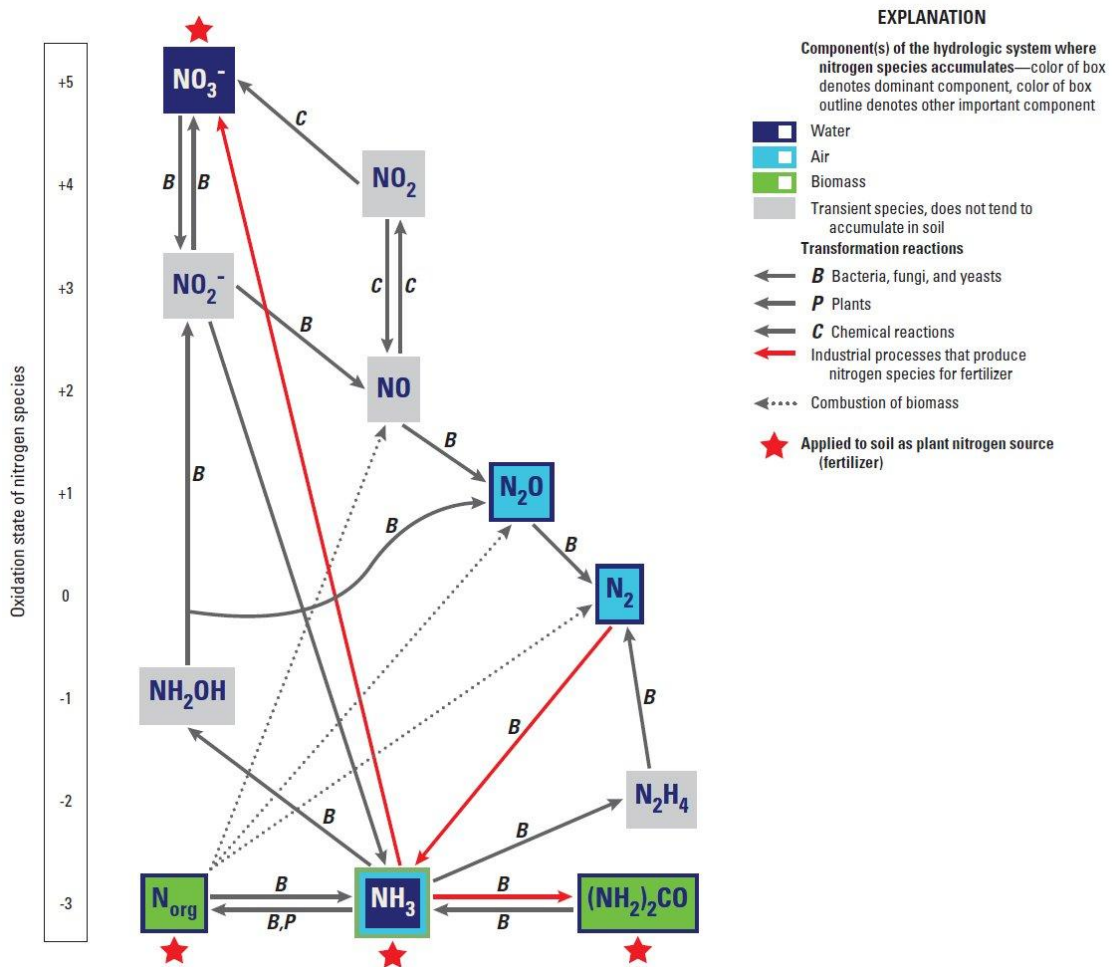
Figure A1.1: The cycle of nitrogen in the agricultural environment showing the transformations between its various chemical forms (boxes) and the transformation pathways that connect them (arrows). The chemical forms of nitrogen are arranged in rows by their oxidation state.
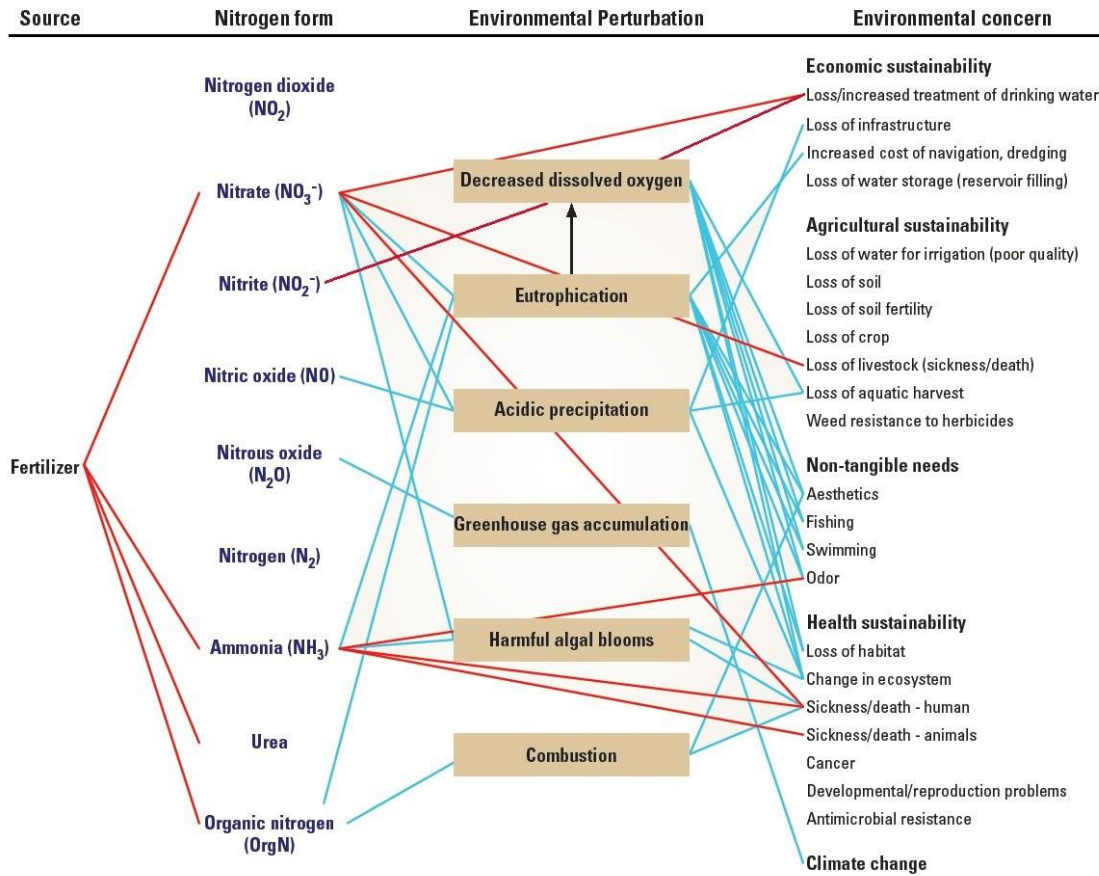
Figure A1.2: The complex connections among nitrogen from fertilizer and various water quality and other environmental concerns. The lines show the many potential pathways that connect the sources of nitrogen used in agriculture with the many different concerns. A realized concern has an adverse effect on economic or agricultural sustainability, aesthetics and recreation, human and ecosystem health, or climate change. The nitrogen which is applied as fertilizer, both chemical and manure, undergoes transformation through chemical and biological reactions to different forms of nitrogen. Each of these forms of nitrogen interacts with the environment in a different way and through different processes. Some forms of nitrogen can cause direct adverse effects on the environment. (Shown by direct lines from a nitrogen form to an environmental concern.) Some of the nitrogen forms can cause environmental perturbations (changes in the condition of the environment), which then can then produce adverse effects on the environment.

Appendix B: Supplemental information for Chapter 2: A comparison of high-resolution specific conductance-based end-member mixing analysis and a graphical method for baseflow separation of four streams in hydrologically challenging agricultural watersheds

This supplemental information provides the analysis of the effects of time interval between continuous SC measurements on the estimated long term average slowflow volume. Eleven different data collection time intervals (15 minutes, 30 minutes, 1 hour, 2 hours, 4 hours, 6 hours, 12 hours, 1 day, 2 days, 7 days, and 28 days) were tested using data records of three different lengths (1-week, 1-month, and 2-years). Results are displayed in Figure B2.1.

Also included are supplemental figures and tables (Figures B2.2 through B2.4 and Table B2.1). Streamflow and stream SC data for the four streams, along with the time-variable plots of SC_{SF} are shown in Figure B2.2. Figure B2.3 presents the comparison of the baseflow hydrograph from the BFI program to multiple slowflow hydrographs (each using different SC_{FF} input values) from SC-EMMA in Granger Drain, WA. Plots of daily slowflow estimates from SC-EMMA using a constant SC_{SF} against those using a time-variable SC_{SF} are presented in Figure B2.4 for all four streams. Table B2.1 shows the range of slowflow index (SFI) estimates resulting from all combinations of commonly used SC_{FF} and SC_{SF} values in each of the four streams.

Effect of time intervals between SC measurements on SC-EMMA results

The temporal resolution (interval between observations) of the SC input data was examined to determine its effect on the calculated SFI for the 2 years of available data at each watershed. To test the impact of lower resolution continuous data, separate calculations of SFI were made after systematically removing specific data points, representing a more coarsely sampled continuous data set. Beginning with the 15-minute data (highest resolution), every other data point was removed to represent 30-minute continuous data. This method of systematic removal of data points was done to produce 11 separate data sets for each stream with data intervals of 15 minutes, 30 minutes, 1 hour, 2 hours, 4 hours, 6 hours, 12 hours, 1 day, 2 days, 7 days, and 28 days. Values for SC_{SF} and SC_{FF} (Table B2.1) were input into equations 3 and 4 to calculate the slowflow

volume and SFI during the 2 years of data for each of the 11 data sets. The length of the data record was also varied for each watershed to determine the effect of data resolution on SFI at differing lengths of record. Three lengths of data record were examined: the entire record (2 years, bi-annual), a 1-month period (the first June – mid-growing season), and a 1-week period (beginning on June 5 during the first year of record at each site). Each length of record was examined separately, but treated similarly by creating separate data sets of progressively greater data collection intervals, as described above.

The time interval between measurements of SC data had essentially no effect on the bi-annual SFI for intervals of 1 day or less. As shown in Figure B2.1A, the calculated SFI for the 2 years for each watershed remained nearly constant ($\pm$ 1%) for intervals ranging from 15 minutes through 720 minutes (12 hours). For an interval of 1 day, the variation was $\leq \pm$ 5% for all watersheds. For data intervals greater than 1 day, the bi-annual SFI diverged substantially from the bi-annual SFIs calculated using the shorter data intervals in all watersheds except for Granger Drain, WA.

The variability in the calculated SFI values for periods of 1 month and 1 week followed a similar pattern as a 2-year period. For a 1-month period, the calculated SFI remained constant in all watersheds ($\pm$ 1%) at SC measurement intervals of $\leq$ 720 minutes (12 hours, Figure B2.1B). When the period was shortened to 1 week for calculating SFI, the same results were obtained (Figure B2.1C).

Based on the four streams, a daily interval for SC measurements in the stream is sufficient to calculate bi-annual SFI by SC-EMMA. In the streams studied, there was little added value in using 15-minute SC intervals compared to 1-day SC intervals when calculating a bi-annual SFI. As a result, the amount of data that is necessary can be reduced, assuming that the goal is to determine the SFI of a stream. However, at time intervals greater than one observation per day, the calculated bi-annual SFI shows increased variability. Using coarser data also may result in loss of extreme data points, which could alter the values chosen for SC_{FF} and SC_{SF}, and change the calculated SFI. When calculating the SFI for a shorter time period (1 month and 1 week in this case), using a 12-hour data collection interval was as effective as the 15-minute data in the watersheds of this study, even in a flashy stream such as Tommie Bayou, MS.

Table B2.1: The range of potential SFI values from SC-EMMA based on all common combinations of input values (SC$_{FF}$ and SC$_{SF}$) for four agricultural watersheds.

Potential input values for SC$_{FF}$ (µS/cm)	Morgan Creek, MD		Tommie Bayou, MS		South Fork Iowa River (SFIR), IA		Granger Drain, WA	
	SC$_{FF}$ input	SFI % (lowest to highest SC$_{SF}$ input)	SC$_{FF}$ input	SFI % (lowest to highest SC$_{SF}$ input)	SC$_{FF}$ input	SFI % (lowest to highest SC$_{SF}$ input)	SC$_{FF}$ input	SFI % (lowest to highest SC$_{SF}$ input)
SC$_{FF}$ (values used in this study)	41	78 to 51	35	69 to 23	67	104 to 64	149	58 to 42
Median rainfall SC	21	81 to 55	10.3	71 to 25	10.6	103 to 66	5.4	68 to 52
SC during highest flow	80	88 to 38	156	53 to 13	129	104 to 61	335	33 to 20
Lowest stream SC	61	74 to 45	60	66 to 21	123	104 to 61	293	41 to 27
Median overland flow SC	No data	NA	No data	NA	81 [1]	104 to 63	339 [2]	32 to 20
Irrigation water SC	NA	NA	1190	114 to 988	NA	NA	99 [3]	62 to 46
Potential input values for SC$_{SF}$ (µS/cm)	**SC$_{SF}$ input**	**SFI % (lowest to highest SC$_{FF}$ input)**	**SC$_{SF}$ input**	**SFI % (lowest to highest SC$_{FF}$ input)**	**SC$_{SF}$ input**	**SFI % (lowest to highest SC$_{FF}$ input)**	**SC$_{SF}$ input**	**SFI % (lowest to highest SC$_{FF}$ input)**
SC$_{SF}$ (Highest stream SC, values used in this study)	236	55 to 38	904	30 to 318	836	66 to 61	842	52 to 20
Median groundwater SC	168	81 to 68	1098	25 to 988	793	70 to 65	727	60 to 26
SC during lowest flow	184*(185#)	73 to 58 (73 to 57)	392*(672#)	71 to 114 (41 to 176)	541*(574#)	103 to 104 (97 to 97)	674*(688#)	65 to 30 (63 to 29)

1 Range 33 to 395 µS/cm
2 Range 118 to 641 µS/cm
3 Range 79 to 126 µS/cm
* median of all SC values if there were multiple SC measurements during times of equal flow
# highest SC value if there were multiple SC measurements during times of equal flow
NA: not applicable

(A) 2-year SFI
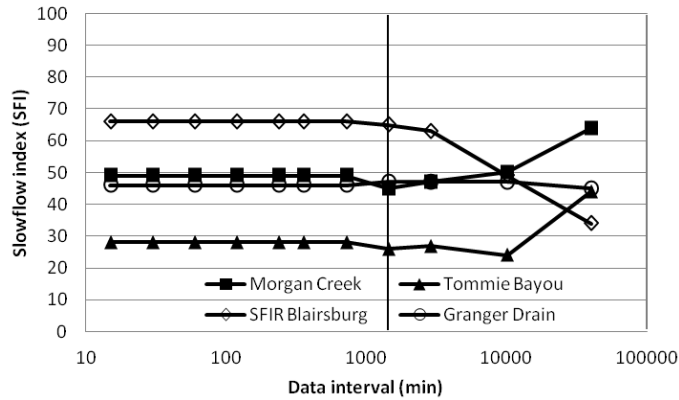
(B) 1-month SFI
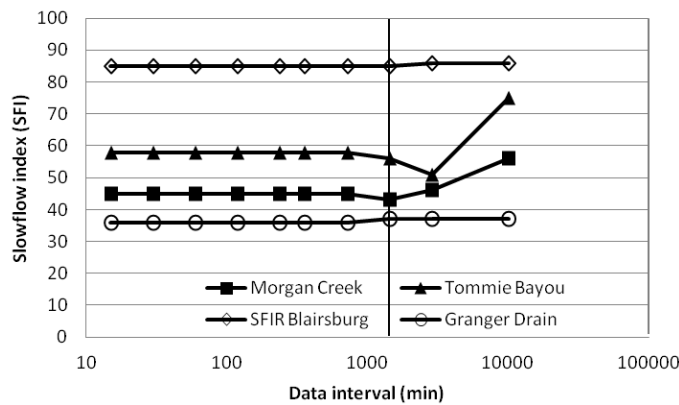
(C) 1-week SFI
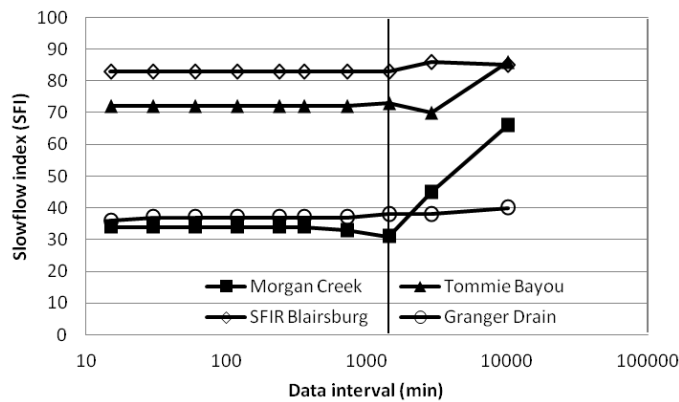
Figure B2.1: Slowflow index (SFI, in percent) for four agricultural streams for **(A)** a 2-year, **(B)** a 1-month, and **(C)** a 1-week time period calculated using various time intervals between the SC inputs values for SC-EMMA. The vertical line represents the calculated SFI based on data resolution of one data point per day. Input values (SC_{FF} and SC_{SF}) for SC-EMMA are shown in Table 2.1 of the main text.

166

(A) Morgan Creek, MD
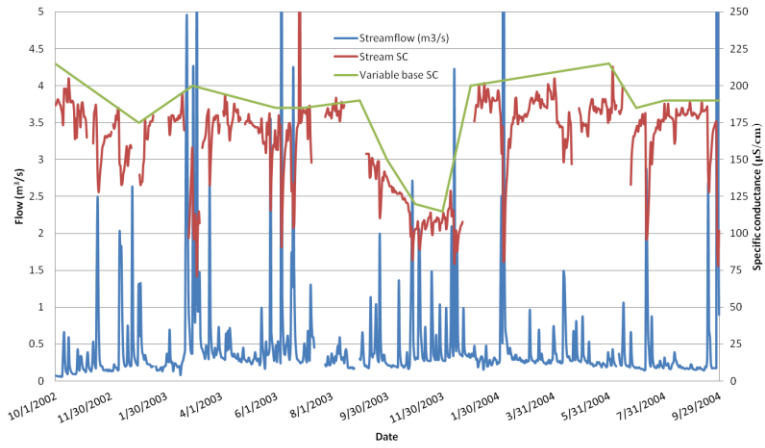
(B) Tommie Bayou, MS
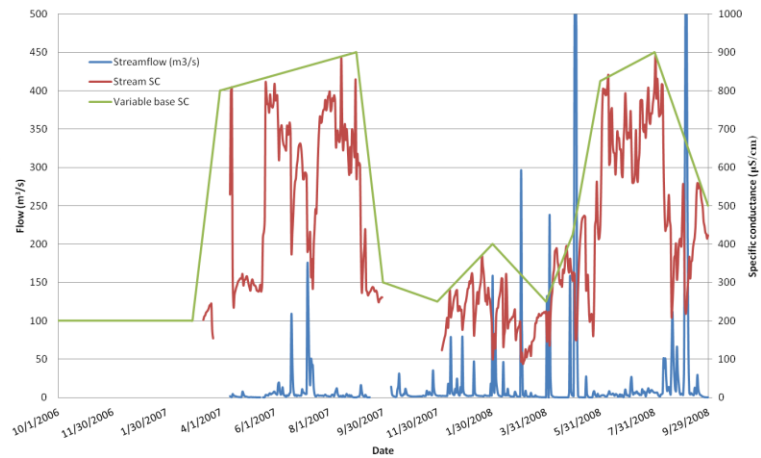
(C) South Fork Iowa River (SFIR), IA
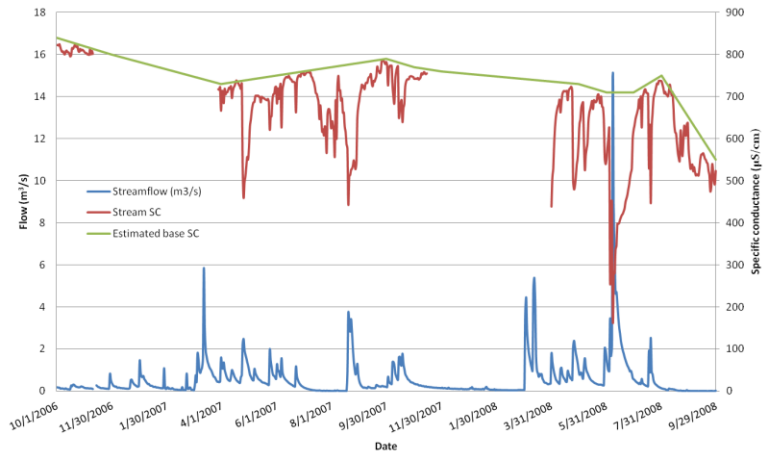
(D) Granger Drain, WA
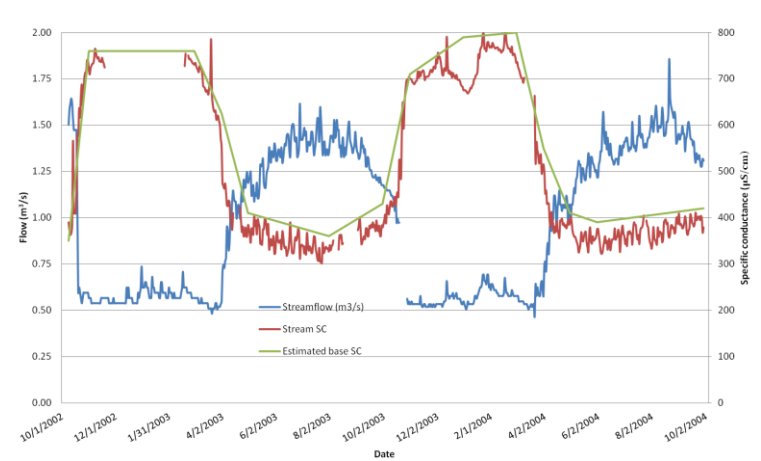
Figure B2.2: Streamflow and specific conductance records from four agricultural streams with the time-variable SC_{SF} values.
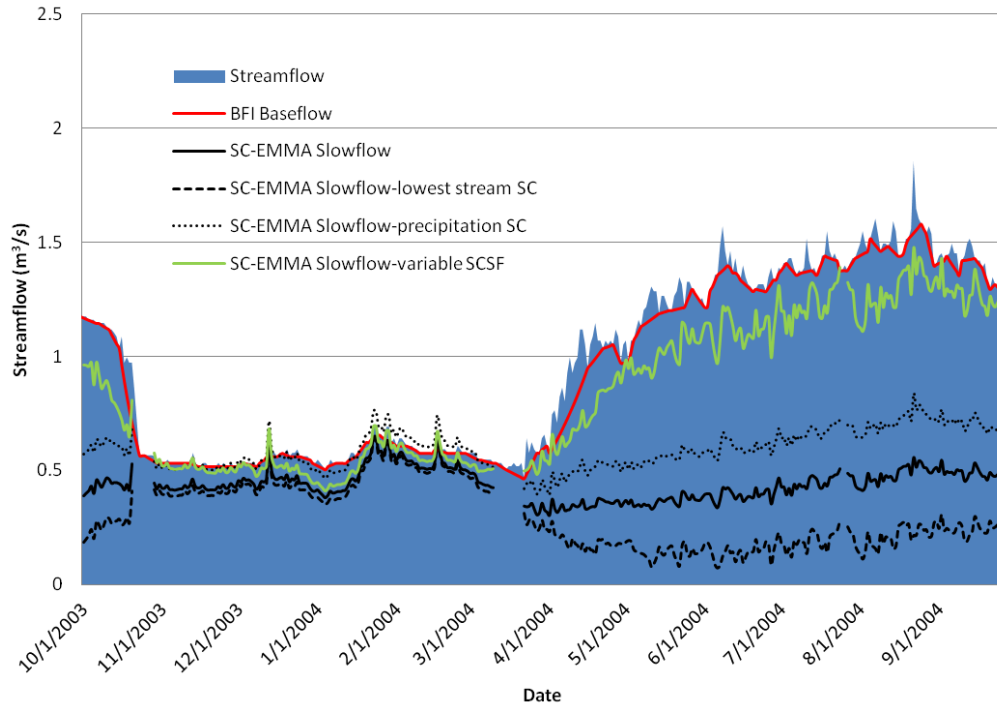
Figure B2.3: Total streamflow hydrograph compared to the baseflow/slowflow hydrographs calculated by the BFI program and SC-EMMA (using multiple potential inputs) for Granger Drain, WA (only one water year displayed). Gaps in the hydrograph represent no data.

