12-2012

Estimation of Solute Fluxes from Ungaged Headwater Catchments in the Catskill Park of New York State

Chris Gianfagna
Syracuse University

Abstract

Predictions of flow and subsequent solute fluxes from ungaged basins have important implications both for water resources management and ecosystem monitoring studies. The Catskill region of New York State is one such place that requires both water resources management and ecosystem monitoring due to its strategic location as the main water-supplying region for New York City. This study examines the differences in chemical mass flux estimates made in ungaged basins using three different chemistry aggregation methods for solute concentrations determined from monthly grab samples. The efficacy of area ratios for predicting flow at the upstream location of a nested pair of stream gages based on flow at the downstream reference gage is also explored. The benefit of data set partitioning and development of separate prediction models for different flow regimes of the reference gage is analyzed, and a threshold of area ratio for use of such a method is established, with implications for use in ungaged basins. This work is focused on the Catskill region, but is likely to be applicable to other temperate, montane systems.

Significant relationships were observed between upstream and downstream flow in all test watersheds. Furthermore, watershed area ratio was the most important basin parameter for estimating flow at the upstream location of a nested pair of stream gages. The area ratio alone explained 93% of the variance in the functional relation slopes that best fit the flow regressions. Data set partitioning was found to be beneficial only for nested pairs with area ratios greater than 0.1, and was determined by analysis of the root mean square error of the different flow prediction models. Five of the fifteen test watershed pairs had a lower root mean square error using the partitioned relationships and these pairs all had area ratios greater than 0.1.

The relative difference between the three different chemistry aggregation methods was found to be relatively small on an annual basis (average difference of 7%) and increase with shorter time steps up to daily flux estimates (average difference of 26%). This finding indicates that simple flow estimation methods based on area ratios are justifiable, and perhaps preferred, for estimation of annual chemical mass fluxes, and that for such estimates of flux, the exact solute chemistry aggregation method matters little on an annual basis.

Estimation of Solute Fluxes from Ungaged Headwater Catchments in the Catskill Park of New York State

By

Chris Gianfagna

B.S. SUNY Plattsburgh 2010

Thesis

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Environmental Engineering

Syracuse University
December 2012

Table of Contents

List of Tables

List of Figures

Figures in Appendix B

Chapter 1: Introduction

Headwater catchments represent the most basic unit of watershed hydrology. Headwater catchments are typically described as small upland watersheds containing a perennial stream channel. A watershed, or catchment, is the areal extent of land which drains to a common point on a stream or river. Headwater catchments provide a variety of watershed functions including regulation of the discharge regime and establishment of the chemical characteristics of downstream reaches (Lowe and Likens, 2005). Many of these catchments however, tend to be ungaged and poorly understood due to their small size (Gomi et al., 2002; Lowe and Likens, 2005). The ability to estimate flow in ungaged catchments and, subsequently, solute fluxes from these watersheds is important for water resources planning and management. In particular, the estimation and modeling of water availability and quality for water supply and ecological assessment requires reliable estimation of flow (Lowe and Likens, 2005). In this particular study, the catchments of interest are located in the Catskill Mountain region of New York State and discharge to water supply reservoirs for New York City. The goal of this study is to enhance the understanding of chemical mass fluxes in headwater streams and to compare methods used to estimate discharge and chemical fluxes in ungaged catchments.

Chemical mass flux is a measurement of the mass flow rate (mass per unit time) of a substance per unit area. Mass fluxes are of interest in both long-term modeling and monitoring studies and can be conveniently thought of as an integrated measurement of all processes within the watershed that act to influence water quality (Semkin et al., 1994). Mass fluxes of nitrogen, sulfur, calcium and other solutes are important water quality parameters, with direct application in establishing and evaluating water quality remediation and pollution abatement programs, such as the establishment and observation of total maximum daily loads (TMDLs).

For chemicals dissolved in stream water (solutes), mass flux is calculated as the product of the volumetric flow rate of the stream and the chemical concentration of the solute, normalized by the watershed area above the point of interest (USGS, 2007). This is illustrated by Equation 1, below:

$$\text{Mass Flux} = \frac{Q \times C}{A} \qquad [1]$$

Where: Q = Volumetric flow rate of the stream
C = Solute concentration in the stream
A = Watershed area above the point of interest on the stream

Any consistent set of units may be used in Equation 1, although convention is to display chemical mass fluxes in SI units (e.g. kg ha^{-1} yr^{-1}).

In gaged catchments flow rate is usually measured at a sufficiently short time step by stream gaging stations to be useful in calculating chemical fluxes at a variety of time scales. Chemical concentrations, however, are generally obtained via grab sampling and laboratory analysis, and measured values are therefore obtained much less frequently due to cost and difficulty of collection and analysis (USGS, 2007). Grab samples are taken as a single withdrawal of water from a stream and provide an instantaneous glimpse at the solute concentrations in the stream at the time of sampling. The use of grab samples leads to the need to aggregate and interpolate concentration values based on sparse measurements.

In ungaged catchments the goal of estimating solute fluxes is even more difficult, as it involves estimation of both the volumetric flow rate and interpolation of chemical concentration values. Historically, volumetric flow rate in ungaged catchments has been estimated using a variety of techniques. Perhaps the earliest and most common technique for estimating daily flow in an ungaged catchment is the area ratio method. Other techniques include empirical regional regression models (Riggs, 1990), use of flow duration curves (Castellarin et al., 2004), and models developed from rainfall-runoff relationships (Post and Jakeman, 1999).

The area ratio method is used to estimate flow in an ungaged catchment when a nearby gaged watershed is present for use as a reference. The method estimates flow at an ungaged location by multiplying the measured flow at the nearby reference gage by the ratio of the ungaged watershed area to the gaged watershed area (Archfield and Vogel, 2010). A major assumption of the area ratio method is that flow scales directly with watershed area. That is, that as watershed area increases, flow rate increases at some fixed rate per unit area. This means that the flow per unit area is the same at both the ungaged location and gaged reference location.

The choice of reference gage in the area ratio method has generally been determined by geographic proximity to the ungaged watershed of interest, or by locating a watershed that should share a similar hydrologic response as the ungaged watershed of interest (Archfield and Vogel, 2010). Mohamoud (2008) suggests choosing the closest stream gage, while Smakhtin (1999) suggests that several reference stream gages should be used in order to smooth out any timing-related issues between the ungaged and reference locations. Recently, Archfield and Vogel (2010) suggested a “Map Correlation Method”, a new technique for identifying the most correlated stream gage based on watershed characteristics and hydrologic response.

Regional regression analysis is an approach by which stream flow in gaged catchments is related to the physical and climatological characteristics of their basins (Riggs, 1990). For an ungaged watershed, these basin characteristics can be determined from maps, field measurements, and climate records. Flow statistics in the ungaged catchments can then be estimated using the relationships developed from the gaged catchments (Riggs, 1990). Riggs (1990) suggested that in order for regional analysis to be successful the study area must meet certain criteria. This includes that the study area contain an adequate number of evenly distributed stream gages that have sufficient periods of record to cover a wide range of flows. Flow in the gaged catchments must also be explained by relatively few basin characteristics and climate records must be adequate to determine precipitation statistics over the study area. Riggs also suggested that the area be adequately mapped in order to estimate pertinent basin characteristics and that the headwaters of both the gaged and ungaged catchments be within the study area. Riggs' inclusion of the caveat that the headwaters be included in the study area reinforces the importance of headwater catchments in establishing and regulating the flow regime of downstream reaches.

Regional regression analysis techniques typically employ multiple regression analysis with several basin and climate characteristics in an attempt to regionalize the hydrologic response characteristics of the area of interest. Regionalization is an attempt to group watersheds together and generalize their physical or chemical characteristics. Common basin and climate characteristics used for estimation of mean annual flow include drainage area, stream channel length, and annual precipitation (Riggs, 1990). Flood discharge characteristics at gaged sites have been examined by Lumia (1991) in New York State using drainage area, main channel slope, percent basin storage, mean annual precipitation, percent forested area, average main channel elevation, and a basin shape index as explanatory variables. Although much work has been done using regional regression models to estimate annual flow and flood statistics (Riggs, 1990; Lumia 1991; Mulvihill et al. 2009; Vogel et al. 1999), relatively little work has been done using regional regression to estimate flows on a daily basis. Regional regression has, however, been used to fill in data gaps in real time or daily flow data sets in which equipment failure or extreme flow events have left the data set incomplete (Dai et al., 2008).

Flow duration curves (FDCs), on the other hand, have been used recently to estimate daily flow in ungaged catchments. Flow duration curves are cumulative frequency curves that show the percent of time a specified flow has been equaled or exceeded in a given period of time (Seary, 1959). They are constructed by ranking the observed flow data and plotting them against their respective durations. Generally, the Weibull plotting formula is used to generate the curves (Castellarin et al. 2004). Both regional and empirical FDCs can be constructed for estimation of flow in ungaged basins. Although the development of regional and empirical FDCs seems similar to flow estimation based on regional regression analysis, the actual estimation of flow is based on the constructed FDC, which is developed from regionalization of established FDCs from nearby gaged catchments, whereas in regional regression estimation of flow is done directly from relationships developed from basin or climate data for a given region.

Generation of regional FDCs requires identification of basin or climate characteristics that explain the FDCs of gaged catchments in a region. These variables are typically identified using multiple regression analysis and a set of gaged stream flow data. Once the variables are identified a daily FDC for the ungaged watershed can be constructed and used to estimate flow. Estimation of a daily flow time series can be done by sequencing the FDC for the ungaged catchment with the time series from a nearby gaged catchment (Mohamoud, 2008).

Rainfall-runoff models have been used extensively to estimate stream flow following rain events and have evolved from the earliest attempts of estimating peak discharge using the

Rational Method to estimation of daily stream flow from more sophisticated rainfall-runoff models (Beven, 2012; Post and Jakeman, 1999). Although varying greatly in complexity, these models all share a common goal that aims to predict runoff (stream flow) from observed rainfall data and basin characteristics (Tan et al., 2005). Rainfall-runoff models may be categorized as empirical, process-based, or conceptual depending on the mechanisms they use to convert the rainfall input data to stream flow output data. Empirical models, also known as "black box" models, use empirical relationships to relate stream flow to rainfall, which are the only two parameters that have physical meaning in the model (Tan et al., 2005). Process-based models attempt to simulate watershed processes by using partial differential equations meant to represent the physical processes taking place within the catchment that control the conversion of rainfall to stream flow (Beven, 2012). Process-based models have the advantage of being based on actual physical processes, but suffer from the fact that they usually contain many parameters which must be measured, assumed, or calibrated (Beven, 2012). Conceptual rainfall-runoff models strive to portray the catchment as a series of connected reservoirs that represent different storage and flow-through processes in the catchment with mathematical equations dictating inflow and outflow of the various reservoirs (Tan et al., 2005).

Recent work in the area of rainfall-runoff modeling includes studies by Post and Jakeman (1999) and Tan et al. (2005), who have attempted to use rainfall-runoff models to predict daily stream flow in Australia. Post and Jakeman (1999) used a lumped parameter, conceptual rainfall-runoff model that utilizes only six parameters, while Tan et al. (2005) used a seven parameter conceptual model called SIMHYD that predicts daily stream flow from daily precipitation and potential evapotranspiration over the study area. Lumped parameter models seek to simplify physical processes by approximating their behavior with reasonable assumptions and mathematics.

The use of rainfall-runoff models for estimating daily discharge requires adequate climate data, usually in the form of daily rainfall data (Post and Jakeman, 1999). In locations where daily precipitation data are unavailable, rainfall-runoff models may not be appropriate for estimating daily flows.

Solute concentrations in streams can easily be measured via grab samples and laboratory analysis, but can rarely be measured with the frequency of discharge measurements due to cost and collection constraints. This leads to the issue of aggregating stream chemistry data for use with daily discharge data. The goal of aggregating the chemistry data is to produce data that are representative of the period over which they are to be applied. Problems can arise when the frequency of stream sampling is inappropriate for the time scale at which flux estimates are to be made. Various methods have been developed to deal with this issue. They include averaging methods, regression-model methods, and the Composite Method (Semkin et al., 1994; Aulenbach and Hooper, 2006). Averaging methods range from the simple application of the annual average concentration to the daily flow value for every day of the study period to period-weighted averaging techniques that apply a single concentration value to each specified period of days within the total study period. Averaging techniques may also include seasonal averaging approaches that apply a single concentration to each specified season for the period of days designated in that season (Semkin et al., 1994). Seasonal averaging techniques may be appropriate for seasonally or biologically mediated solutes, whose concentrations may fluctuate considerably on a seasonal basis.

Regression-model methods relate concentration to daily stream flow via concentration-discharge rating curves or other means. Concentration-discharge curves are plots that compare solute concentrations to the magnitude of stream flow on the date on which they were collected. Ideally, they contain an adequate number of data points to cover the expected range of flows and solute concentration values for the study area and period. Concentration-discharge curves can be used to estimate concentrations on a daily basis based on average daily flow (USGS, 2007), but are only appropriate for solutes that are well correlated with flow (Semkin et al., 1994). Regression methods also assume that the flow and concentration data are independent and have a normal distribution, an assumption which has been shown to be inaccurate in some cases (Dann et al., 1986).

The Composite Method was recently suggested by Aulenbach and Hooper (2006) and combines the regression-model method with a period-weighted approach. In this method, regression model estimates of concentration are adjusted on sample days with known concentrations and the model residuals are applied to the regression model estimates during the period between samples in an attempt to correct errors in the regression model estimates (Aulenbach and Hooper, 2006). Model residuals are the difference in actual and model-predicted values. The Composite Method is a hybrid method that attempts to correct errors made by regression-model methods using data from measured grab samples. The Composite Method is only appropriate for solutes that are correlated to flow rate.

This particular study was conducted in concert with and utilizes data collected and generated for a broader study looking at the response of acidified soils and surface waters to decreases in atmospheric acid inputs (acid rain) in the Northeast United States. The broader study makes use of stream chemistry data from 26 watersheds in the Catskills, 4 of which represent the study sites for this project. The broader study also includes study sites in the Adirondack Mountains of New York and the Hubbard Brook Experimental Forest in New Hampshire, where mitigation strategies were employed to alleviate some of the effects of acid rain, and strives to model the recovery of these regions following decreases in atmospheric acid inputs.

Since the ultimate goal of the study is to estimate solute fluxes in ungaged catchments, several pertinent research questions emerged involving estimation of flow in ungaged catchments, interpolation of concentration values from limited concentration data, and the relative differences in estimated flux values from different chemistry aggregation methods. These questions evolved into three hypotheses, which are explained below.

The first hypothesis, henceforth referred to as Hypothesis 1 (H1), is that when predicting daily flow at the upstream location of a pair of nested stream gages, the area ratio of the two watersheds is the most important predictive watershed characteristic. The second hypothesis (H2) is that the prediction of daily flow at the upstream location is improved when relationships are developed and applied for multiple flow regimes at the downstream gage rather than when a single overall relationship is used. The third hypothesis (H3) regards the relative difference in flux estimates made using different solute chemistry aggregation methods. Hypothesis 3 is that various chemistry aggregation methods will produce similar annual flux estimates, but that agreement among these methods will decrease at shorter time scales. Hypothesis 3 is limited to solutes that do not exhibit strong concentration-discharge relationships, and therefore only considers the three averaging techniques discussed previously. For this study I considered dissolved organic carbon (DOC), nitrate (NO_3^-), total nitrogen (TN), and hydrogen ion (H^+) fluxes in four ungaged watersheds in the Catskill Park of New York State.

Chapter 2: Methods

2.1 Study Site

This study is set in the Catskill Park of New York State. The Catskills region, in which the Park lies, is a mountainous area that contains many small streams. Geologically, the mountains of the Catskills are the remains of the down-cutting of a large plateau by glaciers and the region's numerous streams (Titus, 1998). This plateau, known as the Catskill Delta, was the result of the erosion of the once massive Acadian Mountains, a product of the Acadian Orogeny (Titus, 1998). The figure below shows the location of the Catskill Park in New York State.

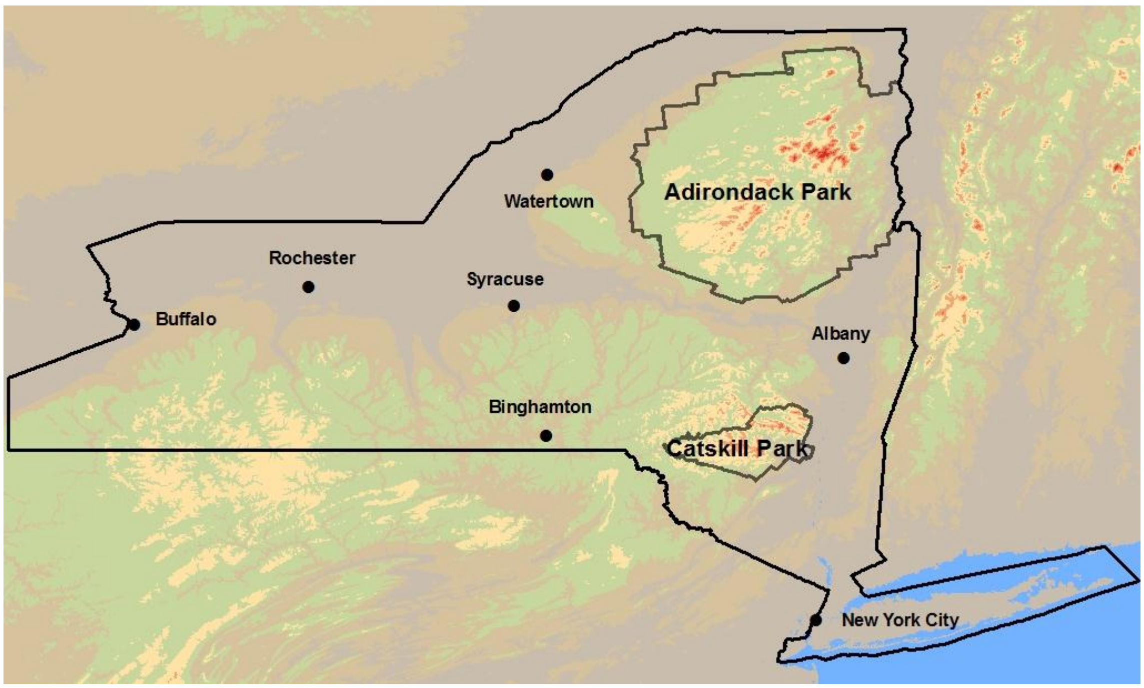

Figure 1: Map of New York State showing the geographic location of the Catskill Park.