(A) Morgan Creek, MD
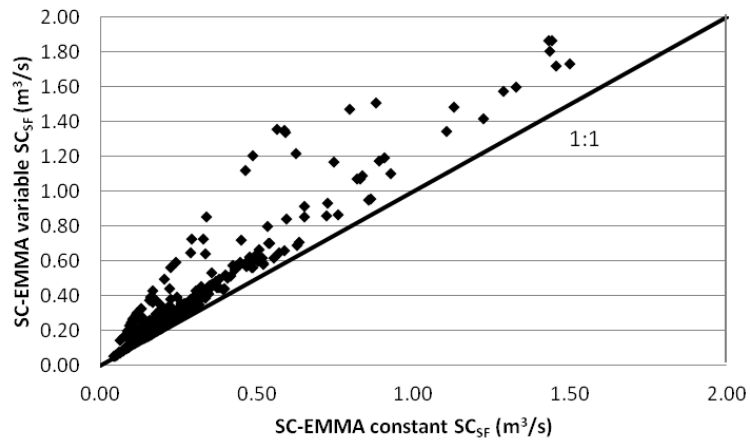
(B) Tommie Bayou, MS
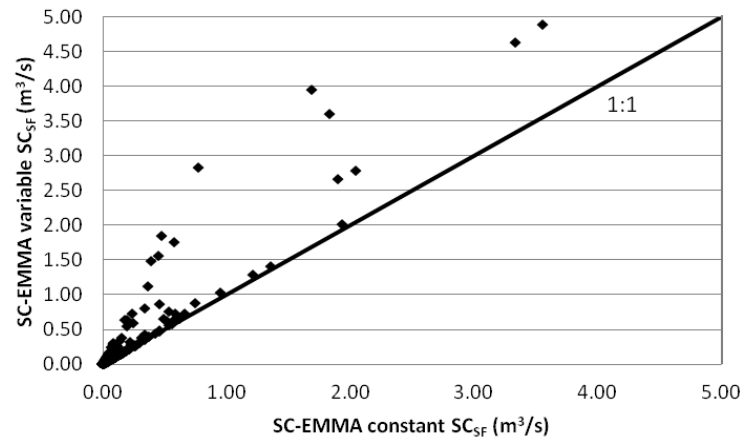
(C) South Fork Iowa River (SFIR), IA
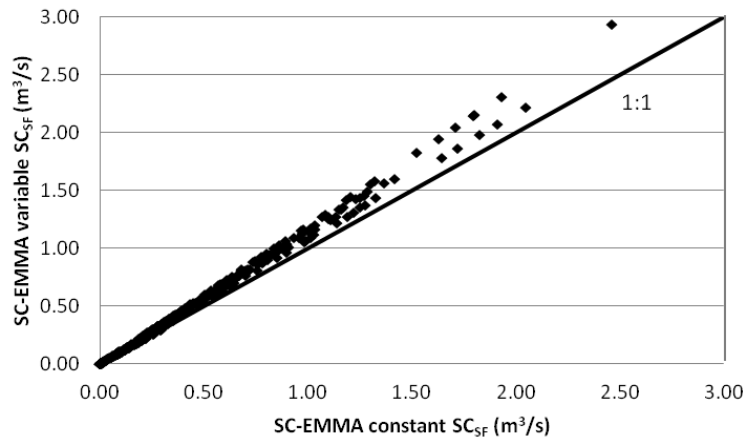
(D) Granger Drain, WA
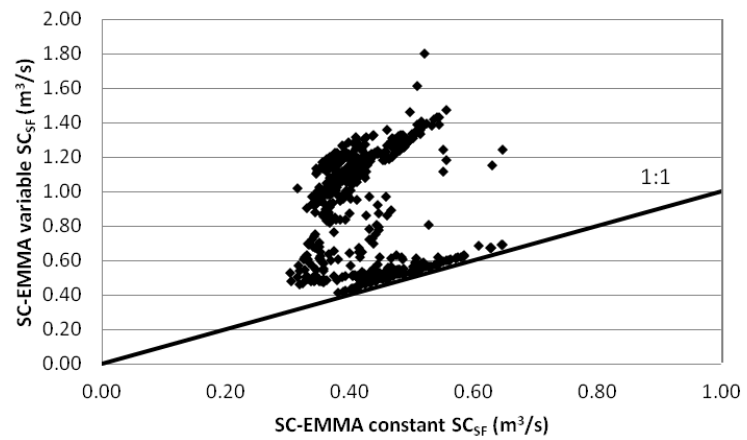
Figure B2.4: SC-EMMA slowflow (m^3/s) calculated using a constant SC_{SF} and a variable SC_{SF}.

Appendix C: Supplemental information for Chapter 3: Estimation of time-variable fast flowpath end-member concentrations for application in chemical hydrograph separation analyses

This supplemental information section contains additional results from the sensitivity analysis of the ratio-EMMA to the temporal relation between streamflow peak and minimum or maximum concentrations in fastflow (Figures C3.1 and C3.2), sensitivity analysis to a user defined range of ratios (Table C3.1), and a discussion regarding the choice of a range of ratios.

Sensitivity analysis of the results to the streamflow peak and max/min concentrations in fastflow

Figures C3.1A-D and C3.2A-D show the ratio-EMMA results compared to the actual concentration and flow values under the conditions where the streamflow peak and minimum $[A]_{FF}$ occurred concurrently, but minimum $[B]_{FF}$ occurred 24 hours later ($t_{Qmax}=t_{[A]min}<t_{[B]min}$); the streamflow peak occurred, but and the minimum $[A]_{FF}$ and $[B]_{FF}$ occurred 24 hours earlier ($t_{Qmax}>t_{[A]min}=t_{[B]min}$); streamflow peak occurred, but and the minimum $[A]_{FF}$ and $[B]_{FF}$ occurred 24 hours later ($t_{Qmax}<t_{[A]min}=t_{[B]min}$); and minimum $[A]_{FF}$ occurred, then streamflow peak occurred 24 hours later, then the minimum $[B]_{FF}$ occurred 24 hours after that ($t_{[A]min}< t_{Qmax}<t_{[B]min}$).

Scenarios in which the timing of the minimum $[A]_{FF}$ occurred concurrently with that of $[B]_{FF}$

In all of the scenarios where chemicals A and B reached simultaneous minimum concentrations in fastflow, during peakflow (>100 m^3/s) the ratio-EMMA results were in close agreement with the known concentrations. The combined median absolute deviation from actual fastflow concentrations during peakflow in both streams was ± 0.04 mg N/L (18%) for chemical A and ± 4.0 µS/cm (29%) for chemical B. There was slightly more discrepancy among the concentrations at low flows (<15 m^3/s), with median absolute deviation from actual concentrations of ± 0.25 mg N/L (15%) for chemical A and ± 29 µS/cm (24%) for chemical B. Percent error was greater among concentration

estimates during peakflow as a result of the actual concentration being low during peak flow.

Scenarios in which the timing of the minimum $[A]_{FF}$ diverged from that of $[B]_{FF}$

In all of the scenarios where chemicals A and B reached minimum concentrations in fastflow at different times, the ratio-EMMA results were in closest agreement with the actual concentrations at peakflow, having a median absolute deviation of $\pm$0.05 mg/L (30%) for chemical A and $\pm$5.3 µS/cm (43%) for chemical B. There was more discrepancy at lower flows, with median absolute deviation from actual concentrations of $\pm$0.32 mg N/L (19%) for chemical A and $\pm$43 µS/cm (37%) for chemical B. As with the scenarios in which the timing of the minimum $[A]_{FF}$ occurred concurrently with that of $[B]_{FF}$, percent error was larger during peakflow than in lowflow due to the smaller actual concentrations during peakflow.

Scenarios in which the timing of the maximum $[A]_{FF}$ and minimum $[B]_{FF}$ occurred concurrently

During peakflow in the streams, the median absolute deviation from actual fastflow concentrations was 0.38 mg/L (4.0%) for chemical A and 37 µS/cm (340%) for chemical B, whereas during lowflow in the streams, the median absolute deviation was 1.3 mg/L (30%) for chemical A and 283 µS/cm (230%) for chemical B.

Sensitivity analysis of the results to a use defined fastflow ratio of $[A]_{FF}$ to $[B]_{FF}$

When defining the range of ratios of $[A]_{FF}$ to $[B]_{FF}$ based on knowledge of the watershed where the ratio-EMMA is applied, and not using the iterative approach to ratio selection, the choice of the range of ratios affects the results of the model. Therefore, the sensitivity of the model to a user defined range of ratios was examined using a wide variety of possible ratios. The sensitivity, as a function of the magnitude and range of the defined ratio of $[A]_{FF}$ to $[B]_{FF}$ was evaluated using the synthetic stream data that were

meant to represent a stream with important additions of groundwater and a stream with important additions of overland flow water. The scenario chosen for both streams was one in which the streamflow peak occurred concurrently with the minimum $[A]_{FF}$ and $[B]_{FF}$ ($t_{Qmax}=t_{[A]min}=t_{[B]min}$).

When the model was run forward (Figure 3.1A-I of the main text), the results were bounded by the upper (larger) extent of the defined ratio, and the model commonly found the lowest fastflow concentrations that fell near that ratio. When the model was run in reverse, the lower (smaller) end of the defined ratio limited the results and commonly led to the highest possible fastflow concentration estimates. Numerous ranges of ratios were tested, some of which did not restrict the limits of the range (unbounded) while others set limits on one or both ends of the range of ratios (bounded).

Unbounded/Unbounded

When the model was run (forward and reverse, taking the mean of the results) with a range of ratios in fastflow that is essentially unbounded on either end (0.0000 to 1000), the model found the lowest and the highest possible fastflow concentration estimates during each two-hour time interval, and therefore defined the limits of all possible fastflow concentrations that could be produced by the model. Those extreme values were then averaged during each two hour time interval to produce the concentration estimates for that interval.

For both synthetic streams, when the ratio was unbounded on both ends, the ratio-EMMA model estimates of fastflow (Q_{FF}) diverged from the actual results to a large extent (Table C3.1). Bi-hourly fastflow estimates differed by a median absolute deviation of ± 0.70 m^3/s (62%) in the groundwater dominated synthetic stream and ± 2.4 m^3/s (25%) in the overland flow dominated synthetic stream. Total fastflow volume had a percent error of 40% in the groundwater dominated synthetic stream and 7.9% in the overland flow dominated stream (Table C3.1). With an unbounded range of ratios, the model produced estimates of slowflow index (SFI, Equation 5 in the main text) that were less than the actual values for both synthetic streams (Table C3.1). The difference

172

between the actual SFI values and the ratio-EMMA estimated values was -13% for the groundwater dominated synthetic stream and was -6.1% when used with the overland flow dominated synthetic stream.

The median absolute deviation between the ratio-EMMA estimated fastflow concentration and the actual values in the groundwater dominated synthetic stream was ± 2.1 mg/L (240%) for chemical A and ± 280 μS/cm (400%) for chemical B and was ± 1.2 mg/L (120%) for chemical A and ± 150 μS/cm (190%) for chemical B in the overland flow dominated synthetic stream.

Chemical A loads in fastflow were estimated for the entire length of record. The percent error for the total load estimate was 1700% in the groundwater dominated synthetic stream and 240% in the overland flow dominated synthetic stream (Table C3.1). Chemical B loads (calculated by converting SC (μS/cm) to total dissolved solids (TDS in mg/L); TDS ~ 0.64*SC by combining the Russell Equation and Langlier approximation) in fastflow were also estimated for the entire length of record. The percent error for the load estimate was 2600% in the groundwater dominated synthetic stream and 370% in the overland flow dominated synthetic stream (Table C3.1).

The deviation between the actual and estimated values in each of these synthetic streams suggest that the high concentration estimates that result from the lower limit of the pre-defined ratio skew the results. In this case, the model does not produce meaningful results.

Bounded/Unbounded

Fastflow estimates produced by running the model with ranges of ratios that were bounded on the low end (0.0121, 0.0128, and 0.0135) and unbounded on the high end (1000) were much more similar to the actual values than when the range of ratios was left unbounded on both ends (Table C3.1). Bi-hourly fastflow estimates differed by a median absolute deviation of ± 0.04 m^3/s (5.0%) in the groundwater dominated synthetic stream and ± 0.45 m^3/s (4.9%) in the overland flow dominated synthetic stream. Total fastflow volume had an average percent error of 0.70% in the groundwater dominated synthetic

stream and 1.0% in the overland flow dominated stream (Table C3.1). The model produced SFI estimates that were less than, but nearly equal to the actual values for both synthetic streams. The difference between the actual SFI values and the ratio-EMMA estimated values was <0.50% for the groundwater dominated synthetic stream and was <1.0% when used with the overland flow dominated synthetic stream (Table C3.1).

The median absolute deviation between the ratio-EMMA estimated fastflow concentration and the actual values in the groundwater dominated synthetic stream was ±0.33 mg/L (32%) for chemical A and ±40 μS/cm (49%) for chemical B, and was ±0.31 mg/L (30%) for chemical A and ±39 μS/cm (48%) for chemical B in the overland flow dominated synthetic stream.

The percent error for load estimates of chemical A had a mean of -30% in the groundwater dominated synthetic stream and -31% in the overland flow dominated synthetic stream. The percent error for chemical B load estimates had a mean of -43% among the scenarios in the groundwater dominated synthetic stream and -45% in the overland flow dominated synthetic stream for the three scenarios (Table C3.1).

Unbounded/Bounded

The results produced by running the model with a range of ratios that were unbounded on the lower end (0.0000), but were bounded on the high end (0.0121, 0.0128, and 0.0135) diverged substantially from the actual fastflow values (Table C3.1). Bi-hourly fastflow estimates differed by a median absolute deviation of ±0.91 m^3/s (93%) in the groundwater dominated synthetic stream and ±3.3 m^3/s (35%) in the overland flow dominated synthetic stream. Total fastflow volume had an average percent error of 42% in the groundwater dominated synthetic stream and 9.5% in the overland flow dominated stream (Table C3.1). Considering all of the tested ranges of ratios that were unbounded/bounded, the difference between the actual SFI values and the ratio-EMMA estimated values was -13% for the groundwater dominated synthetic stream and -7% in the overland flow dominated synthetic stream (Table C3.1).

The median absolute deviation between the ratio-EMMA estimated fastflow concentration and the actual values in the groundwater dominated synthetic stream was ±2.7 mg/L (280%) for chemical A and ±360 µS/cm (450%) for chemical B and was ±1.5 mg/L (150%) for chemical A and ±200 µS/cm (240%) for chemical B in the overland flow dominated synthetic stream.

The percent error for chemical A load estimates had a mean of 1800% in the groundwater dominated synthetic stream and 290% in the overland flow dominated synthetic stream. Chemical B load estimates had a mean percent error of 2700% among the scenarios in the groundwater dominated synthetic stream and 440% in the overland flow dominated synthetic stream for the three scenarios (Table C3.1).

The deviation, as with the results from the unbounded/unbounded analysis, was caused by the high concentration estimates of $[A]_{FF}$ and $[B]_{FF}$ due to the lower limit of the ratio, which skewed the results. The bounds placed on the high end of the ratio resulted in fastflow concentration estimates that were elevated in comparison to the results produced with an unbounded upper end of the range of ratios. When limiting the upper end of the ratio to a value of 0.0128, for example (the known ratio measured from the synthetic data), running the model forward produces concentration estimates that are equal to, or very nearly equal to the actual concentrations from the synthetic data. However, those estimates are then averaged with the estimates produced by running the model in reverse. Running the model in reverse produces the highest possible concentration estimates because the lower end of the ratio was unbounded. This results in an averaged concentration that is too high, producing fastflow concentration estimates which are much greater than the actual values.

Bounded/Bounded

Estimates of fastflow produced by running the model with ranges of ratios that were bounded on the low end (0.0121, 0.0114, 0.0100, and 0.0086), on the high end (0.0135, 0.0142, 0.0156, and 0.0170), and centered on the actual value were in good agreement with the actual values (Table C3.1). Bi-hourly fastflow estimates differed by a

median absolute deviation of ± 0.02 m^3/s (2.6%) in the groundwater dominated synthetic stream and ± 0.20 m^3/s (2.3%) in the overland flow dominated synthetic stream. Total fastflow volume had a mean percent error of 0.60% in the groundwater dominated synthetic stream and 0.70% in the overland flow dominated stream (Table C3.1). In both synthetic streams, the departure from the actual SFI was minimal ($\leq$ 0.1%) in the case where the range of ratios was defined as 0.0121 to 0.0135 (Table C3.1). These ratio values bracket the actual ratio of the chemicals in fastflow (0.0128). As the chosen range of ratios was widened (0.0114 to 0.0142) and then (0.0100 to 0.0156), the results again diverged only slightly, with a maximum discrepancy in SFI of 1.2% seen in the overland flow dominated synthetic stream. As the range of ratios widened more (0.0086 to 0.0170), the divergence between estimated and actual SFI values increased. In the groundwater dominated synthetic stream, the difference between the estimated and actual SFI increased slightly (-3.1%), while in the overland flow dominated synthetic stream the difference increased to a larger extent (-6.5%). At this range of ratios, the lower limit of the range (0.0086) was nearly equal to the actual ratio of $[A]_S$ to $[B]_S$ (0.0083) which was the lowest ratio that produced estimated values (when the ratio is left unbounded). Ratios of $[A]_{FF}$ to $[B]_{FF}$ less than $[A]_S$ to $[B]_S$ cannot characterize the mixture of waters in the stream and therefore the model does not produce results.

The median absolute deviation between the ratio-EMMA estimated fastflow concentration and the actual values in the groundwater dominated synthetic stream was ± 0.14 mg/L (15%) for chemical A and ± 18 µS/cm (23%) for chemical B, and was ± 0.13 mg/L (12%) for chemical A and ± 16 µS/cm (20%) for chemical B in the overland flow dominated synthetic stream.

The percent error for chemical A load estimates had a mean of 22% in the groundwater dominated synthetic stream and 19% in the overland flow dominated synthetic stream. Total load estimates for chemical B had a mean percent error of 32% among the scenarios in the groundwater dominated synthetic stream and 28% in the overland flow dominated synthetic stream for the three scenarios (Table C3.1).

Ranges of ratios that were bounded on both ends, but not centered on the actual ratio value were also tested. These "lopsided" ranges produced fastflow estimates that were in good agreement with actual values (Table C3.1). Bi-hourly fastflow estimates differed by a median absolute deviation of ±0.01 m^3/s (1.0%) in the groundwater dominated synthetic stream and ±0.07 m^3/s (0.88%) in the overland flow dominated synthetic stream. Total fastflow volume had a mean percent error of 0.20% in both the groundwater dominated synthetic stream and the overland flow dominated stream. SFI estimates were <0.3% different than the actual SFI in both synthetic streams (Table C3.1). The difference was greatest when the chosen range of ratios were smaller (0.0114 to 0.0135), resulting in elevated estimates of $[A]_{FF}$ and $[B]_{FF}$.

The median absolute deviation between the ratio-EMMA estimated fastflow concentration and the actual values in the groundwater dominated synthetic stream was ±0.06 mg/L (6.8%) for chemical A and ±6.7 µS/cm (10%) for chemical B, and was ±0.05 mg/L (5.1%) for chemical A and ±6.8 µS/cm (8.3%) for chemical B in the overland flow dominated synthetic stream.

Chemical A loads in fastflow had a mean percent error of 7.0% in both the groundwater dominated synthetic stream and the overland flow dominated synthetic stream. The percent error for total load estimates of chemical B had a mean of 10% among the scenarios in both the groundwater dominated synthetic stream and the overland flow dominated synthetic stream for the three scenarios (Table C3.1).

Choosing a primary ratio when applying a user defined range of ratios

The ratio-EMMA can be a more subjective model with opting for a user defined range of ratios. The choice of range is very important and will establish the accuracy and bias of the results. Although the full optimization of a user defined range of ratios is outside of the scope of this paper, general information is gained from the analyses performed.

A range of ratios could be generated based on known characteristics of the watershed, such as the ratio of the two chemicals in overland flow water. The model

could then be run and the estimated fastflow concentrations could be referenced against any known information about fastflow concentrations in the watershed. The concentrations estimated by the ratio-EMMA should likely be between the lowest measured concentrations in the stream and the chemical concentrations in precipitation (assuming an inverse relationship between the concentration of the chemicals in the stream and streamflow). This is in best agreement with expectations of fastflow concentration for many situations, as the true concentration of the chemicals is expected to be between those two values. The resulting fastflow estimates may not match point samples from within the watershed however, as the estimates produced by the ratio-EMMA model represent the aggregated concentration of the chemicals in fastflow across the entire watershed.

The chosen range of ratios is also a balance between precision/narrowness of the range and the number of values that the ratio-EMMA will not be able to produce. As the selected range of ratios gets smaller, the likelihood that the model will not be able to find a solution that meets all necessary criteria gets larger. A very narrow range of ratios is likely to result in many time intervals where the model cannot find a solution. A wider ratio will result in fewer or no non-values. If the model is unable to produce results (at peak flows for instance), this suggests that the selected ratio may not be representative of the ratio of $[A]_{FF}$ to $[B]_{FF}$. At peak flow, the stream will have large additions of water from fastflow sources. Because of this, the ratio of $[A]_{FF}$ to $[B]_{FF}$ will become more important and could drive the ratio in the stream. This could suggest that altering the ratio by selecting values that are more extreme than those measured in the stream could produce more realistic results for the ratio-EMMA during highflow.

Range of ratios		Groundwater dominated synthetic stream								Overland flow dominated synthetic stream							
		Fastflow ($x10^7$ m^3)†	% Error‡	SFI* (%)	Difference# (%)	Chemical A load ($x10^3$ kg)	% Error‡	Chemical B load ($x10^3$ kg)"	% Error‡	Fastflow ($x10^7$ m^3)†	% Error‡	SFI* (%)	Difference# (%)	Chemical A load ($x10^3$ kg)	% Error‡	Chemical B load ($x10^3$ kg)"	% Error‡
	Actual values	**2.0**		**68.3**		**3.4**		**170**		**4.9**		**23.1**		**11**		**574**	
	Unbounded (0) Unbounded (1000)	2.8	40	55.6	-19	61	1678	4578	2595	5.3	7.9	17.0	-26	39	236	2673	366
	Bounded (.0121) Unbounded (1000)	2.0	-0.6	68.5	0.3	2.6	-25	111	-35	4.9	-0.8	23.7	2.8	9	-26	359	-37
	Bounded (.0128) Unbounded (1000)	2.0	-0.7	68.6	0.3	2.4	-30	96	-43	4.9	-1.0	23.9	3.4	8	-31	311	-46
	Bounded (.0135) Unbounded (1000)	2.0	-0.8	68.6	0.4	2.2	-34	85	-50	4.9	-1.2	24.0	3.9	7	-36	272	-53
	Unbounded (0) Bounded (.0121)	2.9	42	54.9	-20	64	1768	4805	2728	5.4	9.8	15.6	-33	45	294	3165	451
	Unbounded (0) Bounded (.0128)	2.9	42	55.0	-20	63	1753	4765	2704	5.4	9.5	15.8	-32	44	284	3081	437
Range of ratios	Unbounded (0) Bounded (.0135)	2.9	42	55.1	-19	63	1742	4736	2688	5.4	9.3	16.0	-31	43	278	3023	427
	Bounded (.0121) Bounded (.0135)	2.0	-0.1	68.4	0.0	3.3	-3.1	166	-2.5	4.9	0.0	23.1	0.1	11	-1.6	569	-0.8
	Bounded (.0114) Bounded (.0142)	2.0	0.3	68.3	-0.1	3.8	11	197	16	4.9	0.3	22.9	-0.9	12	8.3	646	12
	Bounded (.0100) Bounded (.0156)	2.0	1.3	67.9	-0.6	5.2	52	305	79	5.0	1.5	21.9	-5.1	17	46	978	70
	Bounded (.0086) Bounded (.0170)	2.2	9.7	65.3	-4.5	17	404	1228	623	5.3	8.5	16.6	-28	41	255	2822	392
	Bounded (.0121) Bounded (.0142) - lopsided on the high side of the actual ratio	2.0	0.0	68.3	0.0	3.4	0.1	169	-0.8	4.9	0.0	23.1	0.2	11	-1.4	560	-2.5
	Bounded (.0114) Bounded (.0135) - lopsided on the low side of the actual ratio	2.0	0.3	68.2	-0.2	3.9	14	204	20	4.9	0.4	22.8	-1.3	13	12	675	18

Table C3.1: Comparison of the actual and ratio-EMMA estimates of fastflow volume, slowflow index (SFI), and chemical A and B loads resulting from the selection of different ranges of ratios. Selected ranges of ratios of [A]$_{FF}$ to [B]$_{FF}$ are shown in parentheses. Values were calculated assuming that streamflow peak and minimum [A]$_{FF}$ and [B]$_{FF}$ occur concurrently.
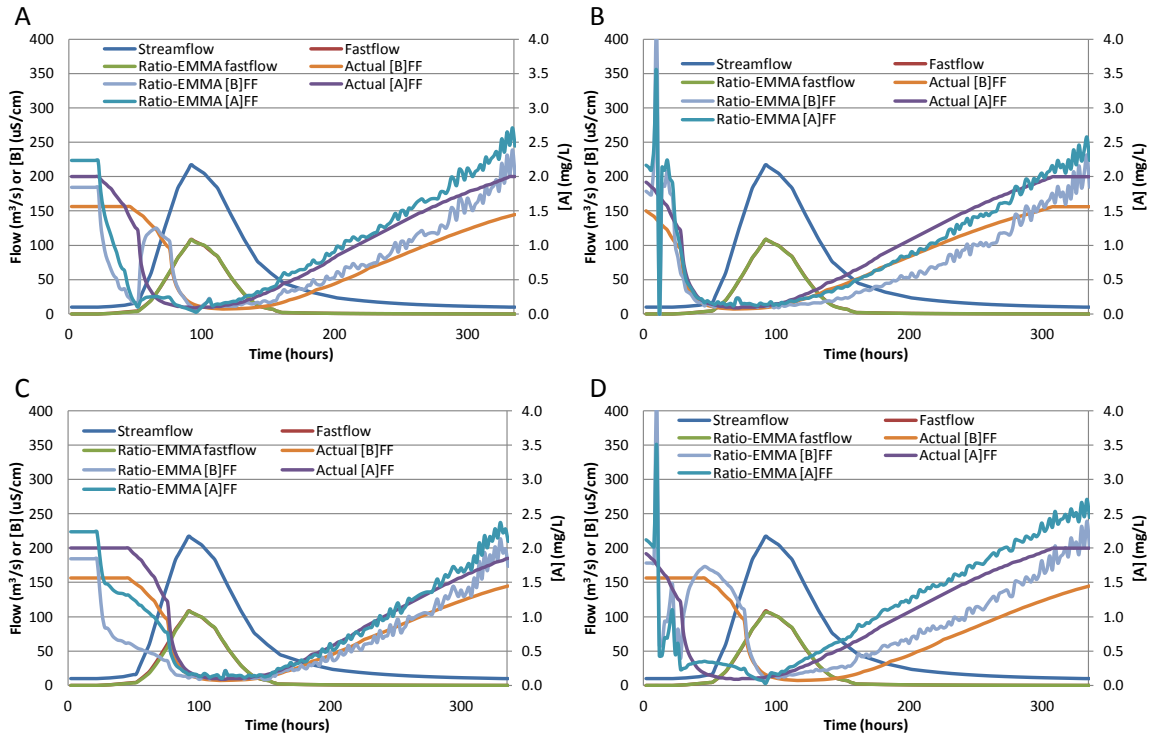
Figure C3.1: Ratio-EMMA estimated slowflow, $[A]_{FF}$, and $[B]_{FF}$ compared to actual values in the groundwater dominated synthetic stream where **(A)** the streamflow peak and minimum $[A]_{FF}$ occurred concurrently, but minimum $[B]_{FF}$ occurred 24 hours later ($t_{Qmax}=t_{[A]min}<t_{[B]min}$); **(B)** the streamflow peak occurred, but and the minimum $[A]_{FF}$ and $[B]_{FF}$ occurred 24 hours earlier ($t_{Qmax}>t_{[A]min}=t_{[B]min}$); **(C)** streamflow peak occurred, but and the minimum $[A]_{FF}$ and $[B]_{FF}$ occurred 24 hours later ($t_{Qmax}<t_{[A]min}=t_{[B]min}$); and **(D)** minimum $[A]_{FF}$ occurred, then streamflow peak occurred 24 hours later, then the minimum $[B]_{FF}$ occurred 24 hours after that ($t_{[A]min}< t_{Qmax}<t_{[B]min}$).

Figure C3.2: Ratio-EMMA estimated slowflow, $[A]_{FF}$, and $[B]_{FF}$ compared to actual values in the overland flow dominated synthetic stream where **(A)** the streamflow peak and minimum $[A]_{FF}$ occurred concurrently, but minimum $[B]_{FF}$ occurred 24 hours later ($t_{Qmax}=t_{[A]min}<t_{[B]min}$); **(B)** the streamflow peak occurred, but and the minimum $[A]_{FF}$ and $[B]_{FF}$ occurred 24 hours earlier ($t_{Qmax}>t_{[A]min}=t_{[B]min}$); **(C)** streamflow peak occurred, but and the minimum $[A]_{FF}$ and $[B]_{FF}$ occurred 24 hours later ($t_{Qmax}<t_{[A]min}=t_{[B]min}$); and **(D)** minimum $[A]_{FF}$ occurred, then streamflow peak occurred 24 hours later, then the minimum $[B]_{FF}$ occurred 24 hours after that ($t_{[A]min}< t_{Qmax}<t_{[B]min}$).

Appendix D: Visual Basic code for the ratio-EMMA

This appendix contains the Visual Basic code which runs the logical algorithm behind the ratio-EMMA model.

Visual basic code for Ratio-EMMA

```
Option Explicit
Public Function FastflowForwardAndReverse()

    Dim wks1 As Worksheet
    Set wks1 = ActiveWorkbook.Sheets("Ratio EMMA")
    Dim wks2 As Worksheet
    Set wks2 = ActiveWorkbook.Sheets("Intermediate calc sheet")

    Dim time As Double        'This starts off the examination of the first row of data and
then                          'continues to the second row, third row,...and so on

For time = 1 To 168           'the number rows in the original data

'measured or assumed stream and end-member values
        Dim Qs As Double              'streamflow
        Dim ConcAs As Double          'concentration of chemical A in stream
        Dim ConcBs As Double          'concentration of chemical B in stream
        Dim ConcAsf As Double         'concentration of chemical A in baseflow
        Dim ConcBsf As Double         'concentration of chemical B in baseflow

    Qs = wks1.Cells(time + 1, 2)
    ConcAs = wks1.Cells(time + 1, 3)
    ConcBs = wks1.Cells(time + 1, 4)
    ConcAsf = wks1.Cells(time + 1, 5)
    ConcBsf = wks1.Cells(time + 1, 6)

    If ConcAs = ConcAsf Then wks1.Cells(time + 1, 7) = 0
'sets ConcAff as 0 if stream concentration is equal to slowflow concentration
    If ConcAs = ConcAsf Then wks1.Cells(time + 1, 8) = 0
'sets ConcBff as 0 if stream concentration is equal to slowflow concentration
    If ConcAs = ConcAsf Then GoTo LastLine2


'loop for generating the TempQsfA and TempQsfB based on CxyF. This has to generate
all 1000 'values
```

182

```
        Dim ConcAffBffTest As Double

    For ConcAffBffTest = 1 To 1001
          Dim TempConcAff As Double
'temporary concentration of chemical A in fastflow

        TempConcAff = wks2.Cells(ConcAffBffTest + 1, 1)
          Dim TempQsfA As Double
'slowflow discharge estimate based on TempConcAff

        If (ConcAsf - TempConcAff) = 0 Then TempConcAff = TempConcAff + 0.00001
        TempQsfA = (Qs * (ConcAs - TempConcAff)) / (ConcAsf - TempConcAff)
        TempQsfA = Round(TempQsfA, 2)
        If TempQsfA < 0 Then TempQsfA = 0
      If TempQsfA > wks1.Cells(time + 1, 2) Then TempQsfA = 0
        wks2.Cells(ConcAffBffTest + 1, 2) = TempQsfA
          Dim TempConcBff As Double
'temporary concentration of chemical B in fastflow

        TempConcBff = wks2.Cells(ConcAffBffTest + 1, 3)
          Dim TempQsfB As Double
'slowflow discharge estimate based on TempConcBff

        If (ConcBsf - TempConcBff) = 0 Then TempConcBff = TempConcBff + 0.00001
        TempQsfB = (Qs * (ConcBs - TempConcBff)) / (ConcBsf - TempConcBff)
        TempQsfB = Round(TempQsfB, 2)
        If TempQsfB < 0 Then TempQsfB = 0
      If TempQsfB > wks1.Cells(time + 1, 2) Then TempQsfB = 0
        wks2.Cells(ConcAffBffTest + 1, 4) = TempQsfB

    Next ConcAffBffTest


'loop to find equal pairings of TempQsfA and TempQsfB

      Dim QsfAQsfBcheck As Double
    For QsfAQsfBcheck = 1 To 1001

        Dim QsfAHold As Double
'temporary slowflow discharge estimate based on chemical A held for comparison against
'the temporary slowflow discharge estimate based on chemical B

          Dim ConcAff As Double        'concentration of chemical A in fastflow
          Dim ConcBff As Double        'concentration of chemical B in fastflow
```

183

```vb
    QsfAHold = wks2.Cells(QsfAQsfBcheck + 1, 2)
    ConcAff = wks2.Cells(QsfAQsfBcheck + 1, 1)
    If ConcAff = 0 Then ConcAff = 0.00001

        Dim QsfBcheck As Double
'sub-loop within QsfAQsfBcheck loop to compare TempQsfB against QsfAHold

    For QsfBcheck = 1 To 1001

If QsfAHold = 0 Then GoTo LastLine1
If wks2.Cells(QsfBcheck, 4) = 0 Then GoTo LastLineA

    If wks2.Cells(QsfBcheck, 4) <= QsfAHold And wks2.Cells(QsfBcheck + 2, 4) >=
QsfAHold And QsfAHold <> 0 And wks2.Cells(QsfBcheck, 4) <> 0 And
wks2.Cells(QsfBcheck + 2, 4) <> 0 Or wks2.Cells(QsfBcheck, 4) >= QsfAHold And
wks2.Cells(QsfBcheck + 2, 4) <= QsfAHold And QsfAHold <> 0 And
wks2.Cells(QsfBcheck, 4) <> 0 And wks2.Cells(QsfBcheck + 2, 4) <> 0 Then GoTo
LastLine1

LastLineA:

    Next QsfBcheck

LastLine1:

    ConcBff = wks2.Cells(QsfBcheck + 2, 3)

    If ConcBff = 0 Then ConcBff = 0.00001

    If ConcAff / ConcBff >= wks1.Cells(time + 1, 18) And ConcAff / ConcBff <=
wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheck, 4) <= QsfAHold And
wks2.Cells(QsfBcheck + 2, 4) >= QsfAHold And QsfAHold <> 0 And
wks2.Cells(QsfBcheck, 4) <> 0 And wks2.Cells(QsfBcheck + 2, 4) <> 0 And
((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) *
ConcBff)) / (wks1.Cells(time + 1, 6) - ConcBff) < wks1.Cells(time + 1, 2) Or ConcAff /
ConcBff >= wks1.Cells(time + 1, 18) And ConcAff / ConcBff <= wks1.Cells(time + 1,
19) And wks2.Cells(QsfBcheck, 4) >= QsfAHold And wks2.Cells(QsfBcheck + 2, 4) <=
QsfAHold And QsfAHold <> 0 And wks2.Cells(QsfBcheck, 4) <> 0 And
wks2.Cells(QsfBcheck + 2, 4) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1,
4)) - (wks1.Cells(time + 1, 2) * ConcBff)) / (wks1.Cells(time + 1, 6) - ConcBff) <
wks1.Cells(time + 1, 2) Then wks1.Cells(time + 1, 7) = ConcAff

    If ConcAff / ConcBff >= wks1.Cells(time + 1, 18) And ConcAff / ConcBff <=
wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheck, 4) <= QsfAHold And
```

wks2.Cells(QsfBcheck + 2, 4) >= QsfAHold And QsfAHold <> 0 And
wks2.Cells(QsfBcheck, 4) <> 0 And wks2.Cells(QsfBcheck + 2, 4) <> 0 And
((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) *
ConcBff)) / (wks1.Cells(time + 1, 6) - ConcBff) < wks1.Cells(time + 1, 2) Or ConcAff /
ConcBff >= wks1.Cells(time + 1, 18) And ConcAff / ConcBff <= wks1.Cells(time + 1,
19) And wks2.Cells(QsfBcheck, 4) >= QsfAHold And wks2.Cells(QsfBcheck + 2, 4) <=
QsfAHold And QsfAHold <> 0 And wks2.Cells(QsfBcheck, 4) <> 0 And
wks2.Cells(QsfBcheck + 2, 4) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1,
4)) - (wks1.Cells(time + 1, 2) * ConcBff)) / (wks1.Cells(time + 1, 6) - ConcBff) <
wks1.Cells(time + 1, 2) Then wks1.Cells(time + 1, 8) = ConcBff

     If ConcAff / ConcBff >= wks1.Cells(time + 1, 18) And ConcAff / ConcBff <=
wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheck, 4) <= QsfAHold And
wks2.Cells(QsfBcheck + 2, 4) >= QsfAHold And QsfAHold <> 0 And
wks2.Cells(QsfBcheck, 4) <> 0 And wks2.Cells(QsfBcheck + 2, 4) <> 0 And
((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) *
ConcBff)) / (wks1.Cells(time + 1, 6) - ConcBff) < wks1.Cells(time + 1, 2) Or ConcAff /
ConcBff >= wks1.Cells(time + 1, 18) And ConcAff / ConcBff <= wks1.Cells(time + 1,
19) And wks2.Cells(QsfBcheck, 4) >= QsfAHold And wks2.Cells(QsfBcheck + 2, 4) <=
QsfAHold And QsfAHold <> 0 And wks2.Cells(QsfBcheck, 4) <> 0 And
wks2.Cells(QsfBcheck + 2, 4) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1,
4)) - (wks1.Cells(time + 1, 2) * ConcBff)) / (wks1.Cells(time + 1, 6) - ConcBff) <
wks1.Cells(time + 1, 2) Then GoTo Reverse

   Next QsfAQsfBcheck

'The same process is now completed with the values of TempConcAff and TempConcBff
'reversed

Reverse:

     Dim QsRev As Double
     Dim ConcAsRev As Double
     Dim ConcBsRev As Double
     Dim ConcAsfRev As Double
     Dim ConcBsfRev As Double

   QsRev = wks1.Cells(time + 1, 2)
   ConcAsRev = wks1.Cells(time + 1, 3)
   ConcBsRev = wks1.Cells(time + 1, 4)
   ConcAsfRev = wks1.Cells(time + 1, 5)
   ConcBsfRev = wks1.Cells(time + 1, 6)