The surficial and bedrock geology of the Catskills consists of both marine and terrestrial deposits. The marine deposits, responsible for the marine shales and limestones found in the area, are the product of deposition in the Appalachian Basin prior to and following the Acadian

Orogeny. The terrestrial sandstone and shale deposits are the result of glacial retreat some 10,000 years ago and the prior advancement of Catskill Delta in the Middle and Late Pleistocine (Titus, 1998; Kudish, 2000). Most of what is seen on the surface in the Catskills today is terrestrial sandstone and shale deposits left by the delta advancement and the glacial retreat (Kudish, 2000).

The geologically recent glacial activity in the Catskills is largely responsible for the region's surficial bedrock and soil characteristics. Most of the region's soils are underlain by glacial till, which is a poorly sorted mixture of glacial sediment derived from subglacial erosion that is deposited by the glaciers at their terminal and lateral extents. The presence of glacial till has had significant influence as a parent material on the development of the soils, as well as their corresponding hydrologic response (Kudish, 2000). Although plot scale heterogeneity in soil texture is common, the overwhelming majority of soils in the Catskills are classified as inceptisols, characterized by a sandy loam texture and poor horizon development (Kudish, 2000). Fragipans, dense cement-like layers that impede root growth and water infiltration, are also fairly common and widespread throughout the region (Kudish, 1979) Spodosols (acid forest soils) may be present at higher elevations in small pockets, and histosols (organic peat soils) have been reported in isolated patches (Kudish, 2000). Soil depth in the Catskills is variable, generally decreasing with elevation, and can range from as shallow as a few centimeters on ridges to well over a meter in valleys.

The modern result of the depositional and glacial history of the Catskills is a mountainous region with many stream valleys, but very few lakes or other natural water storage features. This, along with the region's proximity to New York City, has led to the development of several large water supply reservoirs to supply drinking water for New York City. To support the monitoring and modeling of the water supply for New York City, the study location contains a high density of currently and historically active United States Geological Survey (USGS) stream gaging stations. There are over thirty such currently active gaging stations within the Catskill Park's 1,120 mi^2.

This study focuses on the estimation of chemical mass fluxes in four ungaged headwater catchments distributed throughout the northern and southern Catskills (Figure 2). These catchments include the Fall Brook, Hunter Brook, Rondout Creek, and West Kill watersheds. The Fall Brook and Rondout Creek watersheds are located in Ulster County in the southern Catskills, while the Hunter Brook and West Kill watersheds are located in Greene County in the northern Catskills. All of the watersheds are heavily forested, dominated by the common northern deciduous species (Birch, Beech, and Maple), and lack urban areas or flow-altering structures. Each of these ungaged catchments is referenced to a larger, gaged catchment in the same drainage basin for flow estimation purposes by an area ratio method. The role of the area ratio of the ungaged catchment to the gaged catchment in estimating flows at the ungaged location is explored in H1 and is included in Table 1, which lists the sample and reference gage locations for each of the study sites.

Table 1: Sample site and reference gage locations.

Site	Area Ratio	Site Location (Lat/Long)	Reference Gage Location (Lat/Long)	Reference Gage Name	USGS Station Number
Fall Brook	0.1456	41°57'04", 74°34'04"	41°55'13", 74°34'30"	W. Branch Neversink River at Claryville	01434498
Hunter Brook	0.3883	42°11'06", 74°16'21"	42°11'06", 74°16'38"	West Kill below Hunter Brook near Spruceton	01349711
Rondout Creek	0.1715	41°55'39", 74°22'49"	41°51'59", 74°29'15"	Rondout Creek near Lowes Corners	01365000
West Kill	0.3602	42°10'39", 74°15'42"	42°11'06", 74°16'38"	West Kill below Hunter Brook near Spruceton	01349711

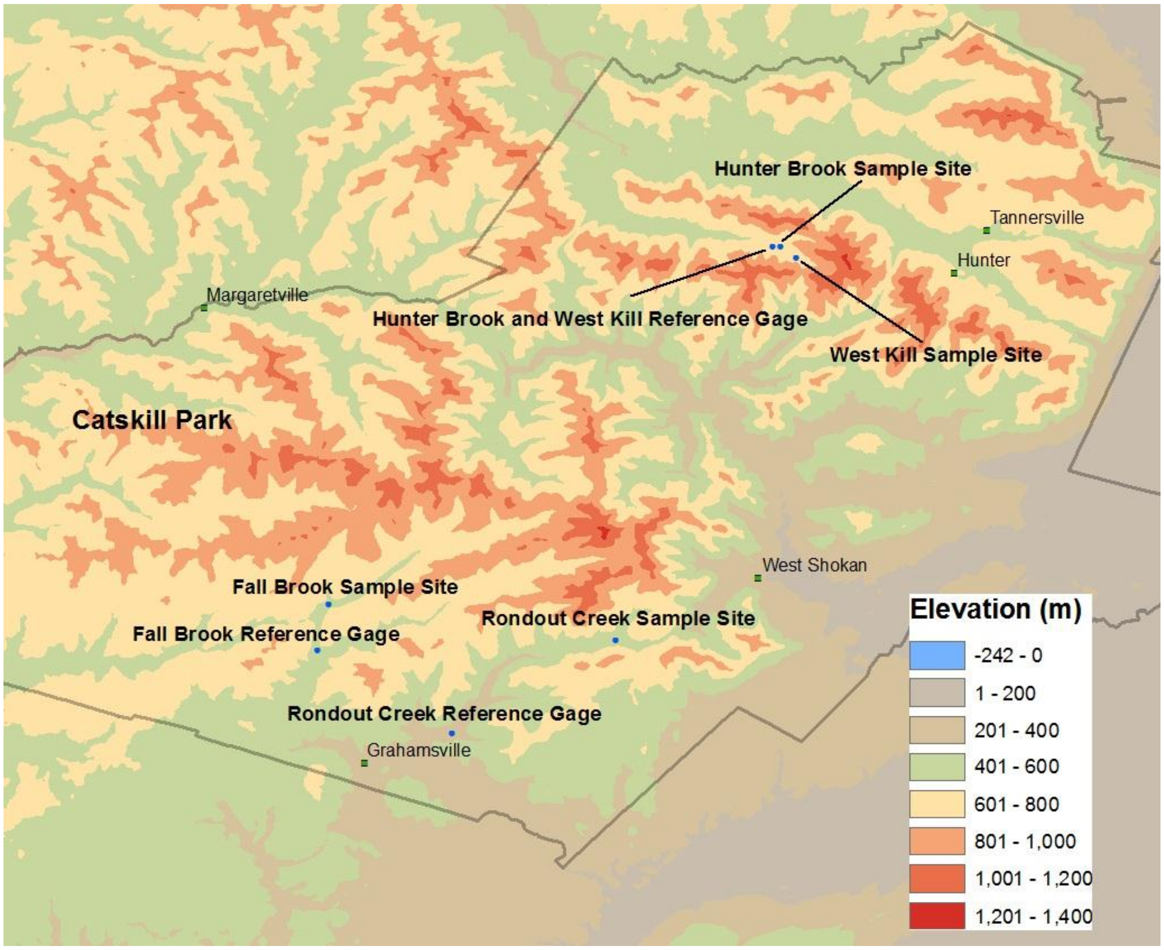

Figure 2: Map of the Catskill Park showing the locations of the four study sites.

2.2 Study Period

The period of interest for the estimation of chemical mass fluxes in this study is the 2011 water year, running from October 1, 2010 to September 30, 2011. However, much of the data used to develop the flow prediction models was generated prior to the study period in the form of historic average daily flow data taken from the USGS. Chemical data for stream water samples taken from outside the study period were also used for the development of concentration-discharge curves to examine general biogeochemical trends with the advantage of having more sample data and to verify the quality control measures being taken in the sampling and analytical methods.

2.3 Methods for Hypothesis 1

The hypothesis that the area ratio of the upstream and downstream watersheds is the most important basin characteristic for estimating flow at the upstream location was tested by locating nested pairs of stream gages in the study region and regressing flow between the paired sites. These pairs were located by inspection of a map containing active USGS gaging stations and with the help of the stream gage tool in the USGS StreamStats program (streamstats.usgs.gov). Criteria for selecting pairs for testing H1 included the requirements that: (1) both gaging stations be located within the political boundary of the Catskill Park; (2) the gaging stations have concurrent periods of record of at least four years; and (3) the watershed of neither gaging station be affected by man-made impoundments or other flow-altering devices.

Fifteen such pairs were identified and used for the testing of this hypothesis (Table 2). These pairs are henceforth referred to as model development pairs or model development sites. Although some of the reference gages for the model development sites are shared with the reference gages for the study sites, these sites are not co-located with any of the four study sites, at which mass flux estimates are to be made in this study. An attempt was made to use a representative set of nested pairs that covered the geographic region of the study site, and to select pairs that provided a variety of watershed area ratios. Table 2, on the following page, lists the gage locations and names for each of the model development pairs and Figure 3 shows their relative locations with the Park.

Table 2: Model development pair gage names and locations.

			Upstream Gages			Downstream Gages		
Pair	**Name**	**Area Ratio**	**Location (Lat/Long)**	**USGS Station Number**	**Gage Name**	**Location (Lat/Long)**	**USGS Station Number**	**Gage Name**
1	Batavia Kill	0.0296	42°17'22", 74°06'59"	01349840	Batavia Kill near Maplecrest, NY	42°18'30", 74°23'25"	01329950	Batavia Kill at Red Falls, NY
2	Biscuit Brook	0.1101	41°59'46", 74°30'01"	01434025	Biscuit Brook above Pigeon Brook at Frost Valley	41°55'13", 74°34'30"	01434498	W. Branch Neversink River at Claryville, NY
3	Bush Kill	0.5681	42°09'03", 74°36'06"	01413398	Bush Kill near Arkville NY	42°08'48", 74°37'25"	01413408	Dry Brook at Arkville
4	Hollow Tree	0.0631	42°08'32", 74°15'55"	01362342	Hollow Tree Brook at Lanesville, NY	42°06'07", 74°18'39"	01362370	Stony Clove Creek below Ox Clove
5	Rondout Creek	0.1399	41°56'13", 74°22'30"	01364959	Rondout Creek above Red Brook at Peekamoose, NY	41°51'59", 74°29'15"	01365000	Rondout Creek near Lowes Corners, NY
6	Winnisook Creek	0.0228	42°00'40", 74°24'53"	01434021	W. Branch Neversink at Winnisook Lake	41°55'13", 74°34'30"	01434498	W. Branch Neversink River at Claryville, NY
7	West Kill	0.1841	42°11'06", 74°16'38"	01349711	West Kill below Hunter Brook near Spruceton, NY	42°13'49", 74°23'36"	01349810	West Kill near West Kill, NY
8	East Kill	0.3678	42°14'57", 74°18'11"	01349700	East Kill near Jewett Center, NY	42°14'13", 74°20'26"	01349705	Schoharie Creek near Lexington, NY
9	Beaver Kill Trib.	0.0051	42°04'59", 74°10'59"	01362465	Beaver Kill Tributary above Lake Hill, NY	42°00'51", 74°16'16"	01362500	Esopus Creek at Cold Brook, NY
10	Little Beaver Kill	0.0859	42°01'10", 74°16'00"	01362497	Little Beaver Kill at Beechford near Mt. Tremper, NY	42°00'51", 74°16'16"	01362500	Esopus Creek at Cold Brook, NY
11	Woodland Creek	0.1073	42°04'47", 74°20'05"	0136230002	Woodland Creek above Mouth at Phoenicia, NY	42°00'51", 74°16'16"	01362500	Esopus Creek at Cold Brook, NY
12	E. Branch Neversink	0.3900	41°58'01", 74°26'54"	0143400680	E. Branch Neversink River Northeast of Denning, NY	41°55'31", 74°32'26"	01434017	E. Branch Neversink River near Claryville, NY
13	Esopus River	0.0242	42°02'01", 74°25'15"	01362192	Panther Mtn. Trib. to Esopus near Oliverea, NY	42°07'01", 74°22'50"	01362200	Esopus Creek at Allaben, NY
14	High Falls Brook	0.0811	41°58'38", 74°31'21"	01434105	High Falls Brook at Frost Valley, NY	41°55'13", 74°34'30"	01434498	W. Branch Neversink River at Claryville, NY
15	Birch Creek	0.1962	42°06′32″, 74°27′08″	013621955	Birch Creek at Big Indian	42°07'01", 74°22'50"	01362200	Esopus Creek at Allaben, NY

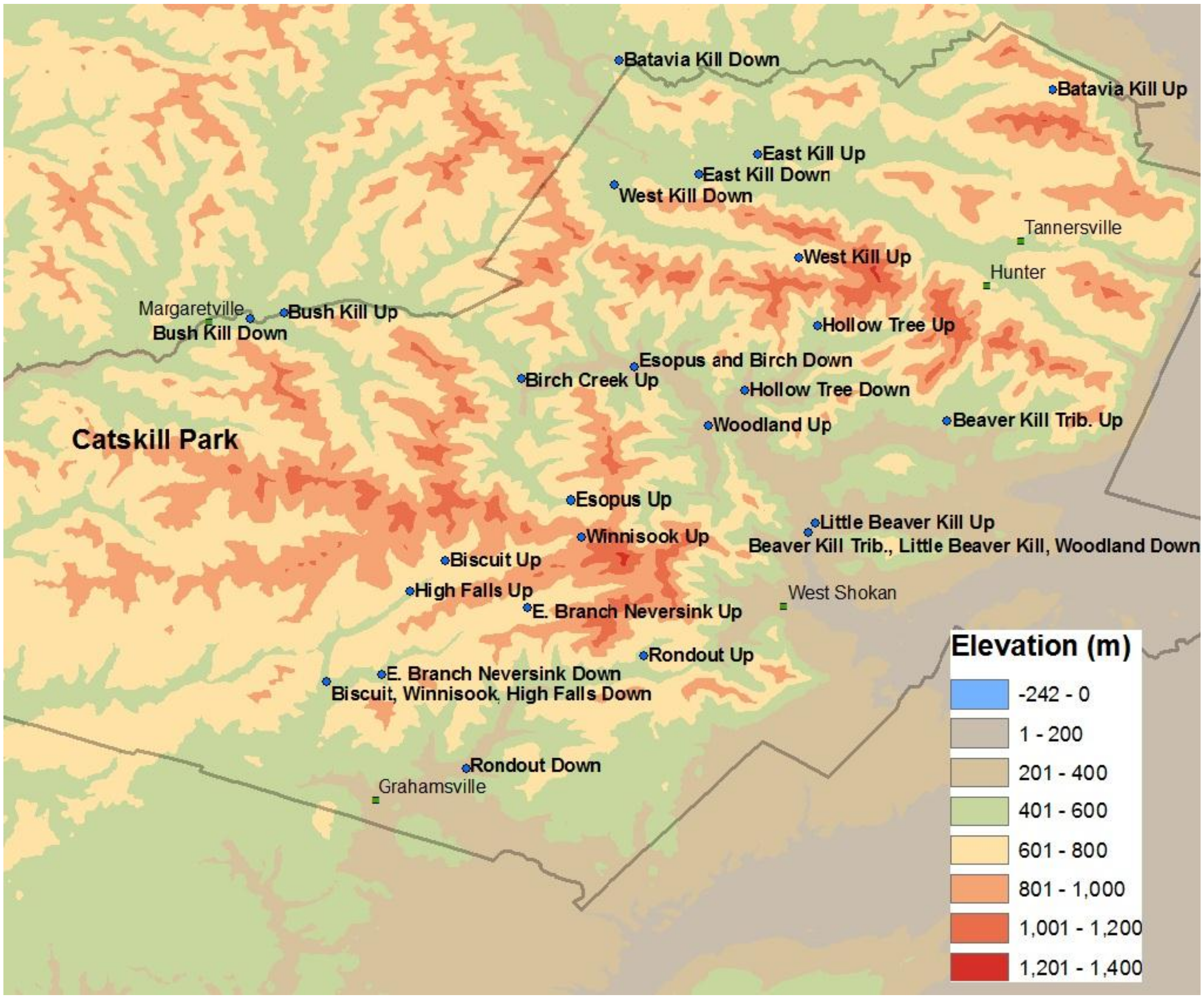

Figure 3: Map of the Catskill Park showing the locations of the model development sites.

Concurrent average daily flow data for each of the model development pairs were downloaded from the USGS National Water Information System website and analyzed using spreadsheet software. The flow at the upstream site was plotted against the flow at the downstream site and a functional relation (slope and intercept) was developed for each pair. A functional relation is a line passing through the plotted data points that minimizes the distance from each point to the line. The slope of the functional relation for each pair was then regressed (plotted) against the area ratio for each pair to determine the importance of the area ratio in predicting flow at an upstream location based on a downstream reference gage. The scaling issue of the area ratio method was verified by normalizing the average daily flows by their respective watershed areas and plotting the functional relation for each pair. If flow truly scales by area and the upstream and downstream flow per unit area are indeed the same, the slope of the functional relation of these area-normalized plots should be equal to 1. For the purposes of a general model, the functional relation slopes for all fifteen area-normalized model development pairs were averaged.