185

```
If ConcAsRev = ConcAsfRev Then wks1.Cells(time + 1, 9) = 0
If ConcAsRev = ConcAsfRev Then wks1.Cells(time + 1, 10) = 0

    Dim ConcAffBffTestRev As Double
  For ConcAffBffTestRev = 1 To 1001
      Dim TempConcAffRev As Double
    TempConcAffRev = wks2.Cells(ConcAffBffTestRev + 1, 5)
      Dim TempQsfARev As Double
    If (ConcAsfRev - TempConcAffRev) = 0 Then TempConcAffRev =
TempConcAffRev + 0.00001
    TempQsfARev = (QsRev * (ConcAsRev - TempConcAffRev)) / (ConcAsfRev -
TempConcAffRev)
    TempQsfARev = Round(TempQsfARev, 2)
    If TempQsfARev < 0 Then TempQsfARev = 0
  If TempQsfARev > wks1.Cells(time + 1, 2) Then TempQsfARev = 0
    wks2.Cells(ConcAffBffTestRev + 1, 6) = TempQsfARev
      Dim TempConcBffRev As Double
    TempConcBffRev = wks2.Cells(ConcAffBffTestRev + 1, 7)
      Dim TempQsfBRev As Double
    If (ConcBsfRev - TempConcBffRev) = 0 Then TempConcBffRev =
TempConcBffRev + 0.00001
    TempQsfBRev = (QsRev * (ConcBsRev - TempConcBffRev)) / (ConcBsfRev -
TempConcBffRev)
    TempQsfBRev = Round(TempQsfBRev, 2)
    If TempQsfBRev < 0 Then TempQsfBRev = 0
  If TempQsfBRev > wks1.Cells(time + 1, 2) Then TempQsfBRev = 0
    wks2.Cells(ConcAffBffTestRev + 1, 8) = TempQsfBRev

  Next ConcAffBffTestRev

    Dim QsfAQsfBcheckRev As Double
  For QsfAQsfBcheckRev = 1 To 1001

      Dim QsfAHoldRev As Double
      Dim ConcAffRev As Double
      Dim ConcBffRev As Double
    QsfAHoldRev = wks2.Cells(QsfAQsfBcheckRev + 1, 6)
    ConcAffRev = wks2.Cells(QsfAQsfBcheckRev + 1, 5)
    If ConcAffRev = 0 Then ConcAffRev = 0.00001

      Dim QsfBcheckRev As Double
    For QsfBcheckRev = 1 To 1001
```

If QsfAHoldRev = 0 Then GoTo LastLine1Rev
If wks2.Cells(QsfBcheckRev, 8) = 0 Then GoTo LastLineARev

   If wks2.Cells(QsfBcheckRev, 8) <= QsfAHoldRev And wks2.Cells(QsfBcheckRev
+ 2, 8) >= QsfAHoldRev And QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <>
0 And wks2.Cells(QsfBcheckRev + 2, 8) <> 0 Or wks2.Cells(QsfBcheckRev, 8) >=
QsfAHoldRev And wks2.Cells(QsfBcheckRev + 2, 8) <= QsfAHoldRev And
QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0 And
wks2.Cells(QsfBcheckRev + 2, 8) <> 0 Then GoTo LastLine1Rev

LastLineARev:

   Next QsfBcheckRev

LastLine1Rev:

   ConcBffRev = wks2.Cells(QsfBcheckRev + 2, 7)

   If ConcBffRev = 0 Then ConcBffRev = 0.00001
   If ConcAffRev / ConcBffRev >= wks1.Cells(time + 1, 18) And ConcAffRev /
ConcBffRev <= wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheckRev, 8) <=
QsfAHoldRev And wks2.Cells(QsfBcheckRev + 2, 8) >= QsfAHoldRev And
QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0 And
wks2.Cells(QsfBcheckRev + 2, 8) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time
+ 1, 4)) - (wks1.Cells(time + 1, 2) * ConcBffRev)) / (wks1.Cells(time + 1, 6) -
ConcBffRev) < wks1.Cells(time + 1, 2) Or ConcAffRev / ConcBffRev >=
wks1.Cells(time + 1, 18) And ConcAffRev / ConcBffRev <= wks1.Cells(time + 1, 19)
And wks2.Cells(QsfBcheckRev, 8) >= QsfAHoldRev And wks2.Cells(QsfBcheckRev +
2, 8) <= QsfAHoldRev And QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0
And wks2.Cells(QsfBcheckRev + 2, 8) <> 0 And ((wks1.Cells(time + 1, 2) *
wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) * ConcBffRev)) / (wks1.Cells(time +
1, 6) - ConcBffRev) < wks1.Cells(time + 1, 2) Then wks1.Cells(time + 1, 9) =
Round(ConcAffRev, 2)
   If ConcAffRev / ConcBffRev >= wks1.Cells(time + 1, 18) And ConcAffRev /
ConcBffRev <= wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheckRev, 8) <=
QsfAHoldRev And wks2.Cells(QsfBcheckRev + 2, 8) >= QsfAHoldRev And
QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0 And
wks2.Cells(QsfBcheckRev + 2, 8) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time
+ 1, 4)) - (wks1.Cells(time + 1, 2) * ConcBffRev)) / (wks1.Cells(time + 1, 6) -
ConcBffRev) < wks1.Cells(time + 1, 2) Or ConcAffRev / ConcBffRev >=
wks1.Cells(time + 1, 18) And ConcAffRev / ConcBffRev <= wks1.Cells(time + 1, 19)
And wks2.Cells(QsfBcheckRev, 8) >= QsfAHoldRev And wks2.Cells(QsfBcheckRev +
2, 8) <= QsfAHoldRev And QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0
And wks2.Cells(QsfBcheckRev + 2, 8) <> 0 And ((wks1.Cells(time + 1, 2) *
187

wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) * ConcBffRev)) / (wks1.Cells(time + 1, 6) - ConcBffRev) < wks1.Cells(time + 1, 2) Then wks1.Cells(time + 1, 10) = ConcBffRev

    If ConcAffRev / ConcBffRev >= wks1.Cells(time + 1, 18) And ConcAffRev / ConcBffRev <= wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheckRev, 8) <= QsfAHoldRev And wks2.Cells(QsfBcheckRev + 2, 8) >= QsfAHoldRev And QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0 And wks2.Cells(QsfBcheckRev + 2, 8) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) * ConcBffRev)) / (wks1.Cells(time + 1, 6) - ConcBffRev) < wks1.Cells(time + 1, 2) Or ConcAffRev / ConcBffRev >= wks1.Cells(time + 1, 18) And ConcAffRev / ConcBffRev <= wks1.Cells(time + 1, 19) And wks2.Cells(QsfBcheckRev, 8) >= QsfAHoldRev And wks2.Cells(QsfBcheckRev + 2, 8) <= QsfAHoldRev And QsfAHoldRev <> 0 And wks2.Cells(QsfBcheckRev, 8) <> 0 And wks2.Cells(QsfBcheckRev + 2, 8) <> 0 And ((wks1.Cells(time + 1, 2) * wks1.Cells(time + 1, 4)) - (wks1.Cells(time + 1, 2) * ConcBffRev)) / (wks1.Cells(time + 1, 6) - ConcBffRev) < wks1.Cells(time + 1, 2) Then GoTo LastLine2

  Next QsfAQsfBcheckRev

LastLine2:

Next time
End Function

Appendix E: Supplemental information for Chapter 5: Fundamental watershed factors influencing the transport of nitrogen to streams

This supplemental information section contains additional information regarding the formation of watershed groups, a complete list of validation and calibration watersheds (Table E5.1), a complete list of variables included in the analyses of watersheds (Table E5.2), categorical variable coefficients for inclusion into multiple linear regression equations (Table E5.3), cluster analysis plots (Figure E5.1), an example regression tree (Figure E5.2) and output from random forest regression (Figure E5.3), and plots of residual from the application of the multiple linear regression equations to the various watershed groups when estimating total nitrogen loads, yields, and concentrations (Figures E5.4-E5.6).

Formation of watershed groups based on land use

Many of the land use variables were extracted by both recursive partitioning (RPART) and random forest regression (RFR) during an initial round of analyses. The land use variables were then used to separate the watersheds into groups in an effort to pull out more information from the data in the form of additional extracted (important) variables. Three separate methods were tested for separating the watersheds into groups. Cluster analysis was performed as the first method to separate the watersheds. The second method of separation involved splitting watersheds into agricultural, developed, and undeveloped based on the dominant national land cover dataset (NLCD) land use within the watershed. And a third method of separation split watersheds into agricultural, developed, and undeveloped groups based on NLCD land use, but incorporated the assumption that agricultural and developed land use more heavily influenced nitrogen in streams. The third method was used in the final research and specific details regarding that method can be found in the Methods section of the main text.

Cluster analysis was performed using the "clara" function in the "cluster" package (Maechler et al. 2015) within R. The random number generator function built into R was used in place of the random number generator included within the clara function. In this

way, the analyses could be reproduced using the set.seed(1) function in R. Three clusters were generated to ensure a large number of watersheds (>100) in each group (Figure E5.1A and B). Because of the ability RFR to determine the amount of variability in the dependent variable explained by the independent variables within each cluster, RFR was used to analyze the resulting clusters. A mean of 55% of the variability in the dependant variables (total nitrogen load, yield, and concentration) was explained when RFR was used to examine the three clusters against each of the three dependent variables. Watersheds in each of the clusters, however, were not related to each other based on a single independent variable, which prevents similar clustering among watersheds not used in this study. Because the clustering in this manner did not prove useful for application to unstudied watersheds, it was not used any further.

Instead, watersheds were grouped based on NLCD land use. Two methods were tested. For each of the methods, land use was classified as agricultural (NLCD land use codes 81 and 82), developed (NLCD land use codes 21, 22, 23, and 24), or undeveloped (NLCD land use codes 11, 12, 31, 41, 42, 43, 51, 52, 71, 72, 73, 74, 90, and 95). Separation of the watersheds into three groups (agricultural, developed, and undeveloped) was completed in two ways: solely on the majority NLCD land use type (Approach 1), and separately by the technique explained in the Methods section of the main text (Approach 2). Both methods resulted in groups of watersheds that were significantly different (rank sum test, $\alpha = 0.05$) from one another with respect to the mean total nitrogen loads, yields, and concentrations.

RFR was used to determine the degree of variability in the data that could be explained within each of the groups of watersheds with respect to total nitrogen loads, yields, and concentrations. Dependent variables applied to the watershed groups created solely based on the majority NLCD land use type explained 45% of the variability in the three dependent variables (Approach 1). Grouping the watersheds as agricultural, developed, or undeveloped (Approach 2) explained an average of 49% of the variability in the data across all load, yield, and concentration analyses. The greater percentage of variability explained when the watersheds were split up into agricultural, developed, or

undeveloped as explained by Approach 2 led to the grouping of watersheds in that manner for the rest of the analyses.

References

Brown, J.B., L.A. Sprague, and J.A. Dupree. 2011. Nutrient Sources and Transport in the Missouri River Basin, with Emphasis on the Effects of Irrigation and Reservoirs1. JAWRA Journal of the American Water Resources Association, 47(5), 1034-1060.

Hoos, A.B. and G. McMahon. 2009. Spatial analysis of instream nitrogen loads and factors controlling nitrogen delivery to streams in the southeastern United States using spatially referenced regression on watershed attributes (SPARROW) and regional classification frameworks. Hydrological Processes, 23(16), 2275-2294.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, M. Studer,P. Roudier. 2015. Cluster Analysis Extended Rousseeuw et al. R package version, 2.0-1

Moore, R.B., C.M. Johnston, R.A. Smith, and B. Milstead. 2011. Source and Delivery of Nutrients to Receiving Waters in the Northeastern and Mid-Atlantic Regions of the United States1. JAWRA Journal of the American Water Resources Association, 47(5), 965-990.

PRISM Climate Group. 2008. PRISM Climate Data, Northwest Alliance for Computational Science & Engineering, Oregon State University, http://prism.oregonstate.edu

Rebich, R.A., N.A. Houston, S.V. Mize, D.K. Pearson, P.B. Ging, and C.E. Hornig. 2011. Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico from Streams in the South-Central United States1. JAWRA Journal of the American Water Resources Association, 47(5), 1061-1086.

Robertson, D.M. and D.A. Saad. 2011. Nutrient inputs to the laurentian great lakes by source and watershed estimated using SPARROW watershed models1. JAWRA Journal of the American Water Resources Association, 47(5), 1011-1033.

Saleh, D. and J.L. Domagalski. 2012. Using SPARROW to Model Total Nitrogen Sources, and Transport in Rivers and Streams of California and Adjacent States, USA. In *AGU Fall Meeting Abstracts* (Vol. 1, p. L06).

Smith, R.A., R.B. Alexander, and G.E. Schwarz. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. *Environmental Science & Technology*, *37*(14), 3039-3047.

Wahl, K.L. and Wahl, T.L., 1995. Determining the Flow of Comal Springs at New Braunfels, Texas, Texas Water '95, American Society of Civil Engineers, August 16-17, 1995, San Antonio, Texas, pp. 77-86.

Wise, D.R., and H.M. Johnson. 2011. Surface-Water Nutrient Conditions and Sources in the United States Pacific Northwest1. JAWRA Journal of the American Water Resources Association, 47(5), 1110-1135.

Table E5.1: Calibration and validation streams within the seven major river basins, including stream gage location and total watershed area.

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
New England and Mid-Atlantic (Calibration watersheds)				
1054200	WILD RIVER AT GILEAD, ME	44.391	-70.979	181.3
1073000	OYSTER RIVER NEAR DURHAM, NH	43.131	-70.919	32.1
1073500	LAMPREY RIVER NEAR NEWMARKET, NH	43.082	-70.935	471.8
1101500	IPSWICH RIVER AT SOUTH MIDDLETON, MA	42.569	-71.027	121.0
1103500	CHARLES RIVER AT DOVER, MA	42.256	-71.260	476.8
1105000	NEPONSET RIVER AT NORWOOD, MA	42.178	-71.201	97.4
1109000	WADING RIVER NEAR NORTON, MA	41.948	-71.177	115.1
1111500	BRANCH RIVER AT FORESTDALE, RI	41.996	-71.563	241.0
1119384	WILLIMANTIC R AT MERROW, CT	41.835	-72.310	259.0
1124000	QUINEBAUG RIVER AT QUINEBAUG, CT.	42.022	-71.956	392.2
1124151	QUINEBAUG R AT WEST THOMPSON, CONN.	41.943	-71.900	433.9
1134500	MOOSE RIVER AT VICTORY, VT	44.512	-71.837	209.5
1137500	BEAVER BROOK 3M SE BETHLEHEM,NH	44.260	-71.633	238.3
1142500	AYERS BROOK AT RANDOLPH, VT	43.935	-72.658	79.9
1163200	OTTER RIVER AT OTTER RIVER, MA	42.588	-72.041	88.0
1169000	NORTH RIVER AT SHATTUCKVILLE, MA	42.638	-72.725	233.9
1169900	SOUTH RIVER NEAR CONWAY, MA	42.542	-72.694	68.1
1170100	GREEN RIVER NEAR COLRAIN, MA	42.703	-72.671	109.3
1171500	MILL RIVER AT NORTHAMPTON, MA	42.319	-72.665	137.0
1184490	BROAD BROOK AT BROAD BROOK, CT.	41.914	-72.550	40.9
1187800	NEPAUG R NR NEPAUG, CT.	41.821	-72.970	62.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
1189000	PEQUABUCK R AT FARMINGTON, CT	41.717	-72.840	128.1
1192500	HOCKANUM RIVER NEAR EAST HARTFORD, CT	41.783	-72.587	191.2
1193500	SALMON RIVER NEAR EAST HAMPTON, CT.	41.552	-72.449	263.3
1196500	QUINNIPIAC RIVER AT WALLINGFORD, CT.	41.450	-72.841	288.7
1206900	NAUGATUCK RIVER NR WATERVILLE,CT.	41.615	-73.058	261.6
1302500	GLEN COVE CREEK AT GLEN COVE NY	40.863	-73.634	30.4
1303000	MILL NECK CREEK AT MILL NECK NY	40.888	-73.564	27.4
1304000	NISSEQUOGUE RIVER NEAR SMITHTOWN NY	40.849	-73.224	72.9
1304500	PECONIC RIVER AT RIVERHEAD NY	40.914	-72.687	213.1
1305000	CARMANS RIVER AT YAPHANK NY	40.830	-72.906	193.2
1306500	CONNETQUOT RIVER NEAR NORTH GREAT RIV	40.748	-73.150	72.2
1308500	CARLLS RIVER AT BABYLON NY	40.709	-73.329	100.8
1309500	MASSAPEQUA CREEK AT MASSAPEQUA NY	40.689	-73.455	97.0
1310000	BELLMORE CREEK NEAR BELLMORE NY	40.679	-73.516	39.5
1310500	EAST MEADOW BROOK AT FREEPORT NY	40.666	-73.570	78.1
1311000	PINES BROOK AT MALVERNE NY	40.666	-73.659	26.4
1349900	BATAVIA KILL AT RED FALLS NEAR PRATTS	42.308	-74.390	140.1
1362200	ESOPUS CR AT SHANDAKEN NY	42.117	-74.388	165.4
1377000	HACKENSACK RIVER AT RIVERVALE NJ	40.999	-73.989	149.7
1380500	ROCKAWAY RIVER ABOVE RESERVOIR AT BOO	40.903	-74.410	305.4
1381000	ROCKAWAY RIVER AT PINE BROOK NJ	40.858	-74.348	317.1
1382500	PEQUANNOCK R AT MACOPIN INTAKE DAM NJ	41.018	-74.402	239.2
1391500	SADDLE RIVER AT LODI NJ	40.890	-74.080	160.1
1393450	ELIZABETH R AT URSINO LAKE AT ELIZABE	40.675	-74.222	50.1

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
1394500	RAHWAY R NR SPRINGFIELD NJ	40.687	-74.312	66.1
1395000	RAHWAY RIVER AT RAHWAY NJ	40.619	-74.283	109.7
1396500	SB RARITAN R ARCH ST AT HIGH BRIDGE N	40.664	-74.897	171.9
1396580	SPRUCE RUN AT NEWPORT NJ	40.725	-74.909	34.6
1396660	MULHOCKAWAY CREEK AT VAN SYCKEL NJ	40.648	-74.969	30.6
1397000	SOUTH BR RARITAN R AT STANTON STATION	40.572	-74.869	389.8
1398000	NESHANIC RIVER AT REAVILLE NJ	40.472	-74.828	67.5
1399500	LAMINGTON (BLACK) RIVER NEAR POTTERSV	40.728	-74.730	81.2
1400000	NB RARITAN R AT NORTH BRANCH NJ	40.600	-74.674	498.9
1401000	STONY BROOK AT PRINCETON NJ	40.333	-74.682	120.5
1407760	JUMPING BROOK NEAR NEPTUNE CITY NJ	40.204	-74.066	17.2
1408000	MANASQUAN RIVER AT SQUANKUM NJ	40.163	-74.155	113.8
1408500	TOMS RIVER NEAR TOMS RIVER NJ	39.986	-74.223	330.2
1409500	BATSTO RIVER AT BATSTO NJ	39.642	-74.650	178.3
1409810	WB WADING RIVER AT MAXWELL NJ	39.675	-74.541	219.8
1410000	OSWEGO RIVER AT HARRISVILLE NJ	39.663	-74.524	185.5
1410150	EAST BRANCH BASS RIVER NEAR NEW GRETN	39.623	-74.441	22.9
1411000	GREAT EGG HARBOR RIVER AT FOLSOM NJ	39.595	-74.851	150.8
1411500	MAURICE R NR MILLVILLE NJ	39.448	-75.073	295.3
1429500	DYBERRY CREEK AT TANNERS FALLS NEAR D	41.653	-75.282	167.7
1439500	BUSHKILL CREEK	41.088	-75.038	305.6
1440000	FLAT BROOK NEAR FLATBROOKVILLE NJ	41.107	-74.952	171.6
1443500	PAULINS KILL AT BLAIRSTOWN NJ	40.979	-74.954	330.2
1445500	PEQUEST RIVER AT PEQUEST NJ	40.831	-74.978	276.5

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
1447500	LEHIGH RIVER	41.130	-75.626	234.2
1447720	TOBYHANNA CREEK	41.085	-75.606	308.4
1451800	JORDAN CREEK NEAR SCHNECKSVILLE, PA	40.662	-75.627	149.1
1457000	DELAWARE R AT RIEGELSVILLE NJ	40.593	-75.188	402.7
1463620	ASSUNPINK CREEK NEAR CLARKSVILLE NJ	40.270	-74.672	94.5
1464000	ASSUNPINK CREEK AT TRENTON NJ	40.224	-74.749	236.7
1464500	CROSSWICKS CREEK AT EXTONVILLE NJ	40.137	-74.600	201.6
1465500	NESHAMINY CREEK NEAR LANGHORNE, PA	40.174	-74.957	541.4
1465798	POQUESSING CREEK AT STATE RD,PHILADEL	40.054	-74.984	55.0
1466500	MCDONALDS BRANCH IN BYRNE STATE FORES	39.885	-74.505	5.7
1467000	NORTH BRANCH RANCOCAS CREEK AT PEMBER	39.970	-74.684	317.6
1467081	SOUTH BRANCH PENNSAUKEN CREEK AT CHER	39.942	-75.001	23.1
1467150	COOPER R AT LAWNSIDE NJ	39.871	-75.016	45.4
1470779	TULPEHOCKEN CREEK NEAR BERNVILLE, PA	40.413	-76.172	182.0
1471000	TULPEHOCKEN CREEK	40.363	-75.968	561.1
1472157	FRENCH CRK-.25MI UPSTRM FR WILSON COR	40.151	-75.602	157.4
1473169	VALLEY CREEK AT WILSON ROAD NEAR VALL	40.081	-75.457	54.7
1476480	RIDLY CREEK AT ROUTE 291,RIDLY,PAPA	39.854	-75.348	91.5
1477800	SHELLPOT CR AT US RT 13 BRDG (GOV PRI	39.753	-75.517	23.1
1478000	CHRISTINA RIVER AT RD 346 BRIDGE	39.638	-75.681	63.1
1479000	WHITE CLAY CREEK, AT DE RT 7 BRIDGE,	39.707	-75.653	234.3
1479820	RED CLAY CREEK, ROAD 252 IN YORKLYN	39.806	-75.681	77.2
1483700	ST JONES RIVER AT DOVER, DE	39.164	-75.519	86.6
1484000	MURDERKILL RIVER, KILLENS POND AT ROA	38.981	-75.530	33.6

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
1485000	POCOMOKE RIVER NEAR WILLARDS, MD	38.389	-75.324	126.9
1485500	NASSAWANGO CREEK NEAR SNOW HILL, MD	38.229	-75.471	113.0
1487000	NANTICOKE RIVER NEAR BRIDGEVILLE, DE	38.728	-75.562	187.2
1488500	MARSHYHOPE CREEK NEAR ADAMSVILLE, DE	38.850	-75.673	124.9
1491000	CHOPTANK RIVER NEAR GREENSBORO, MD	38.997	-75.786	291.8
1493112	CHESTERVILLE BR NR CRUMPTON MD	39.257	-75.940	15.9
1493500	MORGAN CREEK NEAR KENNEDYVILLE, MD	39.280	-76.015	32.7
1495000	BIG ELK CREEK	39.730	-75.848	147.5
1523500	CANACADEA CREEK IN HORNELL (C) @ USGS	42.335	-77.683	153.5
1545600	YOUNG WOMANS CR NR RENOVO PA	41.389	-77.691	123.6
1546500	SPRING CREEK	40.890	-77.794	227.5
1550000	LYCOMING CREEK	41.418	-77.033	456.3
1552500	MUNCY CREEK NEAR SONESTOWN, PA	41.357	-76.535	62.1
1558000	LITTLE JUNIATA RIVER	40.609	-78.137	579.4
1569800	LETORT SPRING RUN AT BONNY BROOK NEAR	40.178	-77.186	59.6
1571500	YELLOW BREECHES CREEK AT NEW CUMBERLA	40.224	-76.860	565.2
1572025	SWATARA CREEK NEAR PINE GROVE, PA	40.533	-76.402	303.8
1576085	LITTLE CONESTOGA CREEK NEAR CHURCHTOW	40.145	-75.989	15.1
1576787	PEQUEA CREEK AT MARTIC FORGE, PA	39.906	-76.328	396.8
1580520	DEER CREEK BRIDGE ON STAFFORD BRIDGE RD.	39.623	-76.164	435.1
1581810	GUNPOWDER FALLS BRID.AT GUNPOWDER ROAD	39.689	-76.781	69.9
1582500	GUNPOWDER FALLS 4 END OF GLENCO RD. O	39.550	-76.636	415.1
1586000	NORTH BRANCH PATAPSCO BRI.AT.MD RT.91	39.504	-76.886	145.8
1589300	GWYNNS FALLS AT BR. ON ESSEX RD.IN VA	39.347	-76.734	88.4

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
1589440	JONES FALLS NEAR BRIDGE FALLS RD. RT.	39.392	-76.662	68.1
1589478	JANES FALLS	39.328	-76.640	144.8
1591000	PATUXENT RIVER NEAR UNITY, MD	39.238	-77.056	91.5
1592500	PAT.R.AT GAG.STN. BELOW THE ROCKY GOR	39.117	-76.875	345.1
1594000	LITTLE PATUXENT RIVER AT SAVAGE, MD	39.134	-76.816	255.3
1594526	WESTERN BRANCH AT UPPER MARLBORO, MD	38.814	-76.749	238.5
1594670	HUNTING CREEK NEAR HUNTINGTOWN, MD	38.584	-76.605	24.1
1594710	KILLPECK CREEK AT HUNTERSVILLE, MD	38.477	-76.735	12.6
1595200	STONEY RIVER BELOW MOUNT STORM DAM	39.216	-79.282	127.3
1597500	SAVAGE RIVER AT MD RT. 135	39.480	-79.069	277.6
1599000	GEORGES CREEK AT FRANK.1 M.NORTH OF W	39.494	-79.045	193.0
1601000	WILLS CREEK	39.718	-78.771	378.5
1609000	TOWNS CR. AT GAGE NEAR BR.-OLDTOWN RD	39.554	-78.554	386.7
1610155	SIDELING HILL CREEK	39.700	-78.317	269.8
1610400	WAITES RUN NEAR WARDENSVILLE, VA	39.043	-78.598	38.9
1615000	OPEQUON CREEK NEAR BERRYVILLE, VA	39.174	-78.078	148.4
1616000	APPROX 0.2 MILES ABOVE RT.7 BRIDGE	39.179	-78.086	44.8
1621050	MUDDY CREEK AT MOUNT CLINTON, VA	38.486	-78.961	43.1
1626000	ROUTE 664 BRIDGE - CITY OF WAYNESBORO	38.057	-78.908	329.0
1627500	SOUTH RIVER AT RT. 778 AT HARRISTON	38.219	-78.836	549.1
1632000	ROUTE 259 BRIDGE	38.637	-78.853	547.3
1632082	DOWNSTREAM OF RT. 257 BRIDGE	38.618	-78.799	119.2
1632900	RT. 620 BRIDGE	38.694	-78.643	249.2
1634500	ROUTE 628 BRIDGE	39.078	-78.326	264.5

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
1635500	RT. 55 BRIDGE	38.959	-78.267	227.1
1636690	PINEY RUN AT RT. # 671	39.311	-77.719	35.5
1637500	CATOCTIN CR NR BR ON MD RT 17 AT GAG	39.428	-77.556	175.2
1638480	ROUTE 663	39.255	-77.577	233.8
1639000	MONOCACY RIVER AT BRIDGEPORT, MD	39.679	-77.234	449.6
1639500	BIG PIPE BRIDGE ON BIGGS FORD RD	39.612	-77.237	267.2
1643590	LIMESTONE BRANCH AT RT. # 15 BRIDGE	39.167	-77.536	20.4
1643700	GOOSE CREEK AT RT. # 611	38.986	-77.795	318.6
1643805	NORTH FORK GOOSE CREEK AT RT. # 722 BRIDGE	39.076	-77.697	98.9
1643880	BEAVERDAM CREEK AT RT. # 734	39.037	-77.723	122.2
1644280	BROAD RUN AT RT. 7 BRIDGE	39.047	-77.433	197.1
1645000	SENECA CR. BRIDGE ON MD. ROUTE 112	39.079	-77.340	268.2
1646000	ROUTE 193 BRIDGE	38.976	-77.246	150.8
1648000	D. C. LINE, MILE 101-9.1, D. C.	38.986	-77.064	170.7
1649500	ANACOSTIA RIVER BRIDGE ON BLADENSBURG	38.939	-76.944	194.4
1651000	ON ANACOSTIA RIVER AT BLANDESBURG MD.	38.943	-76.968	135.1
1652500	GEORGE WASHIGTON PARKWAY BRIDGE	38.841	-77.048	37.1
1653000	BACKLICK RUN AT VAN DORN STREET	38.803	-77.134	87.3
1653600	PISCATAWAY CREEK	38.706	-76.966	101.9
1654000	ACCOTINK CREEK NEAR ANNANDALE VA	38.813	-77.229	60.9
1656000	RT. 806 BRIDGE	38.637	-77.626	243.0
1656100	ROUTE 646 BRIDGE	38.641	-77.512	401.7
1658000	MATTAWOMAN CREEK NEAR POMONKEY, MD	38.596	-77.056	143.8
1658500	SOUTH FORK QUANTICO CR NEAR INDEPENDE	38.587	-77.429	23.7

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
1659000	NORTH BRANCH CHOPAWAMSIC CREEK NEAR J	38.566	-77.430	16.4
1660400	RT. 641 BRIDGE	38.491	-77.434	93.2
1660500	BEAVERDAM RUN NEAR GARRISONVILLE, VA	38.507	-77.429	34.2
1661050	ST CLEMENT CREEK NEAR CLEMENTS, MD	38.333	-76.725	47.8
1665000	RT. 522 BRIDGE ABOVE CULPEPER	38.477	-77.999	46.3
1665500	RAPIDAN RIVER AT RT.29 BR. AT GAGIN STATION (	38.280	-78.341	295.3
1666500	ROUTE 614 BRIDGE	38.325	-78.096	466.2
1669000	20 M UPSTREAM FROM RT 691 BRIDGE (ESS	37.877	-76.901	75.3
1669520	RT. 17 BRIDGE	37.585	-76.604	277.7
1670180	ROUTE 651 - ORANGE COUNTY	38.155	-77.951	111.2
1671100	RT. 685 BRIDGE	37.873	-77.514	279.1
1673550	TOTOPOTOKOY CREEK	37.663	-77.258	67.9
1673800	RT. 208 BRIDGE	38.171	-77.596	201.5
1677000	WARE CREEK NEAR TOANO, VA	37.438	-76.786	18.4
2011400	ROUTE 603 AT GAGING STATION - BATH CO	38.042	-79.882	409.9
2011500	RT. 39 AT GAGING STATION	38.069	-79.897	348.1
2014000	RT. 18 BRIDGE	37.752	-79.997	397.2
2015700	RT. 614 BRIDGE	38.195	-79.571	286.3
2020500	CALFPASTURE RIVER AT DOWNSTREAM OF RT. 42 BRI	37.988	-79.494	373.0
2027000	TYE RIVER AT RT 56/158 BRIDGE	37.715	-78.982	240.4
2027500	RT. 151 BRIDGE AT GAGING STATION NELS	37.703	-79.028	127.9
2027800	RT. 657 AT GAGING STATION	37.610	-78.923	377.9
2030000	HARDWARE RIVER AT RT. 637 BRIDGE	37.813	-78.455	300.4
2031000	RT. 614 BRIDGE, W OF CHARLOTTESVILLE	38.103	-78.593	257.4

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
2032680	RT. 649 BRIDGE	38.088	-78.413	457.7
2036500	RT. 711 BRIDGE	37.598	-77.820	58.7
2038850	HOLIDAY CR NR ANDERSONVILLE VA	37.415	-78.636	23.8
2039000	BUFFALO CREEK AT BUS. RT. 460 BRIDGE-PRINCE E	37.303	-78.407	180.5
2041000	RT. 153 BRIDGE	37.284	-77.869	411.7
2042287	RT. 360 BRIDGE	37.595	-77.382	161.4
2042426	RT. 1 BRIDGE (BROOK ROAD)	37.615	-77.441	98.3
4271815	LITTLE CHAZY RIVER	44.900	-73.410	130.8
4276842	PUTNAM CREEK	43.950	-73.430	145.7
4282650	LITTLE OTTER CR AB MIDDLE BRK RD NR F	44.199	-73.212	151.9
4282795	LAPLATTE RIVER	44.370	-73.210	137.2
4296000	BLACK RIVER AT COVENTRY, VT	44.869	-72.270	301.1

New England and Mid-Atlantic (Validation watersheds)

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
1208500	NAUGATUCK R BELOW FULLING MILLS BK AT	41.502	-73.048	674.1
1208990	SAUGATUCK RIVER NEAR REDDING, CT.	41.295	-73.395	55.1
1209700	NORWALK RIVER AT WINNIPAUK,CT.	41.135	-73.426	77.3
1382500	PEQUANNOCK R AT MACOPIN INTAKE	41.017	-74.396	239.2
1396500	SB RARITAN R AT MIDDLE VALLEY NJ	40.761	-74.821	171.9
1396580	SPRUCE RUN NR GLEN GARDNER NJ	40.678	-74.918	34.6
1445500	PEQUEST RIVER (FIXED)	40.834	-75.061	276.5
1464500	CROSSWICKS C AT GROVEVILLE RD AT GROV	40.167	-74.678	201.6
1467000	N BR RANCOCAS CREEK MOUNT HOLLY	39.989	-74.785	317.6

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
1473169	STATION 437 AT VALLEY CREEK AT ROUTE	40.100	-75.463	54.7
1478000	CHRISTINA RIVER AT SMALLEYS DAM SPILL	39.655	-75.669	63.1
1483700	ST. JONES RIVER AT DELAWARE RT 10 NEA	39.125	-75.494	86.6
1576085	LITTLE CONESTOGA CREEK, SITE 3A, NEAR	40.147	-75.927	15.1
1589300	GWYNNS FALLS	39.271	-76.648	88.4
1589440	JONES FALLS	39.378	-76.644	68.1
1610155	SIDELING HILL CREEK	39.650	-78.344	269.8
1615000	OPEQUON CREEK AT RT. 655 BRIDGE (FREDERICK CO	39.148	-78.090	148.4
1615000	OPEQUON CREEK AT RT. 672 BRIDGE	39.245	-78.040	148.4
1621050	MUDDY CREEK AT RT 726 BRIDGE AT GAGING STATIO	38.487	-78.961	36.8
1634500	CEDAR CREEK AT RT. 55 BRIDGE (FREDERICK/SHENA	39.081	-78.425	266.8
1637500	CATOCTIN CR NR MOUTH AT BR ON MD RT 4	39.332	-77.579	175.2
1643700	RT. 734 BRIDGE	39.013	-77.700	317.6
1649500	HICKEY HILL, D.C.	38.909	-76.956	194.4
1653600	PISCATAWAY CREEK AT PISCATAWAY, MD	38.706	-76.966	102.3
1654000	RT. 790	38.729	-77.203	62.3
1654000	POHICK CREEK AT RT. # 1 BRIDGE	38.700	-77.210	60.9
1656000	LICKING RUN AT RT. # 616	38.619	-77.658	241.9
1664000	MARSH RUN AT RT. # 651 BRIDGE	38.475	-77.773	1605.8
2042287	APPOMATTOX RIVER AT RT.15 BRIDGE W OF FARMVIL	37.340	-78.472	161.1
2042500	CHICKAHOMINY RIVER AT RT. 625 BRIDGE	37.700	-77.514	652.7

South Atlantic-Gulf and Tennessee (Calibration watersheds)

2053200	POTECASI CRK AT NC 11 NR UNION	36.371	-77.026	583.0

202

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
2068500	DAN RIV AT NC 704 NR FRANCISCO	36.515	-80.303	334.0
2069700	GAGE NEAR NETTLE RIDGE, RT 700 BRIDGE	36.571	-80.130	220.0
2070000	NORTH MAYO AT GAGE NEAR SPENCER RT. 629 BRIDGE	36.568	-79.987	280.0
2071530	RT. 708 BRIDGE	36.779	-80.248	70.0
2081500	TAR RIV AT NC 96 NR TAR RIV	36.195	-78.583	433.0
2083800	CONETOE CRK AT SR 1409 NR BETHEL	35.774	-77.464	202.0
2085000	ENO RIVER AT HILLSBOROUGH, NC	36.071	-79.096	169.0
2085070	ENO RIV AT SR 1004 NR DURHAM	36.073	-78.863	365.0
2085500	FLAT RIVER AT BAHAMA, NC	36.183	-78.879	382.0
2090380	CONTENTNEA CREEK NEAR LUCAMA, NC	35.691	-78.110	413.0
2091000	NAHUNTA SWAMP NEAR SHINE, NC	35.489	-77.806	206.0
2092500	TRENT RIV AT SR 1129 NR TRENTON	35.064	-77.461	435.0
2093000	NEW RIV AT SR 1314 NR GUM BRANCH	34.849	-77.520	243.0
2095000	S BUFFALO CRK AT SR 2821 AT MCLEANSVILLE	36.113	-79.672	88.0
2095500	N BUFFALO CRK AT SR 2832 NR GREENSBORO	36.120	-79.708	96.0
2096846	CANE CREEK NEAR ORANGE GROVE, NC	35.987	-79.206	19.0
2097314	NEW HOPE CRK AT SR 1107 NR BLANDS	35.885	-78.966	197.0
2097464	MORGAN CREEK NEAR WHITE CROSS, NC	35.924	-79.115	21.0
2097517	MORGAN CRK AT SR 1726 NR FARRINGTON	35.861	-79.010	106.0
2099000	E FORK DEEP RIV AT SR 1541 NR HIGH POINT	36.037	-79.946	39.0
2099500	DEEP RIV AT SR 1921 NR RANDLEMAN	35.904	-79.854	324.0
2111180	ELK CRK AT NC 268 AT ELKVILLE	36.070	-81.402	124.0
2121500	ABBOTTS CRK AT SR 1243 AT LEXINGTON	35.806	-80.235	451.0
2128000	LITTLE RIV AT SR 1340 NR STAR	35.387	-79.832	275.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
2133500	DROWNING CRK AT US 1 NR HOFFMAN	35.061	-79.494	474.0
2135300	SCAPE ORE SWAMP AT S-31-108	34.108	-80.281	249.0
2137727	CATAWBA RIV AT SR 1221 NR PLEASANT GARDENS	35.686	-82.061	326.0
2138500	LINVILLE RIV AT NC 126 NR NEBO	35.795	-81.890	174.0
2143040	JACOB FORK AT SR 1924 AT RAMSEY	35.591	-81.567	67.0
2143500	INDIAN CREEK NEAR LABORATORY, NC	35.421	-81.265	177.0
2144000	LONG CRK AT SR 1456 NR BESSEMER CITY	35.305	-81.233	83.0
2146381	SUGAR CRK AT NC 51 AT PINEVILLE	35.091	-80.900	168.0
2146530	LITTLE SUGAR CRK AT NC 51 AT PINEVILLE	35.085	-80.882	127.0
2146600	MCALPINE CRK AT SR 3356 SARDIS RD NR CHARLOTTE	35.137	-80.768	104.0
2146750	MC ALPINE CREEK AT U.S. ROUTE 521 IN N.C.	35.063	-80.878	238.0
2146900	TWELVE MILE CRK AT NC 16 NR WAXHAW	34.952	-80.756	199.0
2150495	SECOND BROAD RIV AT SR 1538 NR LOGAN	35.404	-81.872	223.0
2154500	PACOLET RVR RD 978 1.5 MI SE OF FINGERVILLE	35.123	-81.990	300.0
2155500	PACOLET RVR AT S-42-55	35.110	-81.959	549.0
2164000	REEDY RVR AT S-23-30 SE GREENVILLE	34.799	-82.365	127.0
2164110	REEDY RVR ON HWY 418 AT FORK SHOALS	34.653	-82.298	269.0
2165200	S RABON CK ON DIRT RD BETWEEN SC 101 & S-30-76	34.539	-82.176	78.0
2167582	BRDG OVER BUSH RVR ON RD NO. 56	34.153	-81.615	298.0
2169570	GILLS CREEK AT COLUMBIA, SC	33.990	-80.974	153.0
2172300	MCTIER CREEK (RD 209) NEAR MONETTA, SC	33.754	-81.602	39.0
2177000	CHATTOOGA RIVER NEAR CLAYTON, GA	34.814	-83.306	531.0
2186000	TWELVE MI CK AT S-39-51 N OF NORRIS	34.780	-82.758	275.0
2186645	CONEROSS CK AT SC-59	34.636	-82.970	168.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
2187910	ROCKY RIVER AT S-04-244	34.383	-82.577	287.0
2197315	UPPER THREE RUNS CK AT ROAD A AT SRP	33.239	-81.744	526.0
2197342	FOUR MILE CK AT ROAD A-7 AT SRP	33.235	-81.698	34.0
2197400	LOWER THREE RUNS CK AT S-06-20 7.5 MI SW BRNWELL	33.176	-81.481	153.0
2198690	EBENEZER CREEK - HALF MOON LANDING	32.375	-81.222	469.0
2215100	TUCSAWHATCHEE CREEK NEAR HAWKINSVILLE, GA	32.239	-83.502	418.0
2229000	MIDDLE PRONG ST MARYS RI AT TAYLOR FL	30.436	-82.287	320.0
2246000	NORTH FORK OF BLACK CREEK AT SR 21	30.067	-81.864	458.0
2295420	PAYNE CREEK NEAR BOWLING GREEN FL	27.621	-81.826	310.0
2297100	JOSHUA CREEK AT NOCATEE FL	27.167	-81.879	338.0
2297310	HORSE CREEK NEAR ARCADIA FL	27.200	-81.988	559.0
2298608	MYAKKA RIVER AT MYAKKA CITY FL	27.344	-82.157	320.0
2299950	MANATEE RIVER NEAR MYAKKA HEAD FL	27.474	-82.211	167.0
2300500	LITTLE MANATEE RIVER NEAR WIMAUMA FL	27.671	-82.353	382.0
2300700	BULLFROG CREEK NEAR WIMAUMA FL	27.792	-82.352	75.0
2301000	NORTH PRONG ALAFIA RIVER AT KEYSVILLE FL	27.884	-82.100	346.0
2301300	SOUTH PRONG ALAFIA RIVER NEAR LITHIA FL	27.797	-82.118	274.0
2301990	HILLSBOROUGH R AB CRYSTAL SPR NEAR ZEPHYRHILLS FL	28.186	-82.184	210.0
2302500	BLACKWATER CREEK NEAR KNIGHTS FL	28.141	-82.150	282.0
2303000	HILLSBOROUGH RIVER NEAR ZEPHYRHILLS FL	28.150	-82.232	564.0
2306774	ROCKY CREEK AT ST HWY 587 AT CITRUS PARK FL	28.066	-82.566	46.0
2321000	NEW RIVER NEAR LAKE BUTLER	29.998	-82.274	495.0
2326000	ECONFINA RIVER AT US-27	30.238	-83.703	513.0
2330100	TELOGIA CREEK AT S.R. 20	30.427	-84.922	326.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
2334885	SUWANEE CREEK AT SUWANEE, GA	34.032	-84.089	120.0
2336300	PEACHTREE CREEK AT ATLANTA, GA	33.819	-84.408	222.0
2337500	SNAKE CREEK NEAR WHITESBURG, GA	33.530	-84.928	91.0
2342933	SOUTH FORK COWIKEE CREEK AT BARBOUR CO. RD. 79	32.019	-85.296	290.0
2350080	LIME CREEK NEAR COBB, GA	32.034	-83.996	157.0
2366996	ALAQUA CREEK AT NELSON ROAD	30.670	-86.187	101.0
2370000	BLACKWATER RIVER AT HWY 4 NW OF BAKER	30.834	-86.734	531.0
2399200	LITTLE RIVER AT AL. HIGHWAY 273	34.282	-85.673	515.0
2401000	BIG WILLS CREEK AT STATE RD. 227 NEAR REECE CITY	34.098	-86.038	471.0
2401390	BIG CANOE CREEK AT U.S. HIGHWAY 231	33.840	-86.263	365.0
2405500	KELLY CREEK AT U.S. HIGHWAY 231	33.447	-86.387	500.0
2423380	CAHABA RIVER NEAR MOUNTAIN BROOK AL	33.482	-86.713	359.0
2444490	BOGUE CHITTO CREEK NEAR MEMPHIS, ALABAMA	33.092	-88.299	135.0
2453000	BLACKWATER CREEK AT AL. HIGHWAY 257	33.908	-87.257	469.0
2457670	FIVEMILE CREEK AT ABANDONED BRIDGE DOWNSTREAM OF U	33.662	-86.974	238.0
2465493	ELLIOTS CREEK AT AL. HIGHWAY 69 AT MOUNDVILLE	32.997	-87.622	83.0
2467200	SUCARNOOCHEE RIVER NEAR PORTERVILLE AT HWY 45	32.699	-88.485	350.0
2471013	THREEMILE CR AT ZEIGLER BLVD AT SPRING HILL, ALA.	30.706	-88.151	27.0
2479945	BIG CREEK AT COUNTY RD 63 NEAR WILMER, AL.	30.856	-88.334	81.0
2481000	BILOXI RIVER NEAR WORTHAM	30.570	-89.136	249.0
3439000	FRENCH BROAD RIV AT NC 178 AT ROSMAN	35.142	-82.824	176.0
3456991	PIGEON RIV AT NC 215 NR CANTON	35.522	-82.848	337.0
3460000	CATALOOCHEE CRK AT SR 1395 NR CATALOOCHEE	35.667	-83.073	127.0
3463300	S TOE RIV AT SR 1168 NR CELO	35.831	-82.184	111.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
3471500	RT 660 AT RIVERSIDE CHURCH, 50 YDS DOWNST RT 645	36.761	-81.631	197.0
3478400	BELOW DAIRYMAN OFF RT. 11 ABOVE BRISTOL	36.632	-82.128	73.0
3479000	WATAUGA RIV AT SR 1121 NR SUGAR GROVE	36.239	-81.823	238.0
3488000	ON RT 633 BR, 0.5 MI OFF RT 91, 1 MILE UPSTREAM	36.908	-81.702	575.0
3539778	CLEAR CREEK AT LILLY BRIDGE NEAR LANCING, TN	36.103	-84.718	436.0
3550000	VALLEY RIV AT US 74/19/129 AT TOMOTLA	35.137	-83.980	269.0
3573182	SCARHAM CREEK NEAR MCVILLE, ALABAMA	34.298	-86.117	128.0
3598250	NORTH FORK CREEK NEAR POPLINS CROSSROADS, TN	35.584	-86.596	184.0
208524090	MOUNTAIN CREEK AT SR1617 NR BAHAMA, NC	36.150	-78.897	21.0
208524975	LITTLE R BL LITTLE R TRIB AT FAIRNTOSH, NC	36.113	-78.859	253.0
208726005	CRABTREE CRK AT SR 1649 NR RALEIGH	35.845	-78.724	197.0
208925200	BEAR CREEK AT MAYS STORE, NC	35.275	-77.794	148.0
209741955	NORTHEAST CRK AT SR 1100 NR NELSON	35.872	-78.913	54.0
242354750	CAHABA VALLEY CREEK AT CROSS CR RD AT PELHAM, AL.	33.313	-86.806	66.0