Since the nature of this study involves regressing published flow data, it was determined that a functional relation should be used rather than a least-squares regression relationship. Functional relations are used when the regression assumption that there is no error in the independent variable is unacceptable (Webster, 1997). When relating measured flow data at two gaged sites there is obviously error in both the dependent and independent variables. Webster (1997) offers several alternatives for the variance structure in this situation. In this study, I assumed that the errors in the dependent and independent variables were proportional to their respective variances. The form of the linear functional relations derived in this study is the familiar equation:

$$y = \beta_0 + \beta_1 x \qquad [2]$$

Where: y = Flow at the upstream gage (ft^3/s)

β_0 = Intercept Term = Flow at the upstream gage when flow at the downstream gage is zero (ft^3/s)

β_1 = Slope of the functional relation

x = Flow at the downstream gage (ft^3/s)

The slope of the functional relation in the proportional error case can be calculated using the following equation, as presented by Webster (1997):

$$\hat{\beta}_1 = \sqrt{\lambda} \qquad [3]$$

Where: $\hat{\beta}_1$ = Estimated slope of the functional relation

$\lambda = s_y^2/s_x^2$
s_y^2 = Variance of the flow dataset at the upstream gage
s_x^2 = Variance of the flow dataset at the downstream gage

The physical meaning of the intercept of the functional relation pertains to whether or not there is flow at the upstream location when there is no flow at the downstream location. Hydrologically, one would expect the intercept term to be near zero, or perhaps negative, indicating that flow in the upstream catchment ceases before flow in the downstream watershed. Computationally, the intercept of the relation is found by the following equation, also adapted from Webster (1997):

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X} \qquad [4]$$

Where: $\hat{\beta}_0$ = Estimated intercept of the functional relation (ft^3/s)
$\bar{Y}$ = Average of the upstream daily flow dataset (ft^3/s)
$\hat{\beta}_1$ = Slope of the functional relation
$\bar{X}$ = Average of the downstream daily flow dataset (ft^3/s)

Once all of the functional relation slopes and intercepts were calculated, the slopes were plotted against their respective area ratios. The watershed area of each model development site was found on the USGS real-time flow webpage for each gage. In developing these relationships, it may be reasonably assumed that the dependent variable (the watershed area ratio) is known without error, so linear regression was used to determine the equation of the relationship. The statistical significance of the relationship was then tested and a coefficient of determination (R^2) calculated in order to quantify the importance of the area ratio in estimating daily flow at the upstream location of a pair of nested stream gages. The coefficient of determination, or R^2 value, is a measure of how well the model explains the variability in the data set used.

2.4 Methods for Hypothesis 2

The second hypothesis, which suggests that relationships developed from partitioned downstream flow data sets will better predict flow at the upstream gage than relationships developed from the whole data set at once, was tested by calculating and comparing the root mean square error of flow values at the upstream sites predicted by both approaches. This hypothesis was tested using the same fifteen model development pairs used in Hypothesis 1.

The concurrent period of record for each pair was divided in two, with the first half of the data set used for relationship development and the second half of the data set used for relationship testing. An overall functional relation was developed for the entire first half of the data set using the approach outlined in section 2.3. Developing relationships for the partitioned data sets was done by dividing the first half of the data set into low, medium, and high flow regimes, represented by data points that fell within the 5-25%, 26-75%, and 76-95% flow percentiles for each site of the pair. In order for the data point for a particular day to be considered valid, the flow for both the upstream and downstream sites of the pair had to fall within the low, medium, or high range for their respective data sets.

Once the data sets were trimmed and partitioned into flow regimes a functional relationship for each flow regime was calculated and used to predict flow at the upstream site in the second half of the data set. The specific flow regime relationship (low, medium, or high) used for prediction was determined by the magnitude of flow on a given date at the downstream gage location, using low, medium, and high flow ranges based on the flow percentiles at the downstream location. The predicted values of flow in the second half of the data set at the upstream gages were then compared to the actual values of flow, and the root mean square error for each pair was calculated.

Root mean square error (RMSE) is used to measure the difference between values predicted from a model and actual values. In our case, it gives us insight into whether or not partitioning the data sets and developing separate relationships for the different flow regimes actually improves the estimates of flow, as compared to the relationship developed from the overall data set. Root mean square error has the same units as the input data, in this case ft^3s^{-1}, and is calculated using the following equation:

$$\text{RMSE} = \sqrt{\frac{\sum(Y-\hat{Y})^2}{n}} \qquad [5]$$

Where: RMSE = Root mean square error of predicted flow value (ft^3/s)
Y = Actual upstream flow value at gaging station (ft^3/s)
$\hat{Y}$ = Predicted upstream flow value at gaging station (ft^3/s)
n = Number of data points

A smaller RMSE for the partitioned relationships would indicate that the predictive power of the relationships benefitted from partitioning the data sets into flow regimes and applying separate flow regime relationships. A smaller RMSE for the non-partitioned relationships would indicate that partitioning the data sets and developing separate relationships for each flow regime did not improve the predictions of flow.

2.5 Methods for Hypothesis 3

To test the third hypothesis, I calculated chemical mass fluxes using estimated flow values and interpolated concentration values for the four ungaged study sites. First, concentration-discharge relationships were examined for the solutes of interest for this study. Concentration-discharge plots were developed from the measured grab sample concentrations and the estimated flow values on the corresponding sample date. Even though this study only considers the 2011 water year for flux estimates, all available chemistry data from the broader study outlined in Section 1 were used to generate the concentration-discharge curves. This amounted to either 22 data points (for the Fall Brook and Rondout Creek watersheds) or 20 data points (for the Hunter Brook and West Kill watersheds), depending on the number of samples taken prior to data analysis.

I attempted to fit linear, exponential, and power-law relationships to the concentration-discharge data in an effort to determine if any relationship existed between flow and solute concentration for each analyte at each site. Only 3 of the 48 relationships exhibited an R^2 value of 0.5 or greater, suggesting that none of the solutes considered were strongly correlated with flow. The concentration-discharge curves for each solute and site are included at the back of this document in Appendix B.

Daily flow values for the ungaged study catchments were estimated using an adjustment coefficient based on the area ratio-based functional relations developed for the partitioned data sets used in Hypothesis 2. The partitioned data set relationships were used because they generally showed improved prediction of flow at the upstream location of a nested pair in which the area ratio of the upstream and downstream watershed areas is greater than 0.1, which was true for all of the study sites. Since I was always considering catchments nested within a larger gaged basin, the reference stream gage was always chosen as the nearest downstream gage on the same stream network. It should be noted that the adjustment coefficient is the *predicted* functional relation slope from the functional relation slope versus area ratio regression produced to test H1, because data do not exist to actually calculate a functional relation for those sites.

The watershed area of the ungaged catchments was estimated using the watershed delineation tool in the USGS StreamStats program and the area ratio for each study site was determined. The adjustment coefficients were found by inserting the area ratios into the functional relation-area ratio relationships developed for each flow regime in H2. These plots are shown in the Results section of the paper (Section 3.3).

In this method, daily average stream flow was estimated using the following relation:

$$\hat{Q}_u = \gamma Q_g \quad [6]$$

Where $\hat{Q}_u$ = Estimated average daily flow in the ungaged catchment (ft^3/s)
γ = Flow adjustment coefficient
Q_g = Measured average daily flow in the gaged reference catchment (ft^3/s)

Concentration values were interpolated from the measured monthly concentration data using three methods: a period-weighted approach in which a measured concentration value was applied to a period of days before and after the sampling date, an annual average method in which the annual average concentration was applied to all days of the study period, and a seasonal average method, in which a summer and winter average concentration was calculated and applied to the corresponding period of dates. The period used for the period-weighted approach was established by the midpoints in time between each set of adjacent sampling dates. The annual average approach considered the annual average solute concentrations for the 2011 water year. The summer average used in the seasonal average method considered concentration values from April to September, while the winter average considered values from October through March of the 2011 water year.

Hypothesis 3 was tested by calculating the relative difference in chemical mass flux estimates for each chemistry aggregation method on an annual, monthly, and daily basis. Relative difference describes how similar two values are, relative to each other. Fluxes were calculated as the product of estimated average daily stream flow and the interpolated daily concentration value using the above methods, normalized by the watershed area. Monthly and daily relative differences were averaged for comparison to annual relative differences. The relative difference between each method was then averaged over the four study sites to determine the general trends in the relative difference between chemistry aggregation methods. The hypothesis would be accepted if the average relative difference over the four study sites was least for flux estimates made on an annual basis and increased at each the monthly and daily time scales. It should be noted that the downloaded flow data used in Hypotheses 1 and 2 was left in English units for the flow estimation components of this project, but converted to Metric units for calculation of fluxes in Hypothesis 3. This decision was based on the units of the source data and the convention to display chemical mass fluxes in SI units.

2.6 Stream Sampling Methods

Stream sampling of the four ungaged watersheds (study sites) was done monthly by the author or other members of the research team. Sampling locations were kept consistent within a few meters and care was taken not to disturb the sample site prior to sampling. A 2 liter grab sample was taken near the centroid of flow from each stream and the sample bottle was rinsed three times with stream water prior to sample collection.

The samples were filtered within 2 hours using a 0.45-μm membrane filter and vacuum hand pump apparatus. Samples were kept cool prior to filtration and while being transported from the field to the lab. Once in the lab, samples were stored at approximately $4^{o}C$ until processing, which was always within the sample holding time for each particular analyte. Table 3 shows the maximum length of time samples were stored in the lab before processing.

Table 3: Maximum holding time of samples prior to processing.

Analyte	Maximum Length of Time Stored Before Processing
DOC	28 d
NO_3^-	36 hr
TN	28 d
pH	10 d

Quality control was maintained in the form of field blanks, duplicate samples, continuing calibration verification standards (CCVs), and laboratory blanks. Field blanks were brought into the field as 2 liter bottles containing deionized water and were treated identically as stream samples. Duplicate samples were taken from selected sites each month and concentrations compared for consistency. Continuing calibration verification standards and laboratory blanks will be explained in the following section.

2.7 Analytical Methods

The analytes of interest for this study were dissolved organic carbon (DOC), nitrate (NO_3^-), total nitrogen (TN), and hydrogen ion (H^+). All sample processing took place at Syracuse University within the standard holding time for each analyte. Dissolved organic carbon was measured using an Apollo 9000 TOC Combustion Analyzer manufactured by Teledyne Tekmar. This procedure involves acidification and purging of dissolved inorganic carbon (DIC) followed by oxidation of DOC by combustion and measurement of subsequent CO_2 product (Teledyne Tekmar, 2003).

Nitrate was measured using the Dionex DN-500 Ion Chromatography System manufactured by Thermo Fischer Scientific. This ion chromatography procedure uses ion-specific exchange columns to sorb anions and a potassium hydroxide eluent to release them for measurement by a conductivity detector (CESE-IC).

Total nitrogen concentrations were measured using the Apollo 9000 Combustion Analyzer with a TN Module. In this procedure all nitrogen species (nitrate, nitrite, ammonium, and organic nitrogen) are converted to NO gas by combustion and mixed with ozone. The light emitted from the resulting chemoluminescent reaction is directly proportional to the amount of NO in the sample gas, and is measured by a photodiode, which converts the light energy into electrical energy for quantification of nitrogen by a standard curve (CESE-TN).

Hydrogen ion concentrations were calculated following measurement of pH. Hydrogen ion is calculated from pH using the following relationship:

$$[H^+] = 10^{-pH} \qquad [7]$$

Where $[H^+]$ = The concentration of hydrogen ion (mol/L)
pH = The measured pH of the water sample

The pH was measured using a silver-chloride combination electrode manufactured by Thermo Fischer Scientific and the Brinkmann Metrohm 716 DMS Titrino pH Meter. No ion activity correction was applied in the calculation of $[H^+]$ because the ionic strength of the samples was very low for all sites and sampling dates.

Continuing calibration verification standards were maintained for each procedure and were analyzed every tenth sample in the processing order. Laboratory blanks followed CCVs in the DOC, NO_3^-, and TN analyses and were comprised of deionized water. For pH analysis, CCVs took the form of pH buffers. The quality control data and measures from the broader study were used by this study as evidence for quality assurance in the stream sample analytical data. Tables containing more information on the field blanks, duplicate samples, and CCVs can be found in Appendix D at the end of this document.

Chapter 3: Results

3.1 Results for Hypothesis 1

Hypothesis 1, that area ratio is the dominant factor in estimating flows based on reference gages in upstream locations of nested catchments in the Catskill Park, was found to be reasonable. This is first demonstrated by the strength of the flow-flow relationships that were developed for each of the model development pairs. These plots generally produced strong linear relationships, all of which were statistically significant ($\alpha = 0.05$). An example regression for pair 3, Bush Kill, is included below.

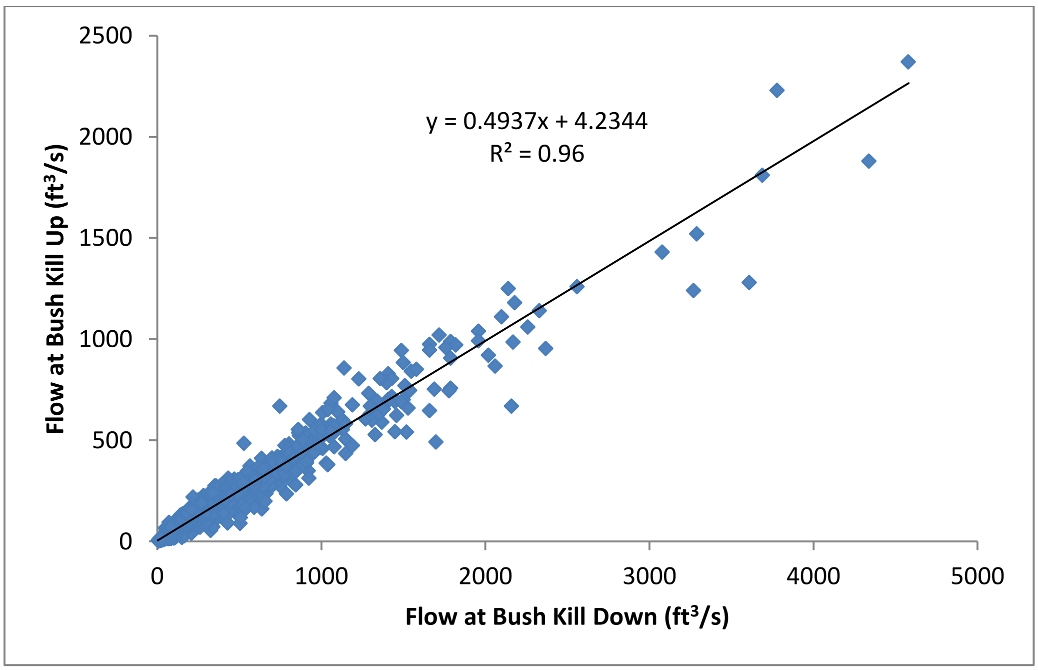

Figure 4: Example flow-flow regression for the Bush Kill model development site.

The line shown on Figure 4 above is the functional relation for the flow to flow comparison. Note the very high coefficient of determination for this particular example ($R^2 = 0.96$). The area ratio for the Bush Kill model development site is 0.568. A table containing the functional relation equations and coefficients of determination for each model development pair can be found in Appendix A at the end of this document.

After establishing the strength of the flow-flow relationships, it was verified that flow in Catskills streams generally does scale directly with watershed area. This is shown by Figure 5, which shows the area-normalized flow to flow relationship for model development pair 3, Bush Kill, and Table 4, which shows the functional relation slopes for the area-normalized flow to flow plots.

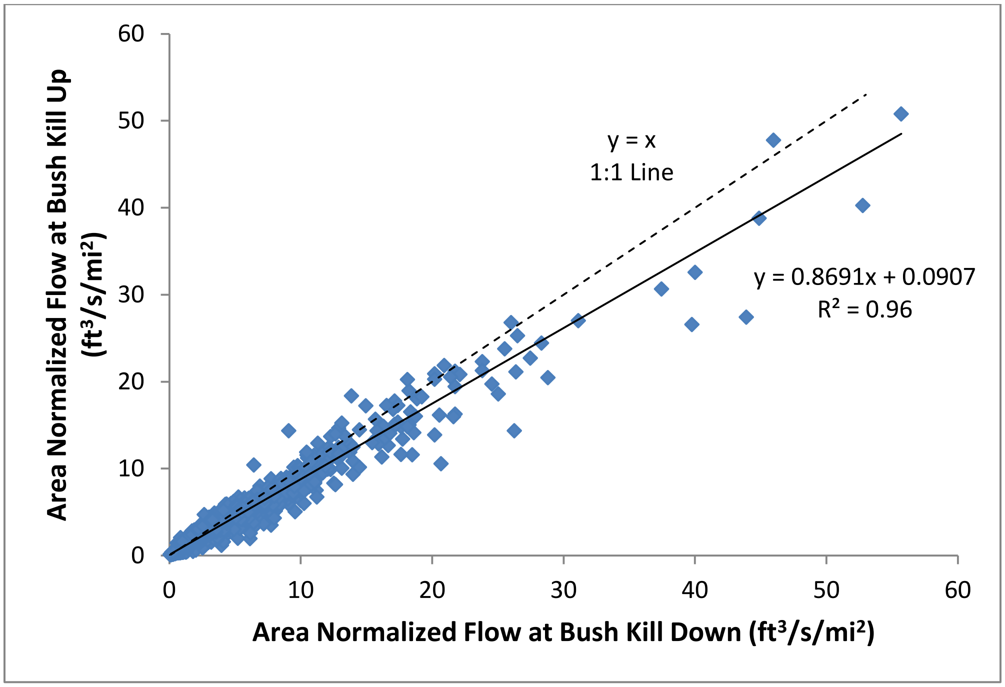

Figure 5: Area normalized flow-flow relationship for Bush Kill.