South Atlantic-Gulf and Tennessee (Validation watersheds)

2121500	ABBOTTS CRK AT NC 47 NR COTTON GROVE	35.748	-80.241	451.0
2186000	12 MI CR AT S-39-52 ABOVE CENTRAL OUTFALL	34.743	-82.802	275.0
3456991	PIGEON RIV AT SR 1642 AT CLYDE	35.535	-82.911	337.0
3478400	BEAVER CREEK	36.594	-82.187	73.0
21677037	LITTLE SALUDA RVR AT 378 E SALUDA	34.008	-81.742	233.0
21677037	LITTLE SALUDA RVR AT S-41-39 NE SALUDA	34.044	-81.698	233.0

Great Lakes, Ohio, Upper Mississippi, and Souris-Red-Rainy (Calibration watersheds)

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
3050000	TYGART VALLEY RIVER NEAR DAILEY, WV	38.809	-79.882	479.1
3075500	YOUGH. R. N OF RT 20 BR.DOWNST.FR. LT. YOUGH.	39.425	-79.421	347.1
3076600	LIT. YOUGH. R. OLD FOOT BR. 0.4M. AB. MOUTH	39.418	-79.419	126.7
3078000	CASSELMAN R. CROS. BY RIVER RD. AT USGS STA.	39.704	-79.140	161.9
3111548	STILLWATER CREEK ABOVE POOL	40.102	-81.146	253.0
3118000	L. CUYAHOGA R. AT AKRON - OTTO ST GAGE (RM 1.85)	41.094	-81.522	111.6
3118500	NIMISHILLEN CRK DST N. INDUSTRY - HOWENSTINE RD.	40.717	-81.347	453.3
3136500	KOKOSING R AT TILDEN AVE GAGE - MT. VERNON	40.406	-82.500	523.2
3157000	CLEAR CREEK NEAR ROCKBRIDGE OH	39.588	-82.578	230.5
3161000	S FORK NEW RIV AT NC 16 AND 88 NR JEFFERSON	36.395	-81.407	531.0
3165000	WILSON CREEK, RT 721 BRIDGE OFF RT 56/16, APPROX	36.600	-81.355	102.0
3175500	WOLF CREEK AT PRIVATELY OWNED LOW WATER BR OFF RT	37.256	-81.014	577.6
3187500	RIGHT FK HOLLY RIVER WV	38.636	-80.466	208.2
3191500	MUDDLETY CREEK OF GAULEY RIVER	38.327	-80.833	104.1
3198350	CLEAR FORK AT WHITESVILLE, WV	37.966	-81.524	162.7
3202750	CLEAR FORK OF GUYANDOT RIVER	37.609	-81.724	326.3
3206600	KIAH CREEK	38.061	-82.266	99.7
3208950	CRANES NEST R OF POUND RIVER VA	37.124	-82.439	172.2
3209000	POUND RIVER VA	37.237	-82.343	572.4
3213500	SLATE CREEK RT. 460 BRIDGE IN GRUNDY	37.279	-82.099	80.3
3220000	MILL CREEK OFF WALDO ROAD - MARYSVILLE (RM 16.8)	40.256	-83.345	461.0
3228805	ALUM CREEK AT COLUMBUS OH	39.945	-82.941	316.0
3230450	HELLBRANCH RUN NEAR HARRISBURG OH	39.831	-83.160	95.8
3237280	UPPER TWIN CREEK AT MCGAW OH	38.644	-83.215	31.6

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
3240000	LITTLE MIAMI RIVER NEAR OLDTOWN OH	39.748	-83.931	334.1
3241500	E FK TODD FORK AT CLARKSVILLE - SR 133/132	39.399	-83.983	163.7
3274750	BIG BLUE RIVER	39.874	-85.439	152.0
3275600	EAST FORK WHITEWATER RIVER AT ABINGTON, IND.	39.733	-84.960	518.0
3282040	HORSE LICK CREEK NEAR LAMERO	37.320	-84.139	200.2
3289300	SOUTH ELKHORN CREEK NEAR MIDWAY	38.141	-84.645	271.9
3295400	SALT RIVER NEAR GLENSBORO	38.002	-85.060	445.5
3302000	POND CREEK NEAR LOUISVILLE	38.120	-85.796	165.8
3325500	MISSISSINEWA RIVER NEAR RIDGEVILLE, IND.	40.280	-84.996	344.5
3327520	PIPE CR	40.722	-86.198	411.8
3340800	BIG RACCOON CR	39.790	-86.959	360.0
3347500	W FK WHITE RIVER	40.182	-84.969	91.9
3351500	FALL CREEK NEAR FORTVILLE, IND.	39.955	-85.868	437.7
3353500	EAGLE CR	39.778	-86.251	450.7
3353700	MILL CR	39.637	-86.641	74.6
3361650	SUGAR CREEK AT CO RD 400 S AT NEW PALESTINE, IN	39.728	-85.879	243.2
3364500	CLIFTY CREEK AT CO RD 1150 E NEAR HARTSVILLE, IN	39.292	-85.696	236.7
3369500	VERNON FK MUSCATATUCK RIV	38.976	-85.620	512.8
3382100	BANKSTON FORK NEAR DORRIS HEIGHTS, IL	37.768	-88.540	380.7
3384450	LUSK CREEK NEAR EDDYVILLE, IL	37.473	-88.548	111.1
4015330	KNIFE RIVER UPSTREAM OF US-61 AT KNIFE RIVER	46.947	-91.795	216.5
4025500	BOIS BRULE RIVER	46.679	-91.595	305.6
4063700	POPPLE RIVER NEAR FENCE, WI	45.763	-88.463	360.0
4072150	DUCK CREEK NEAR HOWARD, WI	44.534	-88.129	279.7

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
4077630	RED RIVER AT MORGAN ROAD NEAR MORGAN, WI	44.898	-88.844	295.3
4085200	KEWAUNEE RIVER NEAR KEWAUNEE, WI	44.457	-87.556	328.9
4086500	CEDAR CREEK NEAR CEDARBURG, WI	43.323	-87.979	310.8
4087240	ROOT RIVER	42.785	-87.830	492.1
4093000	CROOKED CR	41.282	-87.026	321.2
4094000	E BR LITTLE CALUMET RIVER	41.617	-87.126	171.5
4170000	HURON RIVER AT MILFORD, MI	42.579	-83.628	341.9
4170500	HURON RIVER NEAR NEW HUDSON, MI	42.513	-83.676	383.3
4175340	SALINE R AT MAPLE RD BRIDGE; SALINE TWP, SEC 13	42.130	-83.776	176.1
4175600	RIVER RAISIN NEAR MANCHESTER, MI	42.168	-84.076	341.9
4177720	FISH CR	41.559	-84.836	97.1
4197100	HONEY CREEK AT MELMORE OH	41.022	-83.110	385.9
4202000	CUYAHOGA R AT SHALERSVILLE - DIAGONAL RD-SR 303	41.238	-81.303	391.1
4214500	BUFFALO CREEK AT GARDENVILLE NY	42.855	-78.755	367.8
4229500	HONEOYE CREEK AT HONEOYE FALLS NY	42.957	-77.589	507.6
4230500	OATKA CREEK AT GARBUTT NY	43.010	-77.791	518.0
5064900	BEAVER CREEK NR FINLEY, ND	47.595	-97.709	414.4
5099400	LITTLE SOUTH PEMBINA RIVER NR WALHALLA, ND	48.865	-98.007	471.4
5287890	ELM CREEK NR CHAMPLIN, MN	45.163	-93.436	222.7
5320270	LITTLE COBB RIVER NEAR BEAUFORD, MN	43.996	-93.909	336.7
5345000	PRAIRIE CR AT 310TH ST .2 MI UPSTM OF L BYLLESBY	44.501	-92.993	334.1
5378183	JOOS VALLEY CREEK NEAR FOUNTAIN CITY, WI	44.215	-91.665	15.3
5401050	TENMILE CREEK NEAR NEKOOSA, WI	44.263	-89.811	189.8
5406500	BLACK EARTH CREEK AT BLACK EARTH, WI	43.134	-89.733	118.1

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
5427965	GARFOOT CREEK NEAR CROSS PLAINS, WI	43.110	-89.679	8.5
5430150	SUGAR RIVER AT STH 69 DWNSTM BRIDGE	42.949	-89.545	213.9
5431486	TURTLE CREEK AT CARVERS ROCK ROAD NEAR CLINTON, WI	42.598	-88.829	515.4
5451210	SOUTH FORK IOWA RIVER NE OF NEW PROVIDENCE, IA	42.315	-93.152	580.2
5455100	OLD MANS CREEK NEAR IOWA CITY, IA	41.606	-91.616	520.6
5517890	SALT CR	41.499	-87.142	78.5
5527800	DES PLAINES RIVER AT RUSSELL, IL	42.488	-87.925	318.6
5531500	SALT CREEK AT WESTERN SPRINGS, IL	41.826	-87.900	297.9
5536195	LITTLE CALUMET RIVER AT MUNSTER, IND.	41.577	-87.522	233.1
5540095	EAST BRANCH DU PAGE RIVER AT ROUTE 34 AT LISLE, IL	41.801	-88.081	234.1
5543830	FOX (IL) R. BELOW WAUKESHA	42.934	-88.293	326.3
5548280	NIPPERSINK CREEK NEAR SPRING GROVE, IL	42.444	-88.247	497.3
5551700	BLACKBERRY CREEK NEAR YORKVILLE, IL	41.671	-88.446	181.8
5568800	INDIAN CREEK NEAR WYOMING, IL	41.018	-89.836	162.4
5591550	JONATHAN CREEK NEAR SULLIVAN, IL	39.601	-88.546	89.6
5591700	WEST OKAW RIVER NEAR LOVINGTON, IL	39.731	-88.662	290.1
5592575	NORTH FORK KASKASKIA RIVER NEAR PATOKA, IL	38.774	-89.086	114.5
5592800	HURRICANE CREEK NEAR MULBERRY GROVE, IL	38.919	-89.243	393.7
5592900	CROOKED CREEK NEAR ODIN, IL	38.564	-89.050	292.7
5595730	RAYSE CREEK NEAR WALTONVILLE, IL	38.254	-89.040	227.9
5595820	BIG MUDDY RIVER NEAR MT. VERNON, IL	38.309	-88.989	199.2
407809265	TOMORROW RIVER NEAR NELSONVILLE, WI	44.524	-89.338	197.6
423205010	IRONDEQUOIT CR ABV BLOSSOM RD NR ROCHESTER NY	43.145	-77.512	367.8
423205025	IRONDEQUOIT CREEK AT EMPIRE BLVD, ROCHESTER NY	43.176	-77.527	391.1

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
Great Lakes, Ohio, Upper Mississippi, and Souris-Red-Rainy (Validation watersheds)				
3111548	STILLWATER CREEK	40.103	-81.131	253.0
3157000	HOCKING R. DST LANCASTER - 1ST U.S. RT 33 BRIDGE	39.686	-82.574	230.5
3165000	CROOKED CREEK, RT. 635 BRIDGE AT INTERSECTION WITH	36.769	-80.908	102.0
3165000	AT MT CARMEL CHURCH ON RT 650, 1.8 MI OFF RT 660	36.698	-81.058	102.0
3165000	LOW WATER BRIDGE OFF RT 711,APPROX .75 MI W. FOX	36.614	-81.306	102.0
3202750	KNOX CREEK RT. 697 AT STATE LINE	37.471	-82.062	326.3
3206600	EAST FORK TWELVEPLOTE CK	38.052	-82.308	99.7
3206600	BEECH FORK LAKE	38.299	-82.397	99.7
3207800	LEVISA FORK TEEL BR.ON RAILROAD AVE OFF RT 83, .1	37.277	-82.101	769.2
3208500	Russell Fork UPSTREAM OF LAZARUS BR & DOWNSTR RT 7	37.193	-82.288	740.7
3208950	CRANES NEST RIVER N.E. FROM ROUTE 83	37.151	-82.411	172.2
3208950	POUND RIVER VA	37.166	-82.523	172.2
3208950	DISMAL CREEK, RT. 666 BRIDGE	37.243	-82.028	172.2
3208950	MCCLURE RIVER RT. 63 BR N OF CLINCH	37.168	-82.361	172.2
3274750	BLUE R ON INTERSTATE 70 AT MILE POINT 63.76	39.854	-85.477	152.0
3333700	WILDCAT CR	40.486	-86.108	626.8
3347000	W FK WHITE RIVER	40.178	-85.342	624.2
3353500	EAGLE CR	39.735	-86.197	450.7
4072150	DUCK CREEK AT SEMINARY ROAD NEAR ONEIDA, WI	44.465	-88.219	279.7
4094000	E BR LITTLE CALUMET RIVER	41.617	-87.126	171.5
4105000	BATTLE CREEK R@ 9 MILE RD; PENNFIELD TWP SEC 21	42.364	-85.122	624.2
4177720	FISH CR	41.465	-84.814	97.1

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
4232034	IRONDEQUOIT CR AT RAILROAD MILLS, NR FISHERS NY	43.027	-77.478	101.5
5078230	LOST R AT CSAH-5 AT OKLEE	47.844	-95.859	657.9
5406500	SUGAR RIVER AT VALLEY RD.	42.973	-89.567	118.1
5408000	KICKAPOO RIVER AT ONTARIO, WI	43.715	-90.587	688.9
5551700	SOMONAUK CREEK AT SHERIDAN, IL	41.544	-88.687	181.8
5592575	HICKORY CREEK NEAR BLUFF CITY, IL	38.925	-89.039	114.5
5595730	CASEY FORK AT ROUTE 37 NEAR MT. VERNON, IL	38.269	-88.899	227.9
40263491	DRUMMOND BOG DISCHARGE 1MI NE DRUMMOND,WI	46.349	-91.258	169.4
Missouri (Calibration watersheds)				
6295113	ROSEBUD C AT RESERVATION BNDRY NR KIRBY MT	45.360	-106.990	315.2
6623800	ENCAMPMENT RIVER AB HOG PARK CR, NR ENCAMPMENT, WY	41.020	-106.820	186.3
6710605	BEAR CREEK ABOVE MORRISON	39.650	-105.180	451.1
6906150	LONG BR. @ INKWELL ST.	39.890	-92.490	58.9
6906200	E. FK. CHARITON NEAR HWY 36	39.740	-92.510	287.1
6911900	DRAGON C 2.25MI S/BURLINGAME,KAN	38.710	-95.830	292.2
6914950	BIG BULL CK AT I-35	38.750	-94.970	73.6
6921200	LINDLEY CREEK MO HWY 64 BRIDGE	37.750	-93.280	287.1
6935890	CREVE COEUR CREEK NEAR CREVE COEUR, MO	38.680	-90.480	56.4
6935980	COWMIRE CREEK AT BRIDGETON, MO	38.760	-90.430	9.6
Missouri (Validation watersheds)				
6187915	SODA BUTTE CR AT PARK BNDRY AT SILVER GATE	45.000	-110.000	80.0
6298000	TONGUE RIVER NEAR DAYTON, WY	44.840	-107.300	528.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
6332515	BEAR DEN CREEK NR MANDAREE, ND	47.780	-102.760	189.7
6409000	CASTLE CR NR HILL CITY SD	44.010	-103.830	203.0
6720990	BIG DRY CREEK 50 YARDS UPSTREAM OF USGS GUAGE STATION	40.060	-104.830	274.2
6803000	SALT CREEK AT PIONEERS BLVD AT LINCOLN, NE	40.770	-96.710	428.0
6889200	SOLDIER CR. NEAR ST. CLERE	39.370	-95.910	381.9
6889200	SOLDIER C, I RD NR DELIA, KS SC-7	39.260	-95.880	381.9
6893500	BLUE RIVER AT KANSAS CITY, MO	38.950	-94.550	481.8
6893500	BLUE RIVER STATION 40 (BR-40)	39.030	-94.520	481.8
6906300	EAST FORK LITTLE CHARITON R. NEAR HUNTSVILLE, MO	39.450	-92.560	563.9
6918740	LITTLE SAC RIVER NEAR WALNUT GROVE, MO	37.390	-93.410	607.4
6935955	FEE FEE CREEK NEAR BRIDGETON, MO	38.720	-90.440	30.0
6936475	COLDWATER CREEK NEAR BLACKJACK, MO	38.810	-90.250	103.5