The dashed lined in Figure 5 above represents a 1:1 line, or the line that represents perfect scaling of stream flow with watershed area.

Table 4: Area-normalized functional relation slopes for the model development pairs.

Model Development Site	Area Ratio	Area-Normalized Functional Relation Slope
Batavia Kill	0.0296	1.50
Biscuit Brook	0.1101	0.89
Bush Kill	0.5681	0.87
Hollow Tree	0.0631	0.74
Rondout Creek	0.1399	1.22
Winnisook Creek	0.0228	1.33
West Kill	0.1841	1.16
East Kill	0.3678	0.87
Beaver Kill Trib.	0.0051	0.96
Little Beaver Kill	0.0859	0.96
Woodland Creek	0.1073	1.18
E. Branch Neversink	0.3900	1.28
Esopus River	0.0242	1.26
High Falls Brook	0.0811	0.61
Birch Creek	0.1962	0.72
Average	---	1.04

Although some of the functional relation slopes are indeed quite different from 1, the average across the fifteen mode development sites is very close to 1. This was considered verification that, across the Catskill region as a whole, flow does indeed scale directly with area and that area ratio methods are appropriate for the study sites.

The significance of the area ratio in predicting daily flow was then tested and found to be very important. This is demonstrated by Figure 6, which compares the functional relation slope for each model development pair to its corresponding area ratio. Figure 6 also contains a dashed 1:1 line to graphically demonstrate how close the least squares regression slope is to 1.

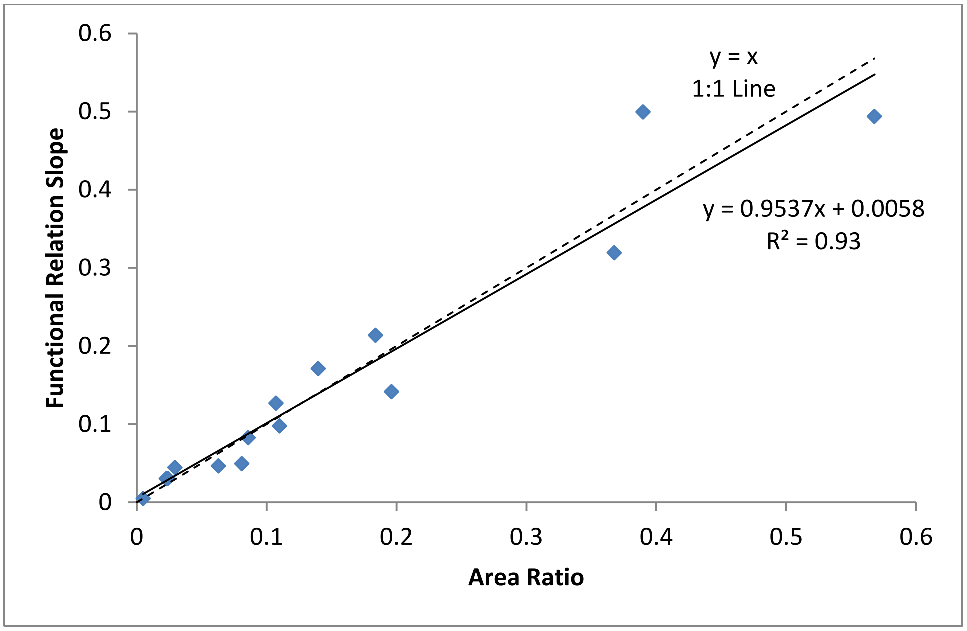

Figure 6: Functional relation slope versus area ratio for the model development pairs.

This relationship was found to be statistically significant ($\alpha = 0.05$, $p = 8.17\text{x}10^{-9}$), with the slope not significantly different from 1 ($p < 0.05$), and the intercept not significantly different from 0 ($p < 0.05$). Also, as can be seen in the figure, the coefficient of determination for this relationship was found to be 0.93, indicating that the area ratio alone accounts for 93% of the observed variation in the slopes of the functional relations relating upstream to downstream flow in the fifteen model development pairs.

.

3.2 Results for Hypothesis 2

The second hypothesis, which suggests that relationships developed for different flow regimes at a reference gage will better predict flow at the upstream gage of a nested pair than a relationship developed from the full data set of the reference gage, does not appear to be a reasonable claim when considering all of the model development sites. This is demonstrated by the fact that only five of the fifteen model development pairs showed improved prediction of average daily flow with relationships developed from the partitioned data sets. Table 5, below, lists the model development pairs, their area ratios, and the root mean square error of the predicted flow for the relationships developed from the full and partitioned data sets.

Table 5: Root mean square error of the predicted flows from the full and flow regime separated relationships.

Pair	Name	Area Ratio	Full RMSE (ft^3/s)	Partitioned RMSE (ft^3/s)
1	Batavia Kill	0.0296	6.30	6.52
2	Biscuit Brook	0.1101	7.56	9.04
3	Bush Kill	0.5681	28.72	28.62
4	Hollow Tree	0.0631	5.75	6.95
5	Rondout Creek	0.1399	11.18	9.15
6	Winnisook	0.0228	3.32	3.35
7	West Kill	0.1841	10.26	9.68
8	East Kill	0.3678	36.23	34.53
9	Beaver Kill Trib.	0.0051	2.10	2.18
10	Little Beaver Kill	0.0859	47.02	52.29
11	Woodland Creek	0.1073	38.03	47.21
12	E. Branch Neversink	0.3900	15.06	14.29
13	Esopus Creek	0.0242	4.48	5.34
14	High Falls Brook	0.0811	2.53	4.47
15	Birch Creek	0.1962	13.77	13.99

Upon further investigation of Table 5 however, it becomes clear that all five of the pairs that did benefit from partitioning have area ratios greater than 0.1. In fact, five of the eight model development pairs with area ratios greater than 0.1 showed improved prediction of average daily flow with relationships developed from the partitioned data sets. It should also be noted that none of the seven pairs with area ratios less than 0.1 showed improvement with the same method. When only considering pairs with area ratios greater than 0.1, partitioning the data sets and developing separate flow regime relationships improves the predictive power of the relationships in five out of the eight cases. This was used as justification for proceeding with data set partitioning and separate flow regime relationship development for flow estimates in the mass flux estimation method outlined in section 2.5. It should be noted however, that the improvement that resulted from partitioning was relatively small except in two cases (Rondout Creek and East Kill).

An example of the flow partitioning for model development pair 12, East Branch of the Neversink River, is included below for illustration. Figures 7-12 below include the full data set regression, the low, medium, and high flow regime partitioned regressions, and plots of the actual versus predicted flow values for the flows predicted from the full data set relationship and the partitioned data set relationships.

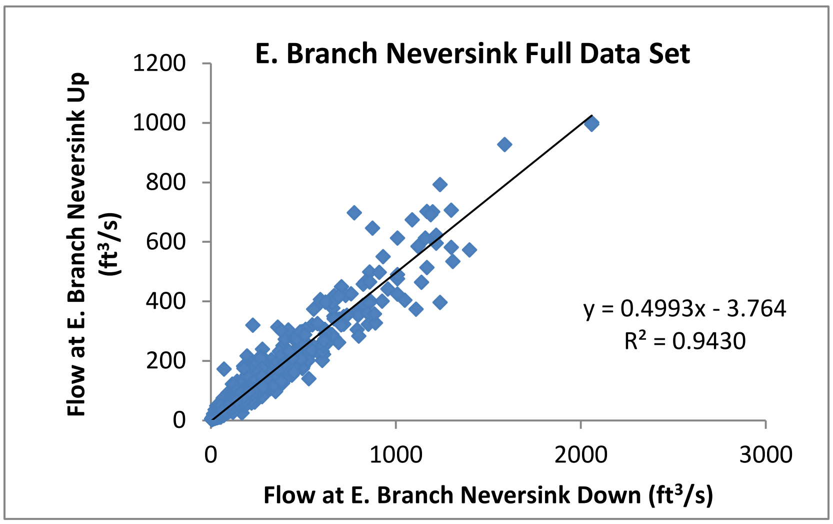

Figure 7: E. Branch Neversink full data set functional relation.

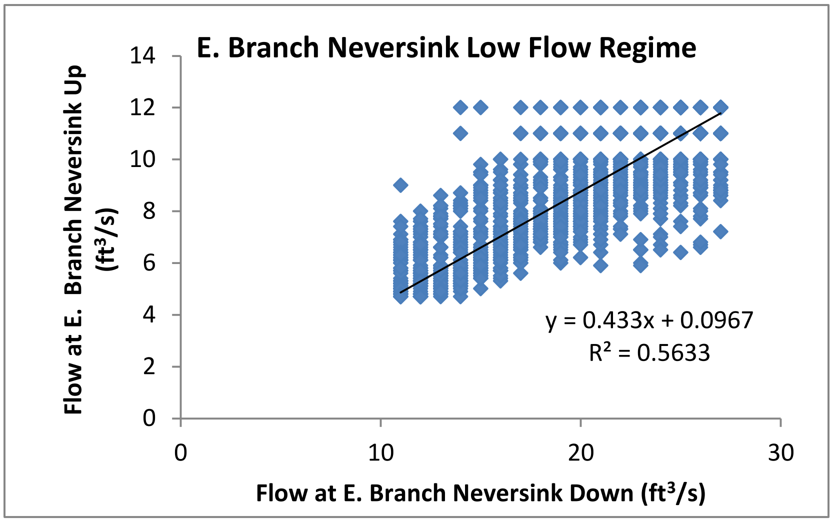

Figure 8: E. Branch Neversink low flow regime functional relation.

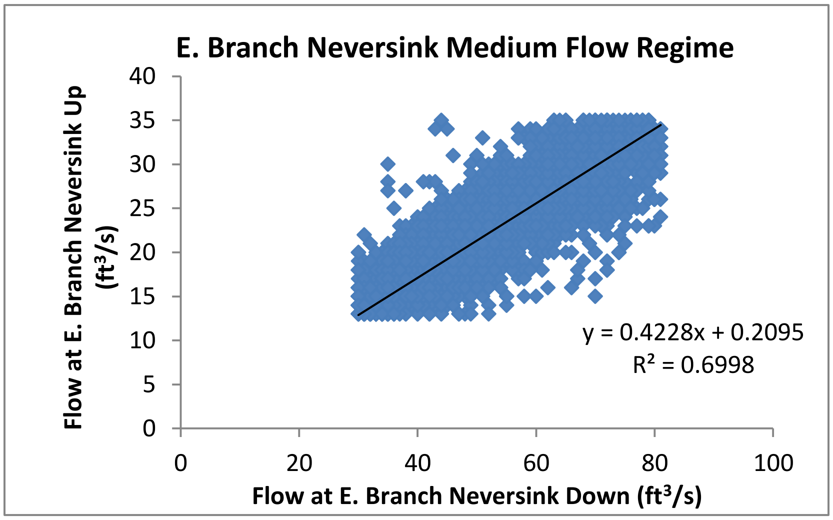

Figure 9: E. Branch Neversink medium flow regime functional relation.

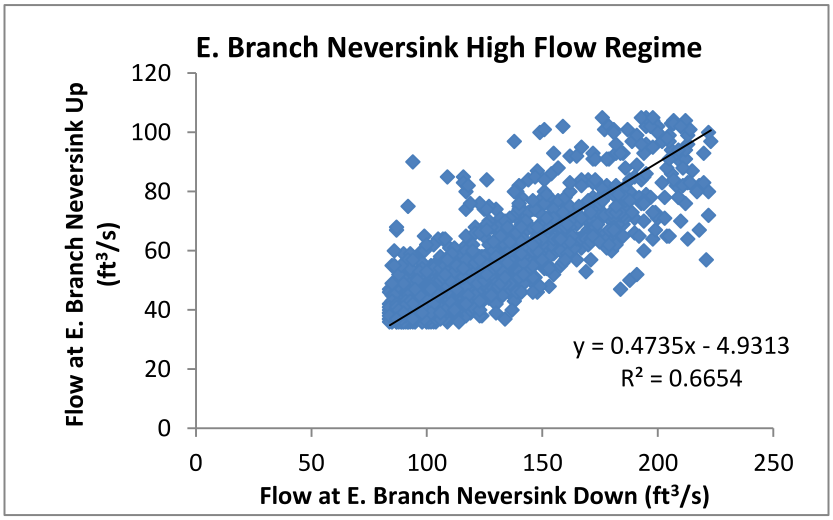

Figure 10: E. Branch Neversink high flow regime functional relation.

The area ratio for E. Branch Neversink is 0.39. It can be seen from Figures 7 through 10 above, that the relationships from the partitioned data sets (Figures 8-10) all have functional relation slopes that are closer to the area ratio than the relationship developed from the full data set (Figure 7). The actual versus predicted flow plots are shown in Figures 11 and 12 below.

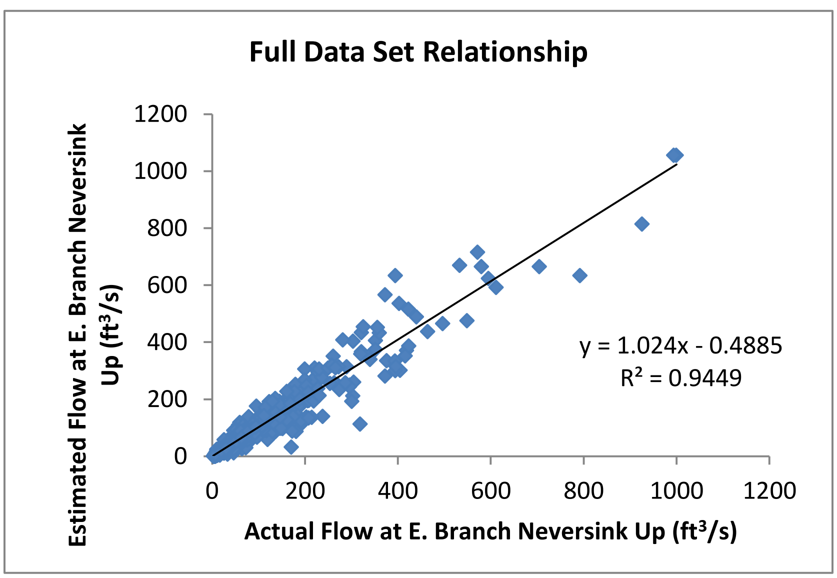

Figure 11: Actual versus predicted flow for E. Branch Neversink using full data set relationships.

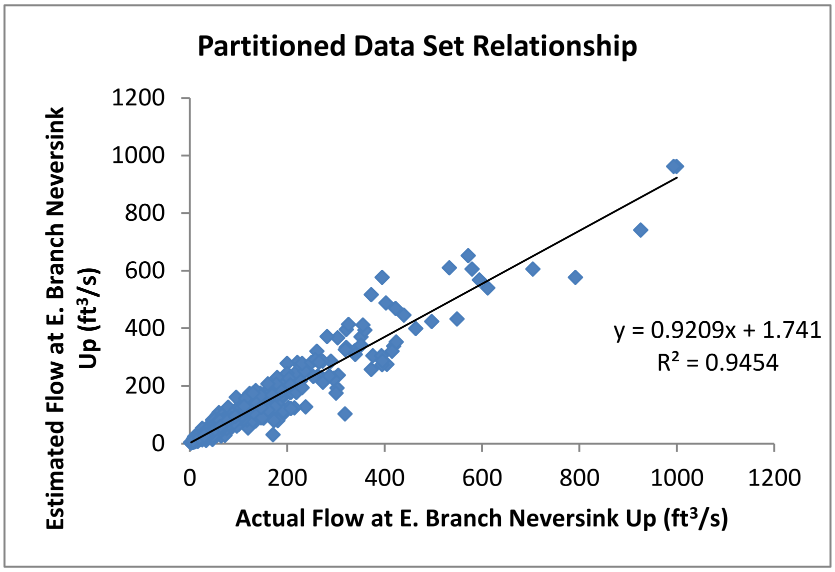

Figure 12: Actual versus predicted flow for E. Branch Neversink using separate flow regime relationships.

Both relationships in Figures 11 and 12 were found to be statistically significant (α=0.05, p=0), with the partitioned data set having a slightly higher R^2 value. Additional evidence in support of data set partitioning is the following plot, which relates correlation coefficients for the medium flow regime partitioned flow data sets to the area ratio for each pair of sites. The correlation coefficient (r) measures the linear dependence of the response variable on the independent variable, and describes how well a linear function can explain the trend in the data.

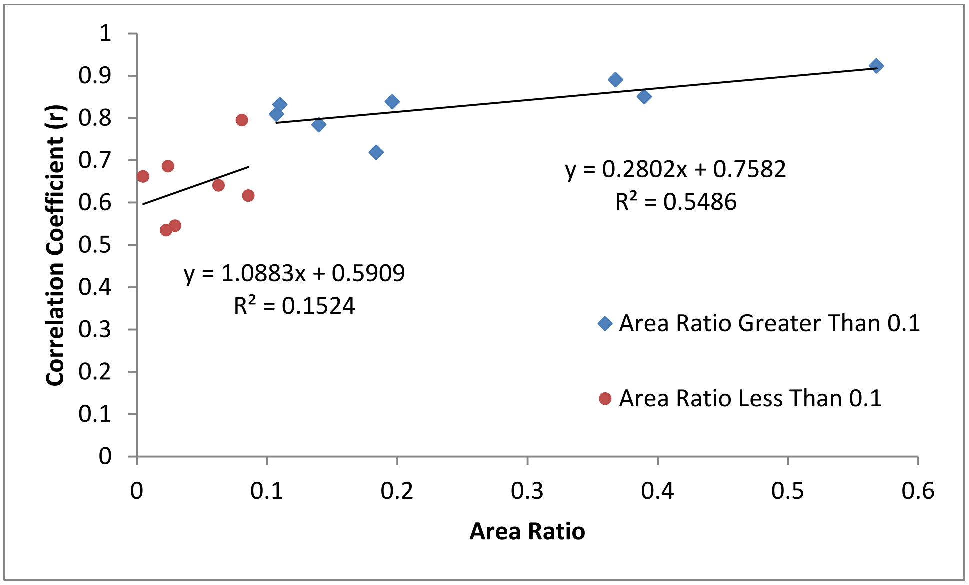

Figure 13: Comparison of area ratio to correlation coefficient.

When plotted as two separate series, it becomes clear that the model development pairs with area ratios less than 0.1 generally had lower correlation coefficients than the pairs with area ratios greater than 0.1. Furthermore, the scatter of the data points is greater for the smaller area ratio pairs. Plots for the low and high flow regimes yielded similar results.

3.3 Results for Hypothesis 3

The third hypothesis, which suggests that various solute chemistry aggregation methods will produce similar annual flux estimates, but that the relative difference among the methods will increase with decreasing time step, was also found to be a reasonable claim. None of the analytes of interest correlated well with flow based on the constructed concentration-discharge curves (Appendix B). It should be noted that these concentration-discharge curves included concentration data points collected over a wide range of flows (generally 5%-80% flow percentiles). Only 3 of the 48 concentration-discharge relationships yielded a coefficient of determination greater than 0.5. Thus the assumption that DOC, NO_3^-, TN, and H^+ are generally independent of flow, or exhibit more complex behavior than can be predicted from a simple concentration-discharge relationship for our study sites, seems reasonable. To test this hypothesis, average daily flows in the ungaged catchments had to be estimated using the adjustment coefficients developed from the functional relation-area ratio relationships for each flow regime. These relationships are shown in Figures 14-16 below.

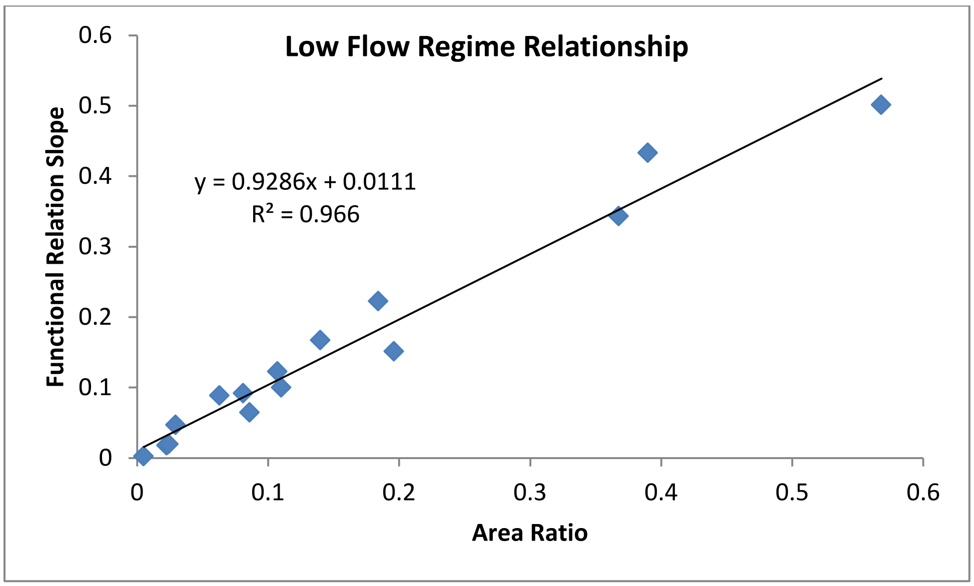

Figure 14: Relationship developed for determining adjustment coefficients at low flow.

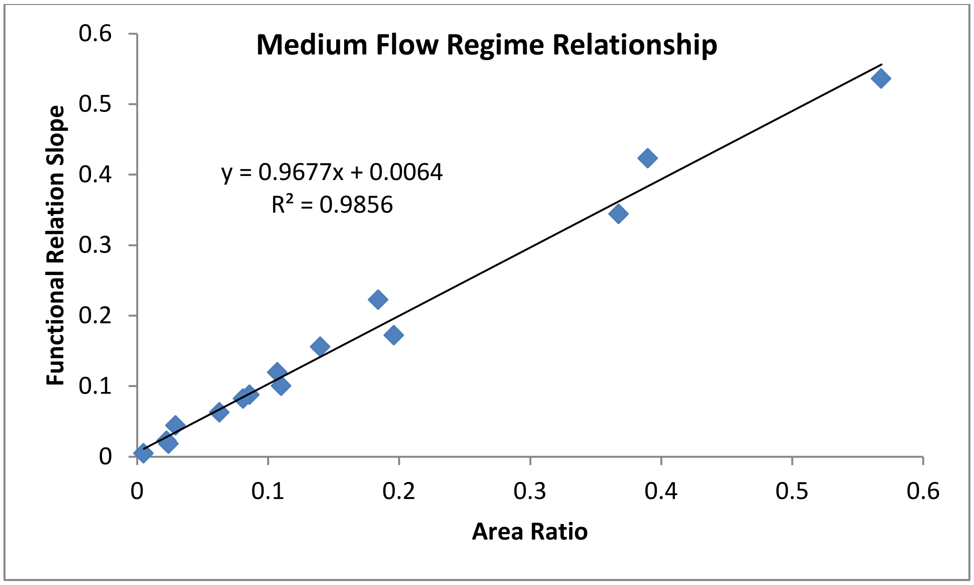

Figure 15: Relationship developed for determining adjustment coefficients at medium flow.

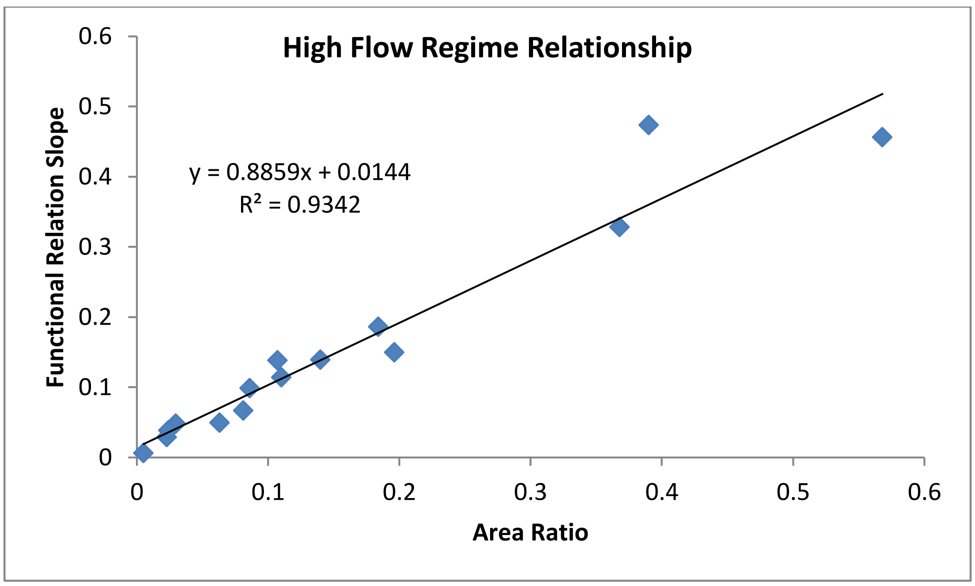

Figure 16: Relationship developed for determining adjustment coefficients at high flow.

The partitioned relationships were used because the study sites all had area ratios greater than 0.1, and as shown by the results for H2, pairs with area ratios greater than 0.1 generally benefit from data partitioning and development of separate flow regime relationships. Table 6 below shows the adjustment coefficients for each study site and flow regime, which were calculated using the area ratio for each study site and the above relationships.

Table 6: Flow adjustment coefficients for each study site and flow regime.

		Adjustment Coefficient		
Study Site	Area Ratio	Low Flow	Medium Flow	High Flow
Fall Brook	0.1456	0.1463	0.1473	0.1434
Hunter Brook	0.3883	0.3717	0.3822	0.3584
Rondout Creek	0.1715	0.1704	0.1724	0.1664
West Kill	0.3602	0.3455	0.3549	0.3335

The hypothesis was tested using the average relative difference in flux estimates for the four analytes for each study site calculated on an annual, monthly, and daily basis using the period-weighted, annual average, and seasonal average approaches. These results are displayed graphically below as bar charts representing the average relative difference over the four study sites in Figures 17-20. In each chart the period weighted approach is denoted by “PW”, the annual average approach by “AA”, and the seasonal average method by “SA”. Numerical tables showing the actual flux estimates and relative difference for each study site and chemistry aggregation method are included in Appendix C at the end of this thesis.

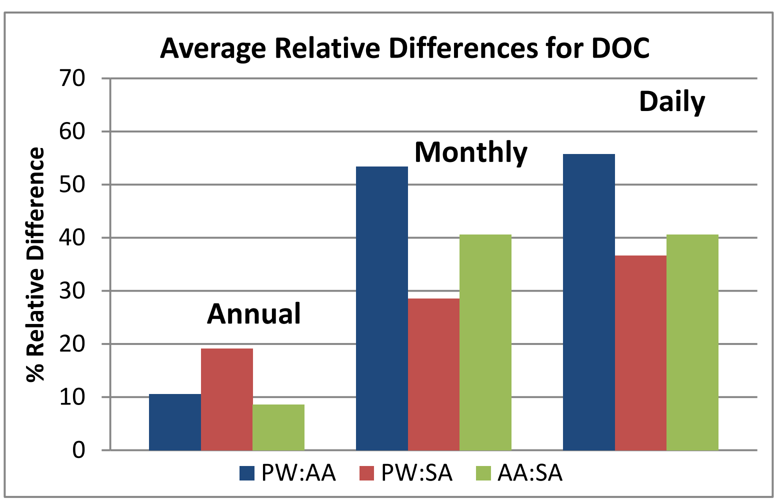

Figure 17: Average relative difference for DOC flux estimates for study sites.

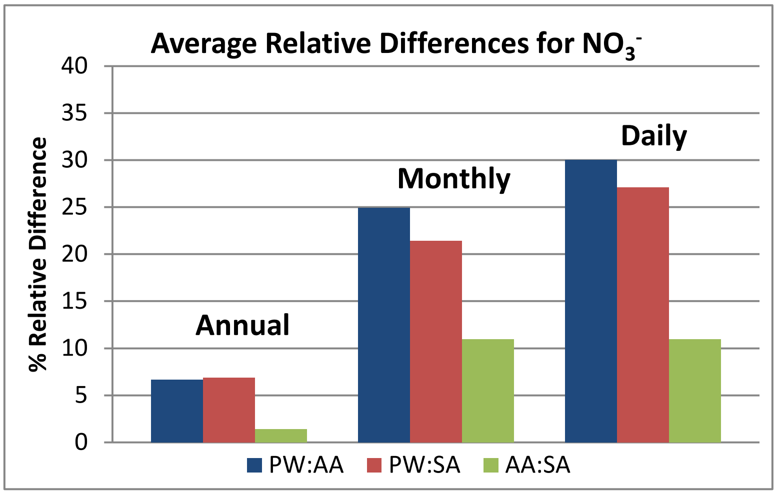

Figure 18: Average relative difference for NO_3^- flux estimates for study sites.

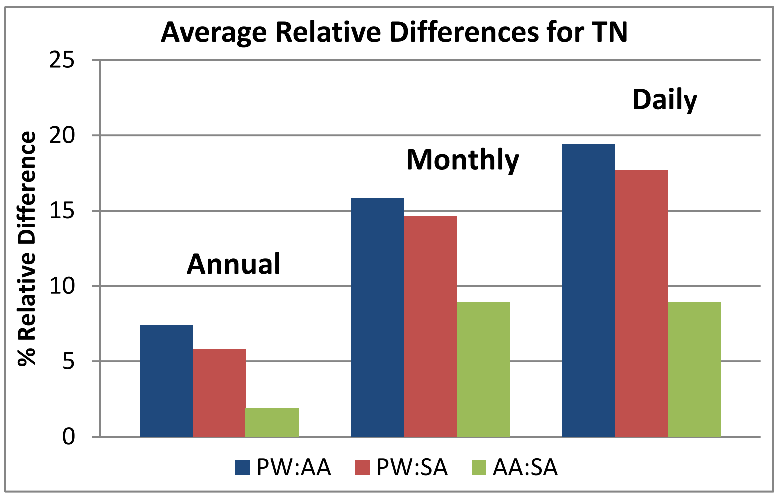

Figure 19: Average relative difference for TN flux estimates for study sites.

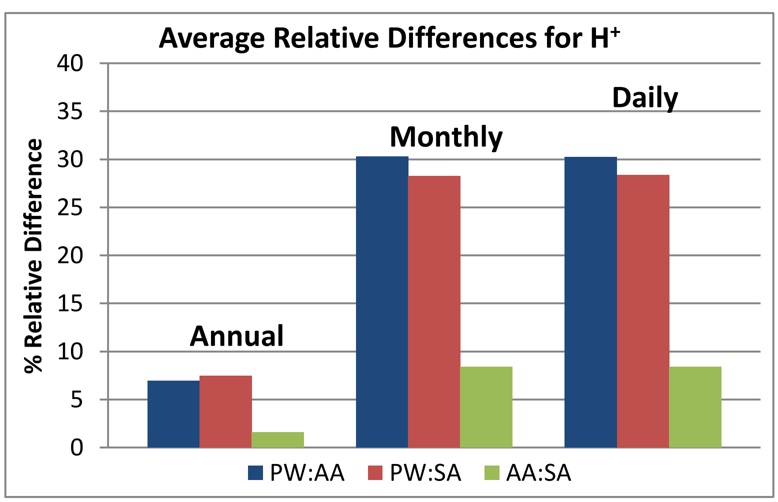

Figure 20: Average relative difference for H^+ flux estimates for study sites.

For all four solutes the average relative differences were smallest when making annual flux estimates. Monthly flux estimates always differed more than the annual estimates, and daily estimates usually differed the most, but the percent relative difference for the monthly and daily flux estimates were generally similar. Therefore, hypothesis 3 is also found to be reasonable.

Chapter 4: Discussion

4.1 Estimation of Flow in Ungaged Basins

Estimating flow in ungaged catchments has been a very active area of research in hydrology over the past decade. With the Predictions in Ungaged Basins (PUB) initiative set forth by the International Association of Hydrological Sciences, many researchers have wrestled with the task of predicting flood statistics, real-time flow, and groundwater characterization in ungaged basins (Wagener, 2006). Some of the models that have emerged recently have been quite complex, utilizing several basin characteristics and meteorological data that requires detailed information on both the basin and the climate.

The results of this study however, suggest that the area ratio of the ungaged to gaged watersheds alone may be adequate basin information for estimating average daily flow based on a reference gage, at least in the Catskill region. This is demonstrated by the results for Hypothesis 1, which show a statistically significant relationship with a high coefficient of determination (R^2 = 0.93) between area ratio and functional relation slope for the model development pairs (Figure 6). The fact that area ratio alone accounts for 93% of the observed variance in the functional relation slopes shows that area ratio is indeed the most important watershed parameter for predicting daily flow based on a reference gage in the Catskills. Other recent studies, which will be explored below, provide conflicting evidence regarding the generality of this result.

In a 2006 study by Mohamoud and Parmar, the authors found that non-linear regional regression equations based on drainage area alone could predict mean annual stream flow with coefficients of determination between 0.95 and 0.98. Their study considered 75 gaged watersheds in the Mid-Atlantic region, and while it demonstrates how important drainage area is in regulating the annual flow regime (Mohamoud and Parmar, 2006), it also alludes to its potential predictive power as an explanatory variable at shorter time scales. In a more recent study by Mohamoud (2008), attempts were made to predict daily stream flow in the Appalachian region by sequencing constructed flow duration curves (FDCs) with stream flow at a gaged reference site. In this study Mohamoud compared flow values predicted from his FDC method and flow values predicted from various forms of the area ratio method to the actual flow values in the study streams. Mohamoud's model utilized multiple regression to identify explanatory basin and climate characteristics from 26 catchments to develop region-specific flow duration curve construction models. Although each point of the flow duration curves was generated using only two explanatory variables, the total number of variables used to construct all of the points on the curves exceeded 20 basin and climate characteristics. These characteristics included land use, geomorphology, soil, geology, and climate characteristics which required the use of geographic information systems (GIS), digital elevation models (DEM), soil survey information, and detailed climate records.

After the development of such a complex model, requiring significant input data, over 20 explanatory variables, and a reference stream gage for stream flow sequencing, the model produced results comparable to those of the area ratio method for the prediction of daily stream flow in the three test watersheds (Mohamoud, 2008). Furthermore, both the predictions made by the FDC method and the area ratio method generally agreed well with the observed flows in the test streams. Although Mohamoud's FDC method does indeed produce good predictions of daily stream flow, I do not believe that the predictions are significantly better than the predictions made from the area ratio method, as evident in his results. This is evidence that despite significant added complexity, some models do not really perform any better than the area ratio method, and that the area ratio alone may provide adequate predictions of daily stream flow in certain regions.

Examples in which flow duration curves significantly outperformed area ratio methods include a 2009 study by the Ohio EPA, which examined TMDLs for the White Oak Creek watershed. The researchers in this study concluded that the area ratio method was inadequate for predicting real time flows. This was based on the fact that predicted flows differed from actual flows by an average of 262% and 64% in the two test watersheds using the area ratio method, while the flow duration curve method they used showed an average error in predicted versus actual flows of 113% and 35% for the same test watersheds (Ohio EPA, 2009).