Lower Mississippi, Arkansas-White-Red, and Texas-Gulf (Calibration watersheds)

7024000	BAYOU DE CHIEN NEAR CLINTON, KY	36.629	-88.964	177.9
7030392	WOLF R 72.6 @ HWY 8113 NEAR LAGRANGE QUAD 432SW	35.032	-89.248	543.9
7031692	FLETCHER CREEK AT SYCAMORE VIEW ROAD AT MEMPHIS	35.169	-89.866	79.0
7048800	RICHLAND CREEK AT GOSHEN, AR	36.104	-94.008	357.4
7052152	WILSON CREEK NEAR BROOKLINE, MO	37.147	-93.375	102.3
7052160	WILSON CREEK NEAR BATTLEFIELD, MO	37.118	-93.404	151.0
7053250	YOCUM CREEK NEAR OAK GROVE, AR	36.454	-93.357	136.8
7055646	BUFFALO RIVER AT WILDERNESS BOUNDARY	35.943	-93.406	148.7
7060710	NORTH SYLAMORE CREEK NR FIFTYSIX AR.	35.992	-92.214	150.5
7083000	HALFMOON CREEK NEAR MALTA, CO.	39.172	-106.389	61.1

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
7103780	MONUMENT C AB N.GATE BLVD AT USAF ACADEMY, CO.	39.031	-104.848	211.6
7104000	MONUMENT CREEK AT PIKEVIEW, CO.	38.918	-104.819	528.4
7145700	SLATE CREEK NEAR WELLINGTON	37.216	-97.348	398.9
7178520	DOG CREEK @ SPAVINAW FLOWLINE	36.249	-95.598	194.0
7184000	LIGHTNING CREEK NEAR OSWEGO	37.237	-95.044	510.2
7191160	SPAVINAW CREEK NEAR MAYSVILLE, AR	36.365	-94.551	228.4
7191179	SPAVINAW CREEK NEAR CHEROKEE, AR	36.342	-94.588	269.4
7191220	SPAVINAW CREEK NEAR SYCAMORE, OK	36.335	-94.641	344.5
7194800	ILLINOIS RIVER AT SAVOY, AR	36.103	-94.344	432.5
7195000	OSAGE CREEK NEAR ELM SPRINGS, AR	36.222	-94.287	336.7
7195855	FLINT CREEK 1/8 MILE BELOW FAGEN CREEK	36.216	-94.604	154.9
7196000	FLINT CREEK NEAR KANSAS, OK	36.186	-94.707	284.9
7196900	BARON FORK AT DUTCH MILLS, AR	35.880	-94.487	105.2
7197360	CANEY CREEK NEAR BARBER, OK	35.785	-94.856	232.1
7247000	POTEAY RIVER AT CAUTHRON, AR	34.919	-94.299	525.8
7247250	BLACK FORK BELOW BIG CREEK NR PAGE, OK	34.774	-94.512	192.7
7247500	FOURCHE MALINE NEAR TED OAK, OK	34.913	-95.156	316.0
7249400	JAMES FORK NEAR HACKETT, AR	35.163	-94.407	380.7
7263295	MAUMELLE RIVER AT WILLIAMS JUNCTION, AR	34.876	-92.775	119.4
7311200	BLUE BEAVER CREEK NEAR CACHE, OK	34.623	-98.564	63.7
7311630	MIDDLE WICHITA RIVER NEAR GUTHRIE, TX	33.796	-100.075	130.3
7311782	S WICHITA RV AT LOW FLOW DAM NR GUTHRIE, TX	33.622	-100.209	577.6
7335700	KIAMICHI RIVER NEAR BIG CEDAR, OK	34.638	-94.613	103.9
7340300	COSSATOT R NR VANDERVOORT AR	34.380	-94.236	232.1

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
7351500	CYPRESS BAYOU AT LINWOOD AVENUE, S. SHREVEPORT	32.344	-93.764	170.9
7352000	SALINE BAYOU AT LA. HWY 507, EAST OF BIENVILLE	32.355	-92.963	398.9
7362587	ALUM FORK SALINE RIVER NEAR REFORM, AR	34.797	-92.934	69.9
7366200	LITTLE CORNEY BAYOU NEAR LILLIE, LA	32.929	-92.633	538.7
7373000	BIG CR AT POLLOCK LA	31.536	-92.408	132.1
7375000	TCHEFUNCTE RIVER NEAR COVINGTON, LA	30.495	-90.170	266.8
7375280	TANGIPAHOA RIVER AT OSYKA AT HWY 584	31.012	-90.461	409.2
7376500	NATALBANY RIVER WEST OF PONCHATOULA, LOUISIANA	30.431	-90.547	205.9
8010000	BAYOU DES CANNES NR EUNICE, LA	30.483	-92.491	339.3
8023400	BAYOU SAN PATRICIO AT LA. HWY. 512, NEAR BENSON	31.874	-93.659	207.7
8045850	CLEAR FK TRINITY RV NR WEATHERFORD, TX	32.740	-97.652	313.4
8049700	WALNUT CREEK AT MATLOCK ROAD 2.6 MI NORTHEAST OF MANSFIELD	32.581	-97.102	162.7
8050400	ELM FORK TRINITY RIVER IMMEDIATELY DOWNSTREAM OF FM 2071 SOUTH OF GAINESVILLE	33.582	-97.128	450.7
8057200	WHITE ROCK CK AT GREENVILLE AVE, DALLAS, TX	32.889	-96.757	172.0
8058900	E FK TRINITY RV AT MCKINNEY, TX	33.244	-96.609	424.8
8067500	CEDAR BAYOU AT US 90	29.972	-94.986	168.1
8070500	CANEY CK NR SPLENDORA, TX	30.260	-95.302	271.9
8071280	LUCE BAYOU IN TRICONTINENTAL PIPELINE RIGHT-OF-WAY 1.13 KM DOWNSTREAM OF CONFLUENCE WITH CAMP BRANCH NORTHEAST OF HUFFMAN	30.109	-95.060	564.6
8074500	WHITEOAK BAYOU AT HOUSTON, TX	29.775	-95.397	223.5

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km²)
8075000	BRAYS BAYOU AT HOUSTON, TX	29.697	-95.412	245.8
8078000	CHOCOLATE BAYOU NR ALVIN, TX	29.369	-95.321	227.1
8104900	SOUTH FORK SAN GABRIEL RIVER AT US 183	30.621	-97.861	344.5
8154700	BULL CK AT LOOP 360 NR AUSTIN, TX	30.372	-97.785	57.8
8155200	BARTON CK AT SH 71 NR OAK HILL, TX	30.296	-97.926	232.3
8155240	BARTON CREEK AT LOST CREEK BLVD	30.274	-97.844	277.1
8155300	BARTON CK AT LOOK 360, AUSTIN, TX	30.245	-97.802	300.4
8155400	BARTON CK ABV BARTON SPGS AT AUSTIN, TX	30.264	-97.772	323.7
8158600	WALNUT CK AT WEBBERVILLE RD, AUSTIN, TX	30.283	-97.655	132.9
8158700	ONION CK NR DRIFTWOOK, TX	30.083	-98.008	321.2
8158810	BEAR CK BL FM 1826 NR DRIFTWOOD, TX	30.155	-97.940	31.6
8158840	SLAUGHTER CREEK AT FM 1826 SOUTH OF AUSTIN	30.209	-97.903	21.3
8162600	TRES PALACIOS RV NR MIDFIELD, TX	28.928	-96.171	375.5
8178565	SAN ANTONIO RIVER AT IH 410	29.322	-98.450	323.7
8178800	SALADO CK AT LOOP 13, SAN ANTONIO, TX	29.357	-98.413	489.5
8181480	LEON CREEK AT IH 35	29.330	-98.584	567.2
8198000	SABINAL RIVER 12.5 MILES NORTH OF SABINAL AND 2.3 MILES DOWNSTREAMM FROM THE MOUTH OF ONION CREEK 4.13 KM NW OF INTERSECTION OF RM187/FM1796	29.493	-99.497	533.5
8200000	HONDO CK NR TARPLEY, TX	29.570	-99.247	247.6
8201500	SECO CREEK AT MILLER RANCH	29.573	-99.403	116.5

Lower Mississippi, Arkansas-White-Red, and Texas-Gulf (Validation watersheds)

7191221	SPAVINAW CREEK NEAR COLCORD, OK	36.323	-94.685	422.2

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
7191222	BEATY CREEK: LOWER	36.366	-94.727	153.3
7247250	BLACK FORK AT HODGEN, OK	34.843	-94.625	192.7
8045850	CLEAR FORK TRINITY R AT FM 51	32.872	-97.745	313.4
8070500	CANEY CREEK AT FM 1485	30.149	-95.192	271.9
8155240	BARTON CK AT BARTON CK BLVD, AUSTIN, TX	30.296	-97.852	277.1
8155300	BARTON CK BL BARTON SPGS, AUSTIN, TX	30.265	-97.765	300.4
8158600	WALNUT CK SP RR BRIDGE, AUSTIN, TX	30.266	-97.657	132.9
8178565	SAN ANTONIO RIVER AT THEO AVE IN SAN ANTONIO	29.388	-98.499	323.7
8178700	SALADO CK AT LOOP 410, SAN ANTONIO, TX	29.516	-98.431	354.8

Pacific Northwest (Calibration watersheds)

470	SWAMP CREEK//USGS GAGING STATI	47.756	-122.234	59.0
484	BEAR CREEK//FIRST RAILROAD BRI	47.670	-122.110	125.0
631	ISSAQUAH CREEK//BRIDGE 99C ON	47.552	-122.048	147.0
404703	MCKAY CREEK AT KIRK ROAD (PEND	45.654	-118.823	515.0
12056500	NF SKOKOMISH R BL STAIRCASE RP	47.514	-123.330	148.0
12112600	BIG SOOS CREEK ABOVE HATCHERY	47.312	-122.165	213.0
12392155	LIGHTNING CREEK AT CLARK FORK	48.151	-116.183	305.0
13018300	CACHE CREEK NEAR JACKSON	43.452	-110.704	29.0
13150430	SILVER CREEK AT SPORTSMAN ACCE	43.322	-114.107	165.0
13238322	NORTH FORK PAYETTE RIVER BL FI	45.035	-116.059	221.0
14200400	LITTLE ABIQUA CREEK NEAR SCOTT	44.956	-122.628	25.0
14206950	FANNO CREEK AT DURHAM	45.403	-122.755	82.0
01G070	MF NOOKSACK R	48.785	-122.112	193.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
03B050	SAMISH.R.NR.BURLINGTON	48.546	-122.337	217.0
13A060	DESCHUTES.R.@.E.ST.BRIDGE	47.012	-122.902	403.0
18A050	DUNGENESS.R.NR.MOUTH	48.144	-123.128	408.0
C320	COVINGTON CREEK	47.320	-122.126	91.0
G320	LITTLE SOOS CRK	47.358	-122.124	20.0
J484	BEAR CREEK	47.718	-122.076	38.0

Pacific Northwest (Validation watersheds)

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
322	NEWAUKUM CREEK// LEFT BANK DOW	47.274	-122.056	82.0
11321	JOHNSON CREEK AT SE 17th AVE (	45.447	-122.642	139.0
412133	WILSON RIVER AT HWY 6 (RIVER M	45.473	-123.735	420.0
3805015	SCOGGINS CREEK @ OLD HWY 47	45.459	-123.154	110.0
12500450	YAKIMA RIVER ABOVE AHTANUM CRE	46.534	-120.467	352.0
13037500	SNAKE RIVER NR HEISE ID	43.613	-111.660	9.0
08C070	CEDAR R AT LOGAN ST BR AT RENT	47.486	-122.208	456.0
08C110	CEDAR RIVER NEAR LANDSBURG	47.391	-121.919	334.0
12A070	CHAMBERS CREEK NEAR STEILACOOM	47.192	-122.572	264.0
16C090	DUCKABUSH RIVER NEAR BRINNON	47.684	-123.010	179.0
24F070	NASELLE RIVER NEAR NASELLE	46.373	-123.746	143.0
27D090	EF LEWIS RIVER NR DOLLAR CORNE	45.815	-122.591	324.0
D474	NORTH CREEK	47.780	-122.187	74.0
GAR100	PAYETTE RIVER AT HARTSELL BRID	44.790	-116.145	384.0
GAR120	NF PAYETTE R AT SHEEP BRIDGE M	44.905	-116.117	384.0

Flow station ID (STAID)	Major river basin/Stream name	Latitude	Longitude	Flow station drainage area (km^2)
California (Validation watersheds)				
10256500	SNOW C NR WHITE WATER CA	33.871	-116.681	28.3
10265150	HOT C A FLUME NR MAMMOTH LAKES CA	37.644	-118.854	130.5
11060400	WARM C NR SAN BERNARDINO CA	34.078	-117.299	499.0
11073360	CHINO C A SCHAEFER AVENUE NR CHINO CA	34.004	-117.727	143.4
11073495	CUCAMONGA C NR MIRA LOMA CA	33.983	-117.599	190.3
11105510	MALIBU C A MALIBU CA	34.043	-118.684	283.7
11132500	SALSIPUEDES C NR LOMPOC CA	34.589	-120.408	122.8
11160500	SAN LORENZO R A BIG TREES CA	37.028	-122.058	292.4
11262900	MUD SLOUGH NR GUSTINE CA	37.263	-120.906	156.0
11264500	MERCED R A HAPPY ISLES BRIDGE NR YOSEMITE CA	37.732	-119.558	470.6
11274538	ORESTIMBA CR AT RIVER RD NR CROWS LANDING CA	37.414	-121.015	465.2
11292000	MF STANISLAUS R AT KENNEDY MDWS NR DARDANELLE CA	38.342	-119.823	241.3
11381500	MILL C NR LOS MOLINOS CA	40.043	-122.099	341.0
11383500	DEER C NR VINA CA	39.947	-122.053	564.0
11391100	SACRAMENTO SLOUGH NEAR KNIGHTS LANDING, CA	38.779	-121.638	257.8
11439500	SF AMERICAN R NR KYBURZ(RIVER ONLY) CA	38.764	-120.328	500.0
11447360	ARCADE C NR DEL PASO HEIGHTS CA	38.642	-121.382	98.5
11451715	BEAR C AB HOLSTEN CHIMNEY CYN NR RUMSEY CA	38.945	-122.345	266.1
11468500	NOYO R NR FORT BRAGG CA	39.428	-123.737	273.5
11475560	ELDER C NR BRANSCOMB CA	39.730	-123.644	16.9
11478500	VAN DUZEN R NR BRIDGEVILLE CA	40.481	-123.890	574.7
11495800	N FORK SPRAGUE RIVER AT POWER PLANT, NEAR BLY, OR	42.502	-120.987	174.2

Table E5.2: Complete list of all variables used by recursive partitioning (RPART) and random forest regression (RFR) analyses. Variables are long-term average values unless otherwise noted in the description.

Dependent variable name	Variable description	Data source
TNLoad	Total nitrogen load (kg N/yr)	1
TNYield	Total nitrogen yield (kg N/km^2/yr)	Calculated
TNConc	Flow-weighted mean nitrogen concentration (mg N/L)	Calculated
Independent variable name (area-normalized, if applicable)	**Variable description**	**Data source**
BGTNLoad	Total nitrogen background load (kg N/yr)	2
BGTNYield	Total nitrogen background yield (kg N/km^2/yr)	2
BGTNConc	Total nitrogen background concentration (mg N/L)	2
NMass (NRate‡)	Mean nitrogen applied to cultivated and pasture agricultural land in 1997 (kg N)	4
Canopy (CanopyPer#)	Mean canopy cover within the watershed (km2)	5
ImpBuff (ImpBuffPer+)	Mean impervious land area within 100m Buffer (km2)	5
NLCDAg (NLCDAgPer#)	Total area of 2006 national land cover dataset (NLCD) agricultural land use classes 81 and 82 (km2)	5
NLCDDev (NLCDDevPer#)	Total area of 2006 national land cover dataset (NLCD) developed land use classes 21, 22, 23, and 24 (km^2)	5
NLCDDevI (NLCDDevIPer[#])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 22, 23, and 24 (km^2)	5
NLCDAgBuff (NLCDAgBuffPer[+])	Total area of 2006 national land cover dataset (NLCD) agricultural land use classes 81 and 82 in the 100m buffer (km^2)	5

Independent variable name (area-normalized, if applicable)	Variable description	Data source
NLCDDevBuff (NLCDDevBuffPer[+])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 21, 22, 23, and 24 in the 100m buffer (km^2)	5
NLCDDevIBuff (NLCDDevIBuffPer[+])	Total area of 2006 national land cover dataset (NLCD) developed land use classes 22, 23, and 24 in the 100m buffer (km^2)	5
HLR	Predominant hydrologic landscape region (HLR) category	6
MaxEcoReg	Predominant level III nutrient ecoregion category within the watershed	5
NLCD	Predominant national land cover dataset (NLCD) land use category	5
NLCDMaj	Predominant major national land cover dataset (NLCD) land use category	5
NLCDMajBuff	Predominant major national land cover dataset (NLCD) land use category in the 100m buffer	5
BuffLU	Predominant land use in the 100m buffer (undeveloped, developed, or agricultural)	Calculated
NLCDBuff	Predominant 2006 national land cover dataset (NLCD) land use in the 100 meter buffer	5
Streamflow	Mean annual streamflow (m^3/s)	1
H2OVolume (H2OYield[#])	Annual stream volume (m^3)	Calculated
Irrigated (IrrigatedPer)	Irrigated land within the watershed (km^2)	6
Buff (BuffPer)	Total Buffer area within the watershed (km^2)	4
Dams (DamsPer)	Major Dams (count)	7
DamStor (DamStorPer)	Total Dam Storage (acrefeet)	7
Strahler	Strahler stream order	2
Area	Total watershed area (km^2)	4
MRB	SPARROW major river basin identifier	1
Lat	Stream gage latitude	4
Lon	Stream gage longitude	4

Independent variable name (area-normalized, if applicable)	Variable description	Data source
Pop2000 (PopPer)	2000 Population density (people/km^2)	8
RoadX (RoadXPer)	Road crossings within the watershed (count)	4
HSGA	Mean percent of soil within the watershed as hydrologic soil group HGA	6
HSGAC	Mean percent of soil within the watershed as hydrologic soil group HGAC	6
HSGAD	Mean percent of soil within the watershed as hydrologic soil group HGAD	6
HSGB	Mean percent of soil within the watershed as hydrologic soil group HGB	6
HSGBC	Mean percent of soil within the watershed as hydrologic soil group HGBC	6
HSGBD	Mean percent of soil within the watershed as hydrologic soil group HGBD	6
HSGC	Mean percent of soil within the watershed as hydrologic soil group HGC	6
HSGCD	Mean percent of soil within the watershed as hydrologic soil group HGCD	6
HSGD	Mean percent of soil within the watershed as hydrologic soil group HGD	6
HSG	Predominant Hydrologic group category	Calculated
PercSand	Mean percent sand in soil within the watershed	6
PercSilt	Mean percent silt in soil within the watershed	6
PercClay	Mean percent clay in soil within the watershed	6
PercOM	Mean percent organic matter in the soil within the watershed	6
AWC	Mean soil available water capacity (cm/cm) within the watershed	6
Kfactor	Mean soil K-factor within the watershed	6
Rfactor	Mean soil R-factor within the watershed	9
Perm	Soil permeability (cm/hr)	6
SoilThick	Soil thickness (m)	6

Independent variable name (area-normalized, if applicable)	Variable description	Data source
SoilEC	Soil electrical conductivity (mS/cm)	4
SRL25	Percent of soil restrictive layer above 25cm	4
SRL35	Percent of soil restrictive layer above 35cm	4
SRL45	Percent of soil restrictive layer above 45cm	4
SRL55	Percent of soil restrictive layer above 55cm within the watershed	4
WTDepth	Depth to water Table (m)	6
ArtDrain (ArtDPer)	Artificially drained land within the watershed in 1992 (km^2)	4
Ditch (DitchPer)	Area of ditched land within the watershed (km^2)	4
SubDrain (TileDPer)	Area of land subject to subsurface drainage within the watershed (km^2)	4
BFI-WAHL	Baseflow index: from Wahl and Wahl (1995) BFI program	10
BFI-NHD	Baseflow index: from national hydrology dataset	6
SubH20Contact	Subsurface water contact time (days)	6
GWRech	Groundwater recharge (mm/yr)	6
Runoff	Mean estimated runoff (mm/yr)	11
InfilOF	Mean annual infiltration excess overland flow (% of total streamflow)	6
SatOF	Mean estimated saturation overland flow (% of total streamflow)	6
FstFrz	Average first freeze (day of year)	9
LstFrz	Average last freeze (day of year)	9
MaxAP	Maximum annual precipitation (mm)	9
MaxPdays	Maximum number of days with precipitation (days)	9
MinAP	Minimum annual precipitation (mm)	9

Independent variable name (area-normalized, if applicable)	Variable description	Data source
MinPdays	Minimum number of days with precipitation (days)	9
AnnualPdays	Mean number of days of days with precipitation per year (days)	12
JanPdays	Number of days of precipitation in January (days)	12
FebPdays	Number of days of precipitation in February (days)	12
MarPdays	Number of days of precipitation in March (days)	12
MayPdays	Mean number of days of precipitation in May (days)	12
JunPdays	Number of days of precipitation in June (days)	12
JulPdays	Mean number of days of precipitation in July (days)	12
AugPdays	Number of days of precipitation in August (days)	12
SepPdays	Number of days of precipitation in September (days)	12
OctPdays	Number of days of precipitation in October (days)	12
NovPdays	Mean number of days of precipitation in November (days)	12
DecPdays	Number of days of precipitation in December (days)	12

‡ Application rate (kg/km2)

Percent of watershed area

+ Percent of 100m buffer area

1 Moore et al., 2011, Hoos and McMahon, 2009, Robertson and Saad, 2011, Brown et al., 2011, Rebich et al., 2011, Wise and Johnson, 2011, Saleh and Domagalski, 2012

2 Smith et al. 2003

3 Personal communications with Mike Wieczorek and Naomi Nakagaki, January 15, 2015

4 Personal communications with Mike Wieczorek, January 15, 2015

5 Personal communications with Mike Wieczorek and Andrew LaMotte, January 15, 2015

6 Personal communications with Mike Wieczorek and Dave Wolock, January 15, 2015

Independent variable name (area-normalized, if applicable)	Variable description	Data source

7 Personal communications with Mike Wieczorek and James Falcone, January 15, 2015

8 Personal communications with Mike Wieczorek and Curtis Price, January 15, 2015

9 PRISM Group, 2008

10 Wahl and Wahl, 1995

11 Personal communications with Mike Wieczorek, Dave Wolock, and Gregory McCabe, January 15, 2015

12 Personal communications with Mike Wieczorek, James Falcone, and Ryan Hill, January 15, 2015

Table E5.3: Coefficient values of categorical variables for insertion into recursive partitioning (RPART) based and random forest regression (RFR) based multiple linear regression models. For example, when using variables extracted by RPART for estimating total nitrogen yield in a developed watershed within hydrologic landscape region (HLR) "2", + 0.0103 is added to the equation. NA represents categories not present in the watershed group.