This study, however, is fraught with inconsistency. The data used to develop the relative difference in flow estimates for the area ratio method were based on the difference in observed and predicted flow from 10 and 12 instantaneous flow measurements for the two test watersheds, while the flow duration curve method utilized a model containing over 50 years of stream data from 10 watersheds in Illinois that has the ability to account for man-made flow-altering devices, such as water withdrawals, which were present in one of the test watersheds. The researchers acknowledge that the area ratio method is inappropriate for such catchments, but compare and display the results equally anyway. The FDC model also requires an estimate of mean annual flow at the ungaged location, which is based on the drainage area, annual precipitation, and annual potential evapotranspiration of the ungaged catchment. Furthermore, the data for the area ratio analysis were collected during summer low flow, which the authors acknowledge as being their worst season for predicting flows based on the area ratio method, as evidenced by the predicted flows during this period having the highest percent error with respect to the observed flows.

Although the model used by the researchers does outperform the area ratio method, it is based on a small sample size (2 catchments, with 10 or 12 measurements in each catchment), with data collected during the worst predictive season (summer), and compares their modeled results to a method that was used in a location inappropriate for its application (the area ratio method should not be used in a catchment with flow altering devices). These shortcomings call into question their conclusions that flow duration curves are significantly better at predicting daily flow in ungaged catchments than methods based on area ratios. One thing the White Oak Creek study did have in common with this study, however, was an inability to produce good flow estimates at very high and low flow percentiles using the area ratio method. This limitation was also demonstrated by their flow duration curve method, perhaps providing evidence that predictions at very high and low flows may be the greatest obstacle to predicting flow in ungaged watersheds.

Prediction of Extreme Flows

Indeed, extreme event prediction proved to be difficult in this study as well. This is evident by the disagreement between the functional relation line and the observed flow data points at the greater than 98th percentile flows. An example of this is demonstrated by model development pair 11, Woodland Creek (Figure 21).

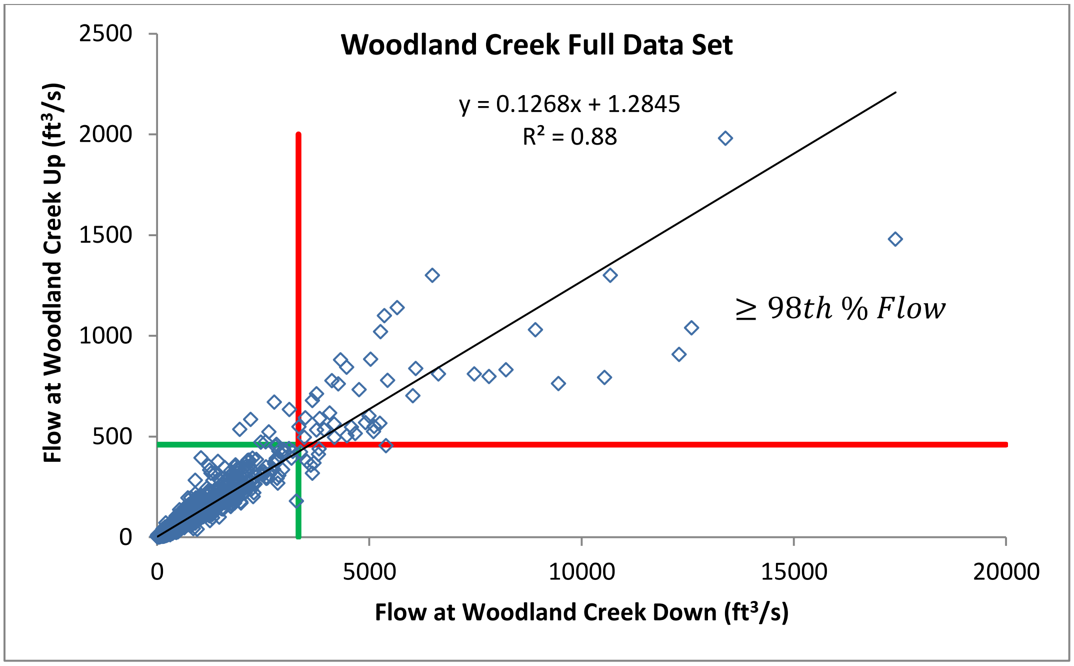

Figure 21: Woodland Creek pair showing the increased scatter of data points in the greater than 98th % flow region denoted by the area to the right and above the red bars.

The data markers in Figure 21 are not shaded in an attempt to portray the density of data points in certain regions of the chart and the line shown is the functional relation. The green bars indicate the area of the chart where flow is below the 98th percentile for both the upstream and downstream gage. The region to the right and above the red bars indicates flow values that exceed the 98th percentile for both the upstream and downstream gages. Note the increased scatter of data points about the functional relation line at flow values greater than the 98th flow percentile. This pattern was observed in most of the model development pairs, but tended to be more dramatic in pairs with low area ratio (<0.1). Although there are relatively few data points in this region, because they represent the highest flow days of the year, they may also represent the highest chemical flux days of the year. Therefore, poor prediction of flow during extreme events may subsequently lead to poor prediction of chemical mass flux. Even on an annual basis, this could drastically affect the quality of the flux estimates, as these few very high flow days could potentially represent the majority of the annual flux (Hinton et al., 1997; Eimers et al., 2008). The implications of this will be discussed further in the chemical flux discussion that follows this section.

Predictions Using Rainfall-Runoff Models

Rainfall-runoff models have also received a fair amount of attention since the PUB initiative emerged. These models have had varying degrees of success for predicting daily stream flow (Post and Jakeman, 1999). In a study by Post and Jakeman (1999), daily stream flow was estimated by regionalizing the parameters of a lumped conceptual rainfall-runoff model. Their model contained only 6 parameters, including time constants related to the unit hydrograph, a temperature variable, and an effective rainfall variable. A unit hydrograph is a flow time series meant to represent the response of a given watershed to a given unit of effective rainfall. Effective rainfall is the amount of precipitation that reaches the stream as runoff and is converted to stream flow. The parameters used by Post and Jakeman (1999) were related to 20 basin characteristics to determine which were important explanatory variables. They found drainage area, basin elongation, channel gradient, basin slope, drainage density, and wetted area to adequately represent their model parameters. This type of model requires the previously stated basin information, daily climate data, and the development of a unit hydrograph for the ungaged location, based on a relationship between drainage area and the unit hydrograph peak determined by the model calibration sites (Post and Jakeman, 1999).

Post and Jakeman (1999) applied their model to estimate daily stream flow in 16 catchments in Australia, with varying degrees of success. In some catchments the coefficient of determination for the actual to predicted flows was as high as 0.72, while in others it was as low as 0.07 (Post and Jakeman, 1999). Their results demonstrate another study in which a fairly complex model, requiring substantial basin and climate data, still does not produce consistently good estimates of daily stream flow. For the Catskill region, where area ratio alone can explain 93% of the variation in daily flow (Figure 6), it is unlikely that more complex approaches, such as those suggested by Post and Jakeman (1999) will greatly improve predictions of flow.

Warranted Complexity and the Value of Data Set Partitioning

A theme throughout this discussion has been how complex models often fail to consistently produce good predictions of daily stream flow. This was experienced in the development of the flow prediction model for this study as well. The flow data sets were originally partitioned into five flow regimes, comprised of the three flow regimes used in the study, plus a flow classification for very low flow (<5% flow) and very high flow (>95% flow). The relationships developed for these extreme flow regimes were generally quite poor, with little predictive potential. On the low end the data points were highly scattered, with little apparent trend, while at the high end, the very highest data points seemed to be driving the regression line and producing unrealistically high coefficients of determination. Poor agreement between the actual and predicted flows at the very low flows was likely also affected by the magnitude and sign of the functional relation intercepts. This is why it was decided to omit these data from the relationships developed for the flow regimes used in the study, and extend those relationships to the extreme ends of the data sets when it came to flow prediction.

Improvements in high-flow predictions may be possible if climate data were incorporated into the high flow regime relationship of the current model. Although this would greatly increase the complexity of the model, it could potentially improve the flow estimates in the highest flow percentiles (>95%). This suggestion is based on the assumption that the very high flow days, especially at upstream gage locations, may be best explained by meteorological data rather than basin data. Consider a small gaged headwater catchment nested within a larger gaged reference catchment and the scenario of convective or orographic rain events. These events can be relatively significant in their rainfall intensity but small in their geographic range. If such an event were positioned so that it caused significant precipitation over the small headwater catchment and relatively little precipitation over the rest of the larger reference catchment it is conceivable that the models presented in this study would under-predict flow at the upstream gage. Incorporation of historic meteorological data could provide justification for the removal of such uncharacteristic data points from the model development phase in the generation of general models, such as in this study, or for the addition of a supplementary flow adjustment coefficient based on storm intensity and duration to better transform flow from the reference gage during events that meet this description.

Even excluding the very highest and lowest flow regimes, it was not clear that data set partitioning and individual flow regime relationship development was beneficial. This is evident in the results for Hypothesis 2, where only 5 of the 15 model development pairs showed improved flow prediction following partitioning and separate relationship development. This is consistent with the modeling philosophy that simpler is often better. However, this analysis suggested that there is indeed an area ratio threshold where data set partitioning and separate flow regime relationship development does generally improve the predictive power of the relationships. The results for Hypothesis 2 show that for nested gage pairs with area ratios greater than 0.1, data set partitioning and separate flow regime relationship development generally improves the predictive power of flow estimates. This was not the case for pairs with area ratios less than 0.1, as none of those pairs benefitted from partitioning. This interesting result was unexpected. It was assumed at the start of the study that using relationships developed for each flow regime would greatly improve the estimates of flow for all model development pairs, regardless of area ratio.

Hortness (2006) suggested that the standard area ratio method only be used when the area ratio between the ungaged and reference watersheds is between 0.5 and 1.5. Others have both extended and restricted this range: Koltun and Shwartz (1987) suggested a limited range of 0.85 to 1.15, while Ries and Friesz (2000) showed that the area ratio method can be used with area ratios as low as 0.3 for low flow estimates. Interestingly, however, of the studies I examined, only Ries and Friesz (2000) provide any scientific evidence for their suggested range. The other studies simply provided area ratio method range guidelines without any justification beyond it being common practice that the range be restricted. Given the lack of evidence supporting the conventions provided by Hortness (2006) and Koltun and Shwartz (1987), neither were followed in this study. The results of Hypothesis 1 and 2, especially Figure 13, which shows the increased correlation of flow data sets at area ratios greater than 0.1, show that perhaps area ratio methods could be used to transform flow at lower area ratios than previously thought, especially if regional relationships comparing area ratio to functional relations and flow adjustment coefficients are developed, as in this study.

Area-ratio-based methods were successful in the Catskills for several reasons. These reasons are based primarily on the fact that the region is relatively small and is likely hydrologically homogenous. This is largely because the soil, climate, topography, and basin characteristics are similar throughout the region, creating a fairly predictable hydrologic response. The soils in the Catskills are almost entirely inceptisols (sandy loams), with the presence of fragipans a common occurrence (Kudish, 1979). The soil texture and the presence of fragipans, along with the generally shallow soil depths found at higher elevations, can act to decrease infiltration rates and promote surface ponding and runoff in upland catchments. Despite the plot-scale heterogeneities in soil texture, they are likely fairly homogenous on the catchment scale (sandy loams, frequently underlain by pans), leading to generally flashy hydrologic response across the region. This is in line with what McDaniel et al. (2008) concluded in a study regarding flashy upland watersheds in Idaho containing fragipans. These researchers determined that shallow soil depths underlain by fragipans were responsible for the flashy hydrologic response observed in their study sites.

Since the region is relatively small (1,120 mi^2), it is not surprising that the climate would be similar across the region. This leads to similar weather patterns, erosional settings, and soil development conditions for the entire region. Basin characteristics, including topography, stream channel characteristics, watershed storage, and land use conditions are also similar across the entire region and likely act to influence the hydrologic response. Since the Catskills are not true mountains (in an orographic sense), the topography is best explained by alluvial and glacial erosion rather than mountain building processes. This has led to relatively similar channel slopes and stream channel characteristics across the region. Stream channels in the Catskills tend to be relatively straight, steep and well defined, therefore decreasing travel time in the channel and increasing the likelihood of correlation between upstream and downstream stream flows.

The almost uniform lack of storage features in Catskill watersheds also influences hydrologic response. Catskill watersheds rarely contain lakes, wetlands, or other natural water storage features that would act to slow the hydrologic response, therefore contributing to the flashy response that characterizes the region. Land use conditions also play an important role in regulating the hydrologic response and are one of the key reasons that area ratio methods are successful in the Catskills. Both the model development sites and study sites used in this study are largely forested and lack urban areas. This helps to increase the hydrologic homogeneity of the sites, and is representative of the region as a whole, which is generally forested and lacks urban centers.

These factors combine to control the relatively simple hydrologic response observed in the Catskill region, where watershed area ratio alone can be used to describe and predict flow at the upstream location of a nested pair of stream gages. This may not be true in hydrologically more complex systems with longer, more sinuous stream channels or regions with significant groundwater contributions, but the findings of this study strongly support the use of area ratio methods for estimation of daily flow in the Catskills given the above characteristics of the region.

4.2 Chemical Flux Estimation in Ungaged Basins

The results from Hypothesis 3 show that for all of the three methods used to estimate chemical mass fluxes, the relative difference between flux estimates generally increases with decreasing time scale. This was the expected result.

Concentration-Discharge Relationships

It was surprising to find that none of the analytes considered in Hypothesis 3 related well to discharge. This is counter to some general geochemical trends and to what many other researchers have found. For example, it is generally accepted that the concentrations of base cations and other mineral weathering products tend to decrease with increasing flow (Walling and Webb, 1986; Johnson et al, 1969). Other solutes may respond differently to changes in flow. Sullivan et al. (1986) and Johnson et al. (1969) showed how hydrogen ion concentrations tend to increase with flow in acidified environments, such as the Catskills, and Walling and Webb (1986) demonstrated how leaching of organic horizons can increase NO_3^- concentrations during periods of high flow. Semkin et al. (1994) also reference several studies that show a direct relationship between DOC and flow rate.

The concentration-discharge relationships in this study were developed from nearly two years of monthly samples. Perhaps more data points would have demonstrated a significant relationship between flow and concentration for some of the solutes of interest. Or perhaps the Catskills truly are chemostatic, as Godsey et al. (2009) suggest many catchments are, at least with respect to weathering products. In their study, Godsey et al. (2009) point out that discharge in their test watersheds varies by several orders of magnitude, while solute concentrations typically only vary by factors ranging from 3 to 20. Determination of the true mechanisms of solute concentration control in the Catskills are beyond the scope of this study, but some suggestions as to why the solutes of interest acted the way they did can be made.

Dissolved organic carbon concentrations in the grab samples ranged from roughly 0.5 to 8 mg/L, with the vast majority of samples varying between 0.5 and 3.5 mg/L. These are reasonable concentrations for temperate forested ecosystems (Mulholland, 2003) and I assumed that DOC and flow rate were not well correlated because of basin characteristics and flow path considerations. Mulholland (2003) lists several studies that connect in-stream DOC concentrations to the flow path of the water to the stream. He suggests that flow paths that pass through organic rich soil horizons (O horizons) enrich the water with DOC, while flow paths that pass through lower horizons with higher DOC sorbing capacities do not. Mulholland attributes stream DOC concentrations to the presence of wetlands, the flow path of the runoff through the soil, and the channel slope. The study sites in this study do not contain large areas of wetlands and the Catskills as a whole are characterized by soils that lack a well-developed organic horizon. Therefore, it seems reasonable that even if DOC concentrations are somewhat regulated by flow rate in some locations, given the lack of wetlands and organic soil horizons, volumetric flow rate alone is not an adequate predictor of DOC concentrations for streams located in the Catskills.

The behavior of nitrate and total nitrogen in the samples, and their apparent lack of relationship with flow, is probably best explained by biological interactions and seasonal fluctuations in the available nitrogen pool. Nitrate has been shown to increase with flow in some studies (Walling and Webb, 1986), and to be unrelated to flow in others (Anderson et al., 1997). Rusjan et al. (2008) explain how during certain times of the year, accumulation of labile nitrogen species (NO_3^-, NO_2^-, NH_3) may be flushed by large meteorological events, while at other times in the year when biological demands for nitrogen are highest, an increase in flow rate can dilute nitrogen concentrations in stream flow. This explanation places an emphasis on both biological and seasonal controls of nitrogen. Thus simple concentration-discharge relationships may not sufficiently explain nitrate and total nitrogen concentrations in the Catskills, and grab samples likely represent a reasonable picture of the true nitrogen dynamics in the study sites.

Sampling Frequency and the Role of Grab Samples

Since no measured flux values were available for comparison in the study sites, the confidence level of the flux estimates made in this study is difficult to quantify. However, some qualitative statements regarding the quality of the method and its appropriateness for estimating fluxes at different time scales can be made. Firstly, one must acknowledge the limitations of flux estimates made using monthly grab samples, especially at finer temporal scales. The sampling frequency of this study however, is not unrealistic for agencies or municipalities with limited funding to spend on monitoring projects. Clearly, increasing the sampling frequency would increase confidence in the chemical mass flux estimates. However, for annual flux estimates, increased frequency may not necessarily greatly improve the quality of the estimates. This is manifest by the fact that these estimates were made in ungaged catchments, where flow is also estimated rather than measured. Therefore, if flux estimates are to be made in ungaged catchments using a relatively simple flow estimation scheme, monthly stream sampling may represent an adequate sampling frequency.

At finer temporal scales, a more rigorous sampling scheme may be necessary to accurately characterize the chemical concentrations at the study sties. This can be achieved through the use of autosamplers and composite samples. Autosamplers are devices placed at the study site that automatically take a stream sample based on temporal or flow weighted sampling cues, and composite samples are samples that contain several sub-samples taken at different times throughout the day. These techniques have the advantage of capturing the daily variation in solute concentrations which, as some recent studies suggest, can be significant.

Rusjan and Mikos (2010) used a 15-min sampling frequency to show how nitrate concentrations vary daily in a forested stream. They found that not only are diurnal fluctuations great, but that they vary considerably in their magnitude and timing over the course of the year as well. This has important implications for studies using grab samples to characterize solute concentrations, as the time of day and time of year could significantly change the concentration of particular solutes in a grab sample.

Another benefit of automated sampling techniques is that they can be programmed to respond to storm events and increase the sampling frequency during a storm to better characterize the solute concentrations during the event. This is important because a few major storm events could be responsible for the majority of the annual flux in certain cases (Grayson and Holden, 2011). Clark et al. (2007) list several studies that show that the majority of the DOC transported from some watersheds occurs in relatively few major storm events. These studies include Hinton et al. (1997), who showed that storms account for between 29% and 68% of the seasonal DOC flux depending on the season, and that a single large storm in one of his test watersheds accounted for 31% of the fall seasonal flux. This is consistent with what Eimers et al. found in a 2008 study which showed a single storm event being responsible for 66% of the annual DOC flux in a watershed in Ontario, Canada. These finding reinforce the importance of accurate flow estimates, especially of the very high flow values, as these few extreme days could represent the majority of the annual flux.

Although my data do not show any clear relationship between flow and DOC for streams in the Catskills, it is conceivable that even without this relationship the highest flow days are also likely to be the highest chemical mass flux days, simply due to the magnitude of flow. This becomes increasingly important in catchments that discharge to water supply reservoirs, such as the catchments used in this study, because DOC can have important implications for water treatment and disinfection using chlorination, specifically by potentially increasing the presence of disinfection byproducts like trihalomethanes and haloacetic acids in the drinking water supply.

I believe the sampling scheme (grab samples) and flow estimation methods used in this study are appropriate for the Catskills because the solutes I considered are not well correlated to flow and the sampling frequency is adequate for the goal of estimating mass fluxes on an annual basis. Furthermore, the flow estimation methods are appropriate because flow in ungaged catchments in the Catskills can be explained very well by just the area ratio of the ungaged to reference watersheds.

4.3 Study Limitations

This study utilizes data and information specific to the Catskill Park in New York State. Therefore, application of these approaches outside of the Catskill region would require validation. Furthermore, the method used to estimate fluxes for Hypothesis 3 is only recommended for sample sites that have area ratios with a reference gage greater than 0.1. This is because only nested pairs with area ratios greater than 0.1 benefitted from flow partitioning and the adjustment coefficients used in the flow estimates for Hypothesis 3 were based on the relationships developed from the partitioned data sets.

Limitations in the design of this study include the use of grab samples for obtaining concentration data for mass fluxes made at fine temporal scales. Potential limitations of the grab sample technique are that the sample taken is not representative of the actual in-stream conditions for the period which it is expected to represent. This can happen with biologically and seasonally mediated solutes, in which temperature or meteorological anomalies could produce unusually high or low solute concentrations that happened to coincide with the sampling times. These potentially misleading data points could then lead to over- or under-estimation of fluxes for the period in which they were taken. The effect of this should decrease with increasing sampling frequency.

Another potential limitation of this study was the limited success in estimating very high and very low flow values (>95% and <5%). Although the implications of poor estimates of low flow aren’t very important with regard to mass flux estimates, poor estimates of high flow are. They can lead to inaccurate estimates of flux, with the potential of providing misleading images of ecosystem health, solute processing, and geochemical relationships.

Chapter 5: Conclusions

After careful consideration of the results of this study, some important conclusions can be drawn. First, based on the results of Hypothesis 1, it is clear that the area ratio is indeed the most important basin characteristic for estimating flow at the upstream location of a nested pair of stream gages based on a reference gage in the Catskills. This is supported by Figure 6, which shows that area ratio explains 93% of the variance in the functional relation slopes for the model development pairs. Furthermore, based on the examples cited in the flow estimation discussion section, I believe that area ratio is the only basin parameter required to make reasonable estimates of flow in ungaged catchments for the purpose of estimating annual fluxes based on monthly grab samples. These examples demonstrate how the use of very complex models still does not produce consistently better estimates of daily average flow than methods based on area ratios, and that the added complexity is, in many cases, unwarranted, especially when considering the goal of estimating fluxes at coarse temporal scales.

Also, as evidenced by the results from Hypothesis 2, if the area ratio of the upstream to downstream gaged catchments is greater than 0.1, partitioning the reference gage flow data set and developing separate low, medium, and high flow regime relationships is beneficial for the estimation of flow at the upstream location of the nested pair. My results do not show the same for pairs with area ratios less than 0.1, so I therefore suggest that the area ratio method used in this study could be extended for predicting flow in ungaged basins to pairs with area ratios as low as 0.1, especially if separate flow regime relationships are developed.

Finally, for solutes that don't relate well to flow, but do show high temporal variability, the relative difference among the different chemistry aggregation methods examined in this study is relatively small on an annual basis (generally < 10%), with differences increasing at shorter time steps. This leads to the conclusion that for annual flux estimates, the choice of chemistry aggregation method is not critical to the outcome. Any of the three methods suggested in this study could be used and the results would be similar. Rather than spending time and energy choosing a chemistry aggregation method, efforts should be made to maximize the confidence in flow estimates and grab sample quality, which will have a far greater influence on the quality of the flux estimates than the specific method used to aggregate the chemistry data.

Appendix A: Functional Relation Equations for Model Development Pairs

The following table summarizes the functional relation information for the fifteen model development pairs. These data were generated from the full data sets of the model development pairs.

Table 7: Functional relation slopes, intercepts, and R^2 for the model development pairs.

			Functional Relation		
Pair	Name	Area Ratio	Slope	Intercept	R^2
1	Batavia Kill	0.0296	0.0444	-0.0266	0.6788
2	Biscuit Brook	0.1101	0.0978	-0.0578	0.8771
3	Bush Kill	0.5681	0.4937	4.2344	0.9615
4	Hollow Tree	0.0631	0.0465	1.3002	0.7990
5	Rondout Creek	0.1399	0.1709	-0.7907	0.8831
6	Winnisook	0.0228	0.0303	-0.6825	0.7335
7	West Kill	0.1841	0.2136	-0.0664	0.9040
8	East Kill	0.3678	0.3191	1.6284	0.9661
9	Beaver Kill Trib.	0.0051	0.0049	-0.3599	0.8151
10	Little Beaver Kill	0.0859	0.0825	-0.9163	0.8243
11	Woodland Creek	0.1073	0.1268	1.2845	0.8782
12	E. Branch Neversink	0.3900	0.4993	-3.7639	0.9430
13	Esopus	0.0242	0.0306	-1.3443	0.5300
14	High Falls	0.0811	0.0496	2.0685	0.8726
15	Birch Creek	0.1962	0.1417	4.4022	0.9062

Appendix B: Concentration-Discharge Curves and Tables for Study Sites

The following figures show the concentration-discharge curves for each solute and study site.

Dissolved Organic Carbon:

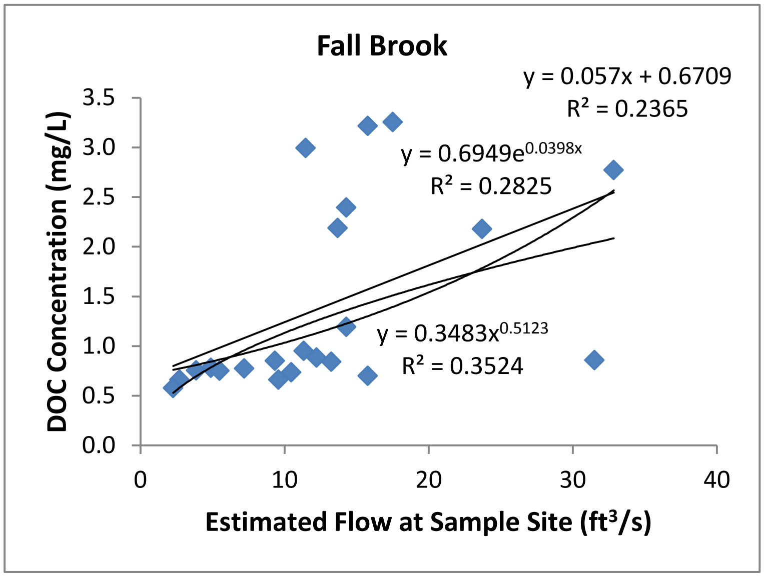

Figure 22: Q-C Curve for DOC at Fall Brook.

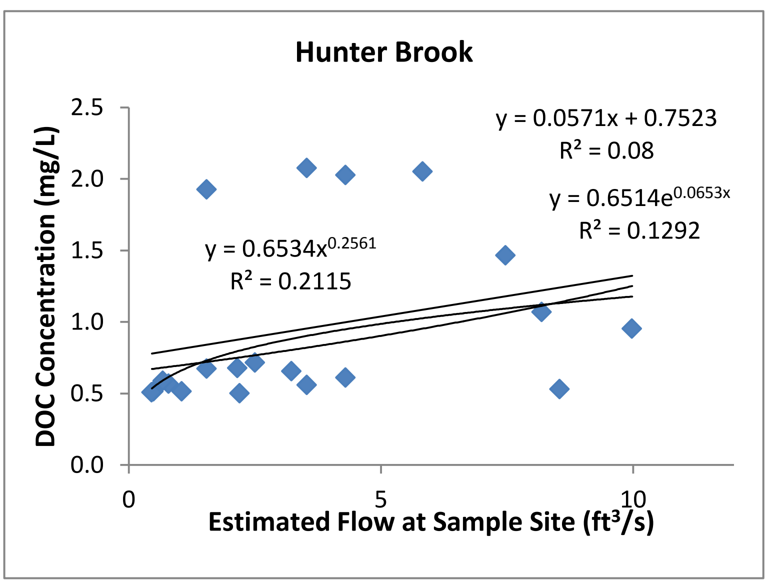

Figure 23: Q-C Curve for DOC at Hunter Brook.

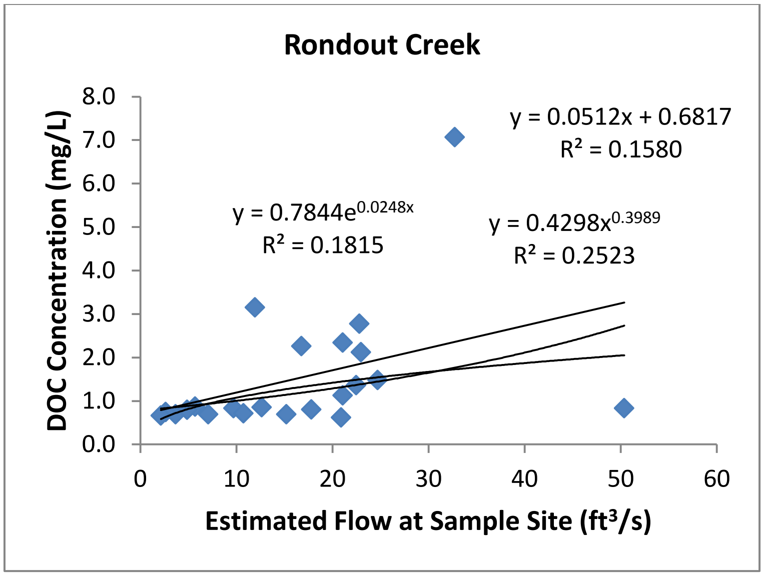

Figure 24: Q-C Curve for DOC at Rondout Creek.

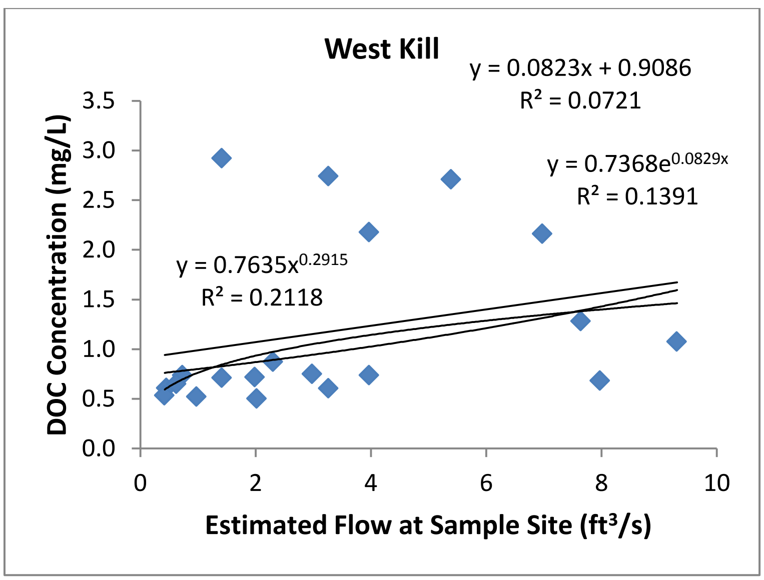

Figure 25: Q-C Curve for DOC at West Kill.

Nitrate:

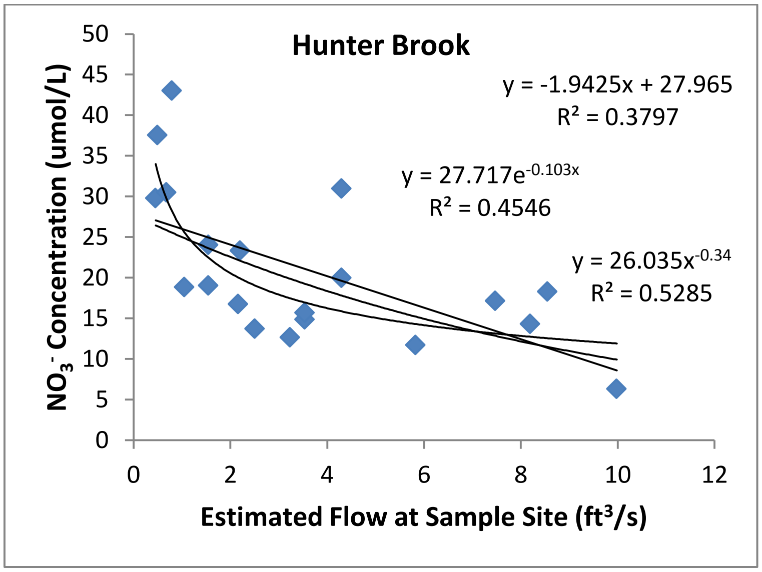

Figure 26: Q-C Curve for NO_3^- at Fall Brook.

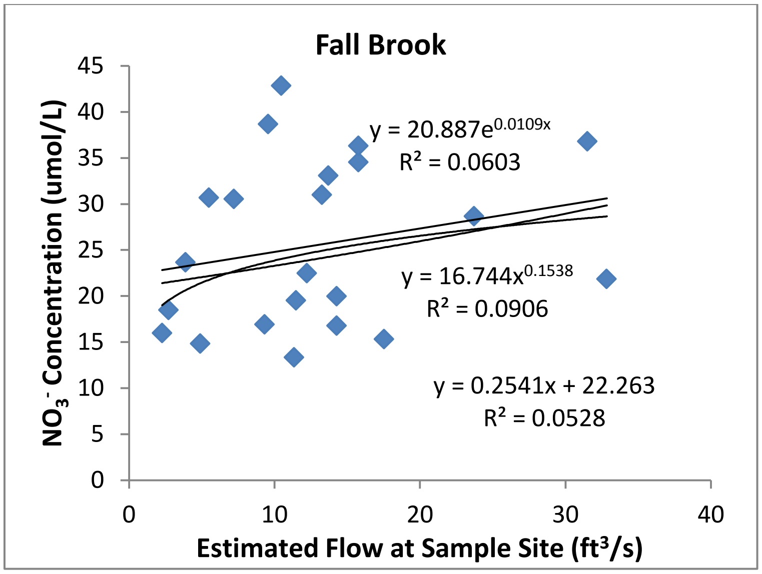

Figure 27: Q-C Curve for NO_3^- at Hunter Brook.

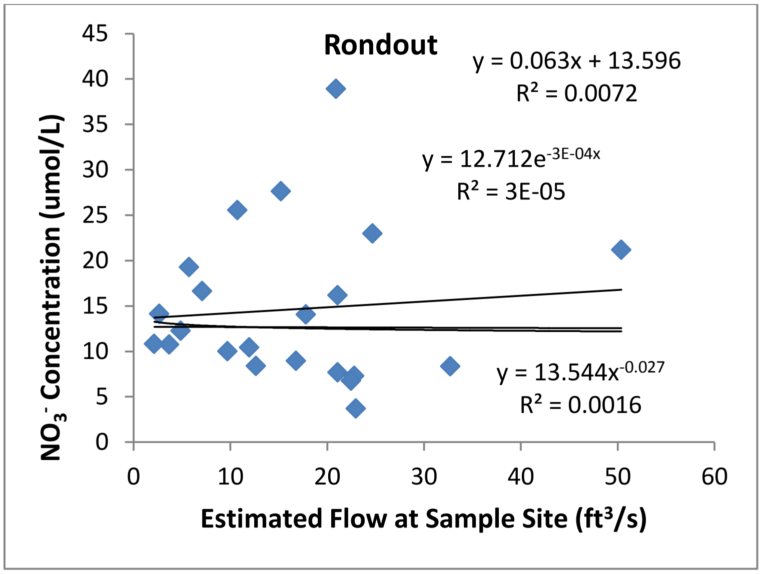

Figure 28: Q-C Curve for NO_3^- at Rondout Creek.