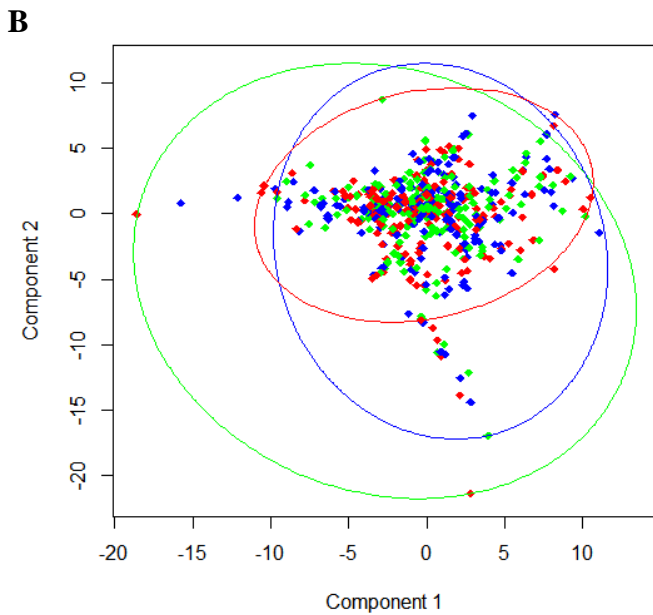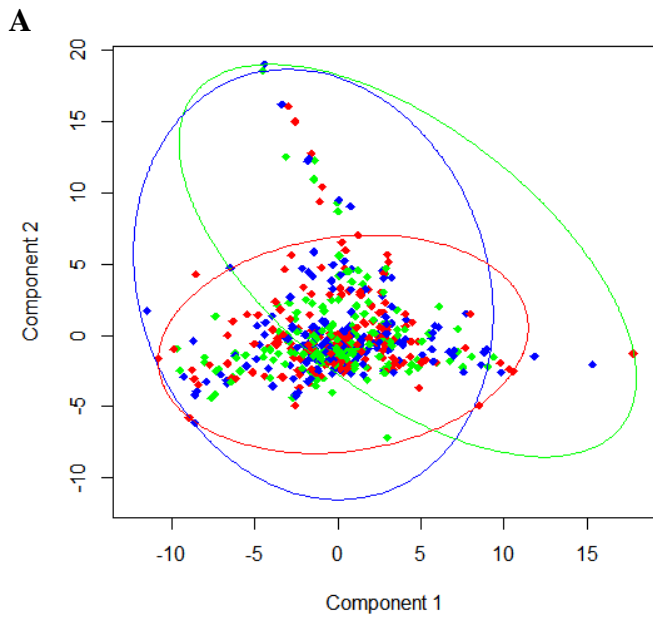| Categorical variable name	Variable category	Total nitrogen yield[#]		Total nitrogen concentration[#]					
		Developed watersheds		Agricultural watersheds		Developed watersheds		Undeveloped watersheds	
		RPART	RFR	RPART	RFR	RPART	RFR	RPART	RFR
HLR	2	0.0103	0.0483					-0.244	
	3	0.175	0.380					NA	
	4	0.0266	0.316					-0.0346	
	5	-0.418	-0.585					NA	
	6	0.361	0.154					NA	
	7	0.00526	0.285					-0.281	
	8	-0.208	0.865					-0.660	
	9	0.113	0.185					-0.341	
	10	-0.398	0.332					-0.736	
	11	0.185	0.487					-0.391	
	12	-0.239	0.0195					-0.320	
	13	-0.190	0.849					-1.04	
	15	-0.747	0.000					-0.356	
	16	0.310	0.482					-0.357	
	17	NA	NA					-1.398	
	18	NA	NA					-0.398	
	19	NA	NA					-0.275	
	20	NA	NA					-0.327	
MaxEcoReg	2		0.294			0.171			NA
	3		NA			NA			-0.430
	4		-0.147			0.425			-0.0754
	5		-0.426	0.448	0.179	0.378			0.00
	6		0.481	0.575	0.189	0.542			NA
	7		0.052	0.358	-0.0119	0.226			0.0696
	8		0.364	NA	NA	0.373			-0.0188
	9		0.258	0.303	-0.054	0.302			-0.0908
	10		0.869	0.309	-0.167	0.706			NA
	11		0.604	0.272	0.00966	0.603			0.00559
	12		0.486	0.214	-0.402	0.267			-0.212
	14		0.506	0.275	-0.113	0.435			-0.211
NLCDBuff	22					0.289			
	23					-0.0767			
	24					0.756			
	41					0.0272			
	42					-0.205			
	43					-0.0130			
	81					0.188			
	82					0.470			
	90					-0.142			
MRB	2							-0.0424	
	3							0.0565	
	4							0.132	
	5							-0.0547	
	7							0.0466	

Figure E5.1: Cluster plots for the set of explanatory variables used for **(A)** load estimation and **(B)** yield and concentration estimation. Watersheds are represented by the colored points and are associated with one of three clusters. Colored ellipses contain a majority of the correspondingly colored watersheds within the specific cluster.
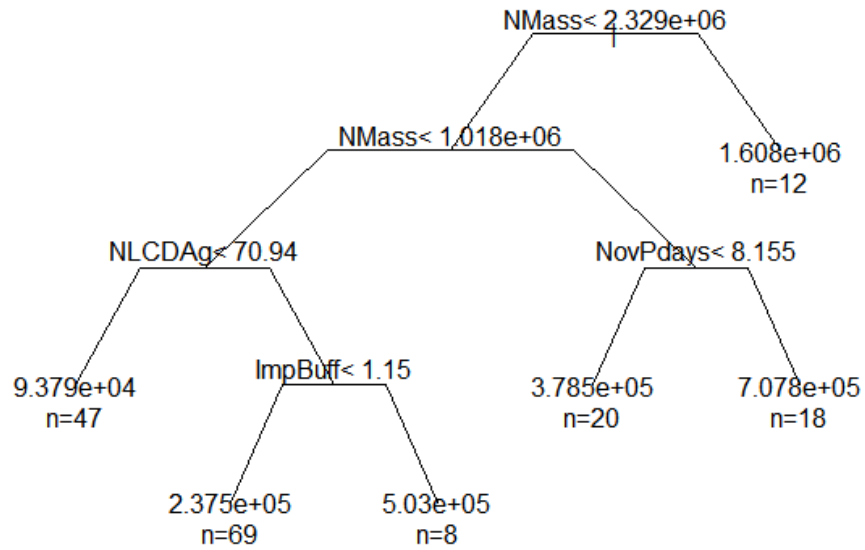
Figure E5.2: A regression tree for total nitrogen loads in agricultural watersheds as produced using recursive partitioning (RPART). The variables that were extracted at each node were considered the important variables for determining total nitrogen load in agricultural streams. Variable names are explained in Table E5.2.
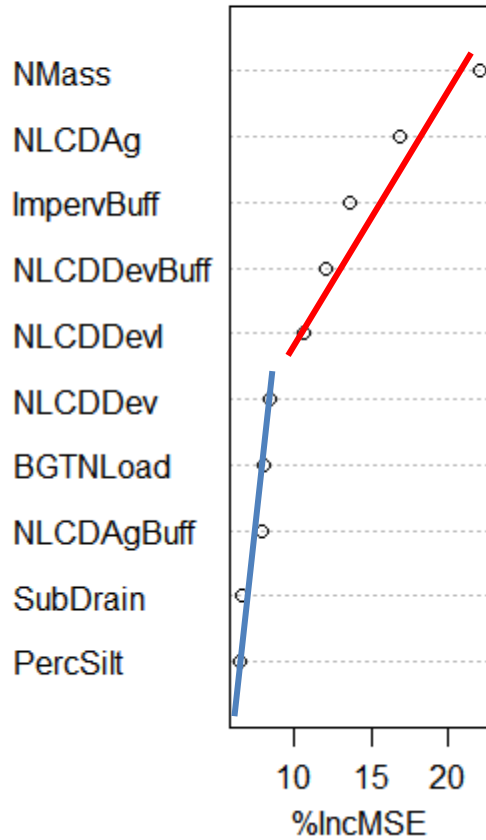
Figure E5.3: A plot of variable importance for total nitrogen loads in agricultural watersheds, as produced using random forest regression (RFR). The X-axis represents the percent increase in mean squared error (%IncMSE) of dependant variable predictions after randomly permuting all of the values within each independent variable. The visual break in slope was used to define which variables are important and which are not. Variables with values that fall along the top line were considered important for estimating loads in agricultural streams, whereas variables with values that fall along the bottom line were considered not important. Lines were drawn manually. Variable names are explained in Table E5.2.
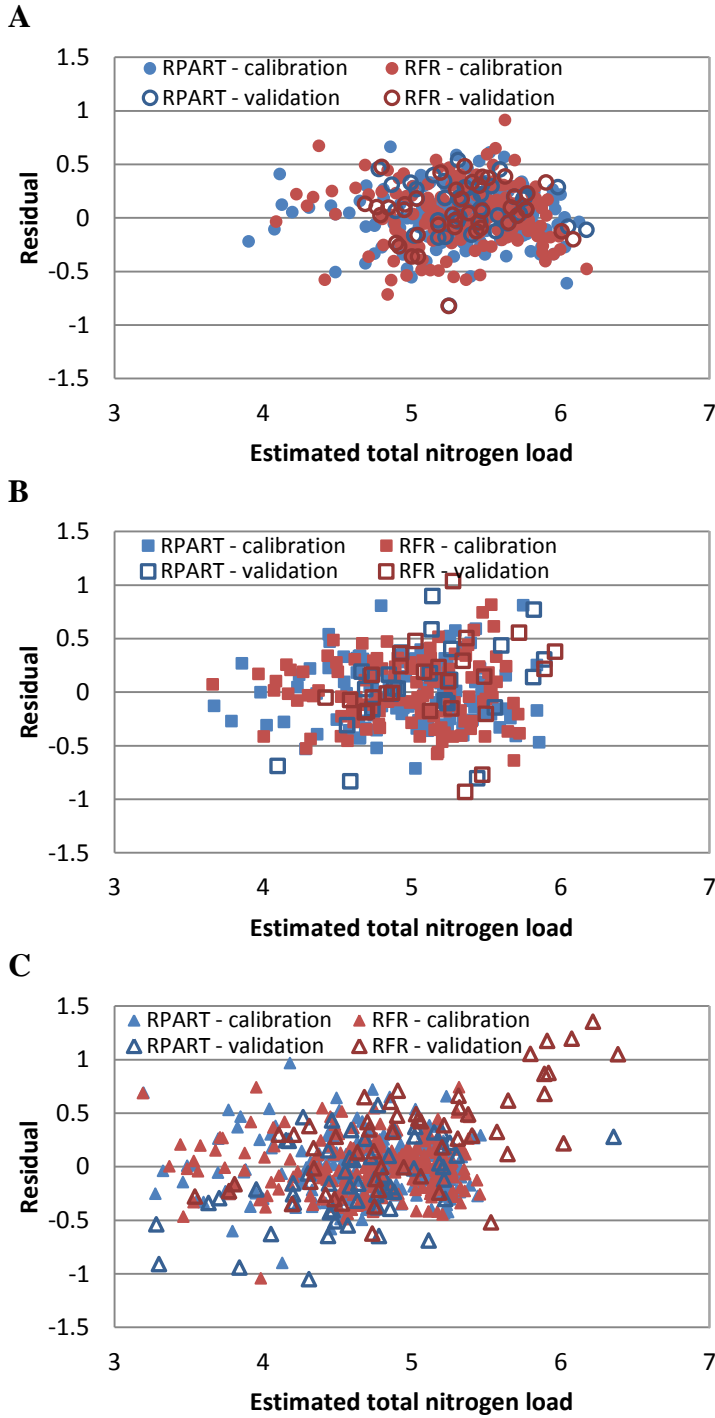
Figure E5.4: Residuals (log10-transformed) from total nitrogen load (log10-transformed) estimates from multiple linear regression equations for calibration and validation data sets compared to estimated load values for **(A)** agricultural, **(B)** developed, and **(C)** undeveloped streams using important variables from recursive partitioning (RPART) and random forest regression (RFR). Note the differing values on the X-axes.
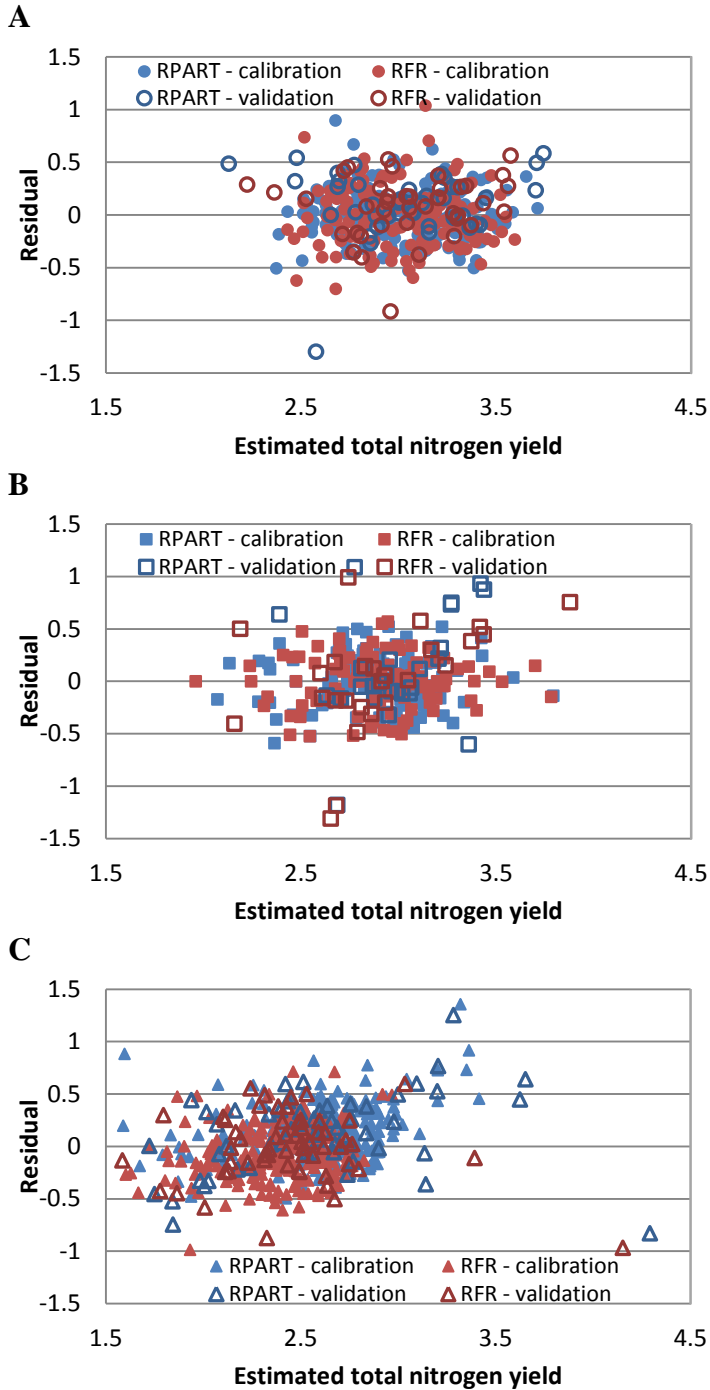
Figure E5.5: Residuals (log10-transformed) from total nitrogen yield (log10-transformed) estimates from multiple linear regression equations for calibration and validation data sets compared to estimated yield values for **(A)** agricultural, **(B)** developed, and **(C)** undeveloped watersheds using important variables from recursive partitioning (RPART) and random forest regression (RFR). Note the differing values on the X-axes.
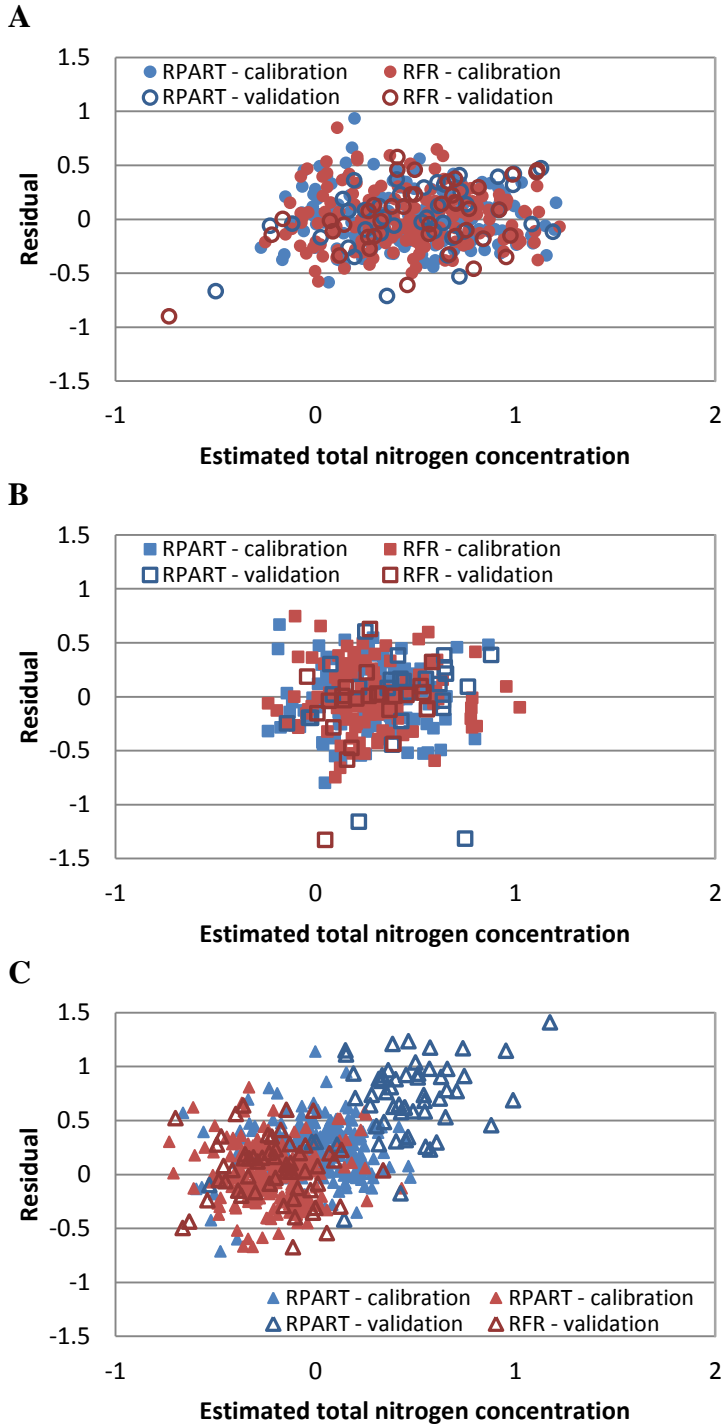
Figure E5.6: Residuals (log10-transformed) from total nitrogen concentration (log10-transformed) estimates from multiple linear regression equations for calibration and validation data sets compared to estimated concentration values for **(A)** agricultural, **(B)** developed, and **(C)** undeveloped streams using important variables from recursive partitioning (RPART) and random forest regression (RFR). Note the differing values on the X-axes.