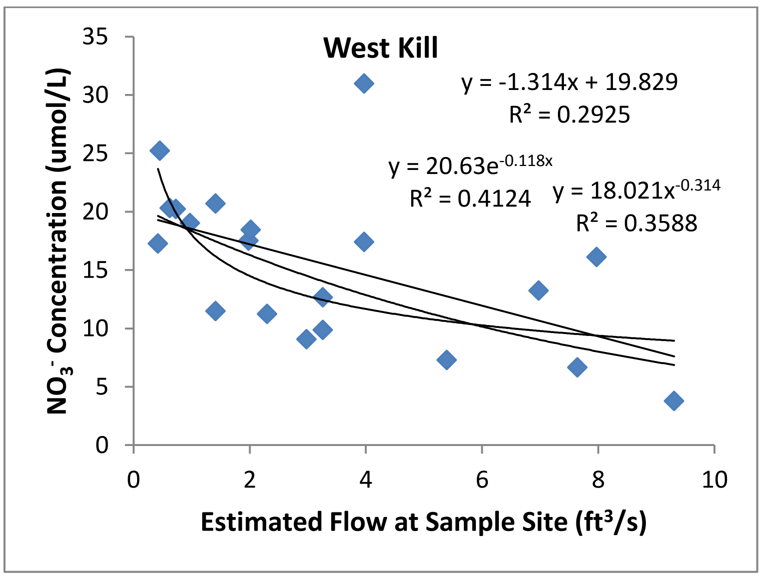

Figure 29: Q-C Curve for NO_3^- at West Kill.

Total Nitrogen:

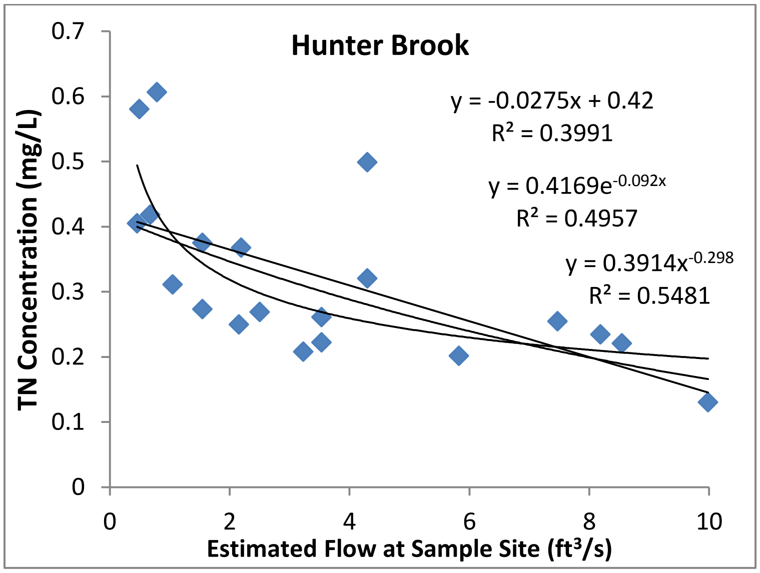

Figure 30: Q-C Curve for TN at Fall Brook.

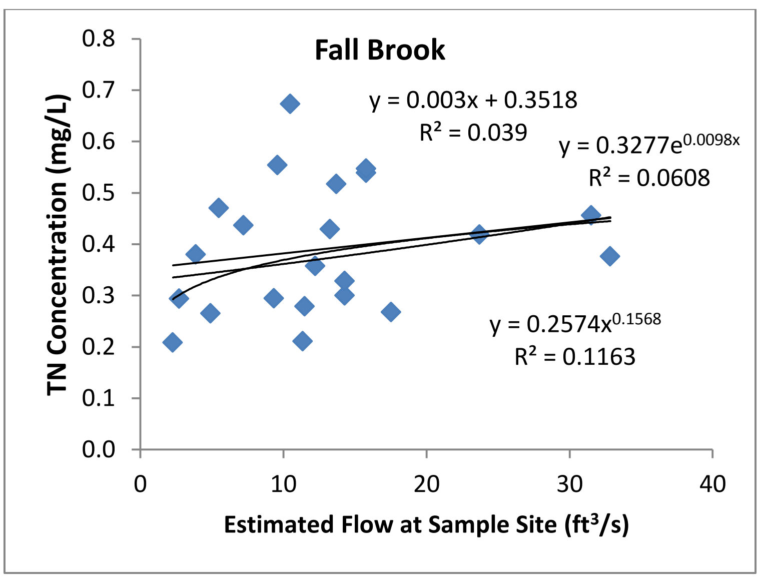

Figure 31: Q-C Curve for TN at Hunter Brook.

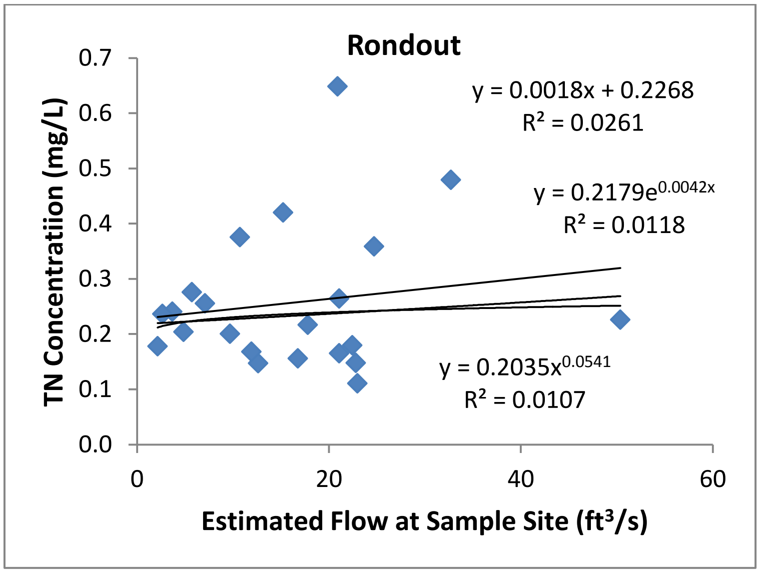

Figure 32: Q-C Curve for TN at Rondout Creek.

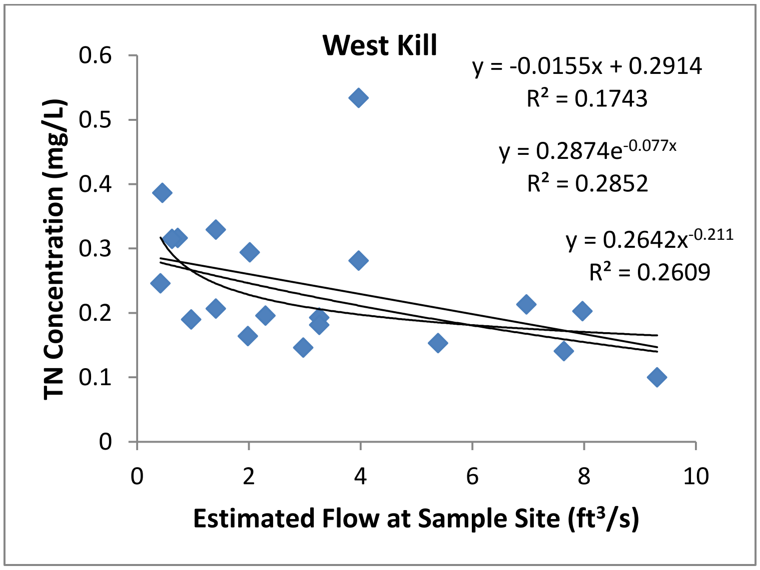

Figure 33: Q-C Curve for TN at West Kill.

Hydrogen Ion:

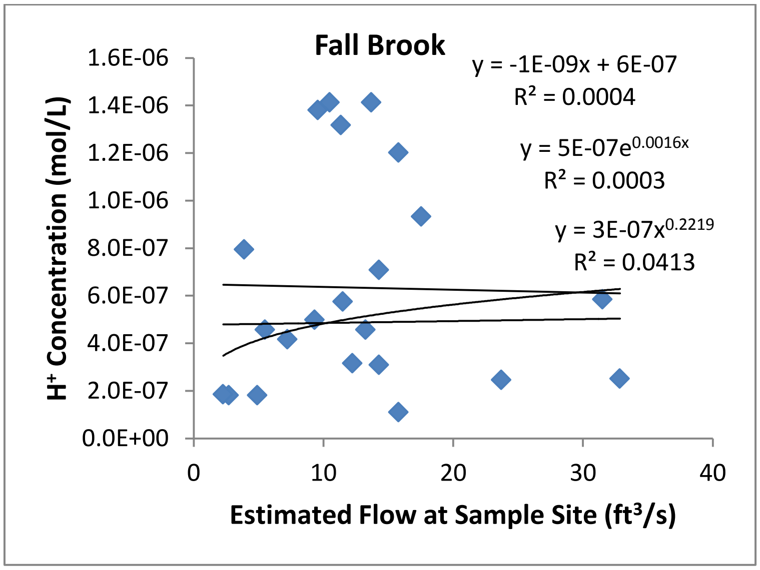

Figure 34: Q-C Curve for H^+ at Fall Brook.

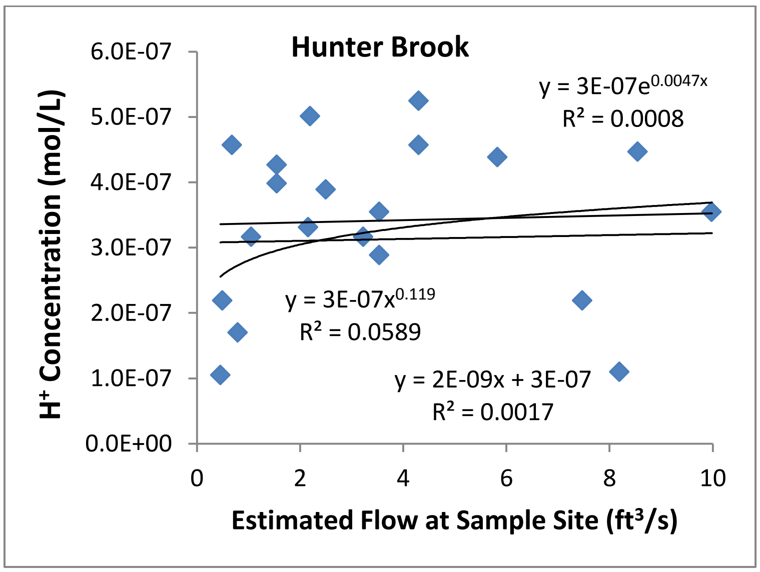

Figure 35: Q-C Curve for H^+ at Hunter Brook.

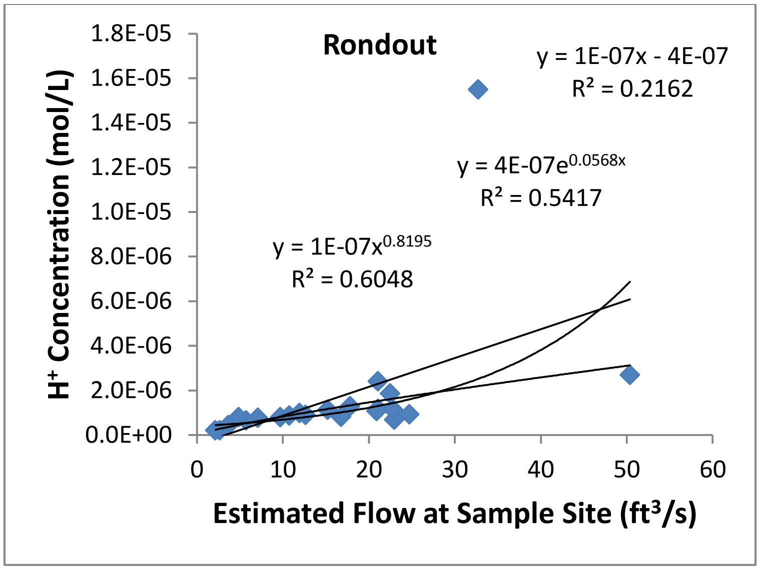

Figure 36: Q-C Curve for H^+ at Rondout Creek.

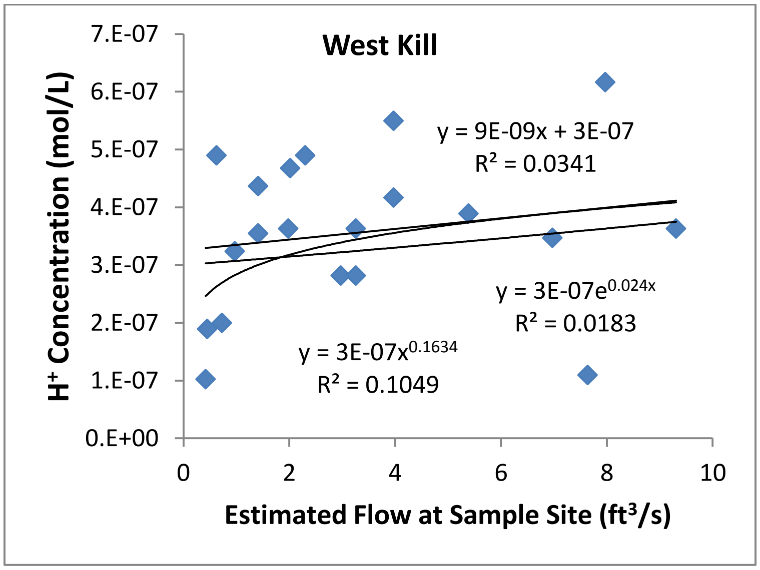

Figure 37: Q-C Curve for H^+ at West Kill.

The following tables show the flow values and solute concentrations for each solute and study site.

Fall Brook:

Table 8: Flow rates and analyte concentrations for samples taken at Fall Brook.

		Solute Concentration				
Date	Estimated Flow (ft^3/s)	DOC (mg/l)	NO_3^- (umol/l)	TN (mg/l)	H^+ (mol/l)	Duplicate Sample (Y/N)
6/29/2010	4.91	0.78	14.80	0.26	1.8E-07	Y
7/27/2010	2.72	0.66	18.46	0.29	1.8E-07	N
8/23/2010	15.79	3.21	34.53	0.54	1.1E-07	N
9/18/2010	2.28	0.58	15.96	0.21	1.9E-07	N
10/16/2010	32.85	2.77	21.84	0.38	2.5E-07	N
11/13/2010	12.23	0.88	22.47	0.36	3.2E-07	N
12/11/2010	13.26	0.84	30.98	0.43	4.6E-07	N
1/8/2011	7.22	0.77	30.51	0.44	4.2E-07	N
2/11/2011	5.49	0.75	30.68	0.47	4.6E-07	N
3/16/2011	31.52	0.86	36.79	0.46	5.9E-07	N
4/9/2011	23.72	2.18	28.64	0.42	2.5E-07	N
5/14/2011	14.29	2.39	19.96	0.30	3.1E-07	N
6/11/2011	11.49	2.99	19.50	0.28	5.8E-07	N
7/7/2011	17.53	3.25	15.31	0.27	9.3E-07	N
8/3/2011	3.89	0.75	23.63	0.38	7.9E-07	N
9/17/2011	14.29	1.19	16.77	0.33	7.1E-07	N
10/23/2011	11.35	0.95	13.32	0.21	1.3E-06	N
11/16/2011	9.34	0.85	16.90	0.29	5.0E-07	N
12/21/2011	13.70	2.19	33.05	0.52	1.4E-06	N
1/21/2012	15.79	0.70	36.28	0.55	1.2E-06	N
2/11/2012	9.58	0.66	38.65	0.55	1.4E-06	N
3/12/2012	10.48	0.74	42.83	0.67	1.4E-06	N

Hunter Brook:

Table 9: Flow rates and analyte concentrations for samples taken at Hunter Brook.

Date	Estimated Flow (ft^3/s)	Solute Concentration DOC (mg/l)	NO$_3^-$ (umol/l)	TN (mg/l)	H$^+$ (mol/l)	Duplicate Sample (Y/N)
6/29/2010	0.49	0.51	37.55	0.58	2.2E-07	N
7/27/2010	0.79	0.57	43.04	0.61	1.7E-07	N
9/18/2010	0.46	0.51	29.82	0.41	1.0E-07	N
10/17/2010	8.19	1.07	14.34	0.23	1.1E-07	N
11/14/2010	3.23	0.66	12.68	0.21	3.2E-07	N
12/11/2010	3.54	0.56	14.86	0.22	3.5E-07	N
1/9/2011	2.16	0.68	16.76	0.25	3.3E-07	N
2/12/2011	1.05	0.51	18.84	0.31	3.2E-07	N
3/17/2011	8.55	0.53	18.30	0.22	4.5E-07	N
4/10/2011	7.47	1.47	17.17	0.25	2.2E-07	Y
5/15/2011	3.54	2.07	15.70	0.26	2.9E-07	N
6/11/2011	1.55	1.93	19.06	0.27	4.3E-07	N
7/7/2011	5.83	2.05	11.72	0.20	4.4E-07	Y
8/4/2011	0.68	0.59	30.49	0.42	4.6E-07	N
10/22/2011	9.98	0.95	6.34	0.13	3.5E-07	N
11/17/2011	2.50	0.72	13.72	0.27	3.9E-07	N
12/22/2011	4.30	2.03	20.00	0.32	4.6E-07	N
1/22/2012	1.55	0.67	24.05	0.37	4.0E-07	N
2/10/2011	2.20	0.50	23.34	0.37	5.0E-07	N
3/13/2012	4.30	0.61	30.97	0.50	5.2E-07	N

Rondout Creek:

Table 10:Flow rates and analyte concentrations for samples taken at Rondout Creek.

		Solute Concentration				
Date	Estimated Flow (ft^3/s)	DOC (mg/l)	NO_3^- (umol/l)	TN (mg/l)	H^+ (mol/l)	Duplicate Sample (Y/N)
6/29/2010	3.69	0.69	10.74	0.24	4.37E-07	N
7/27/2010	2.66	0.74	14.14	0.24	2.04E-07	N
8/23/2010	32.75	7.06	8.35	0.48	1.55E-05	N
9/18/2010	2.15	0.66	10.83	0.18	2.14E-07	N
10/16/2010	22.99	2.12	3.71	0.11	6.76E-07	N
11/13/2010	12.64	0.85	8.37	0.15	8.91E-07	N
12/11/2010	17.82	0.80	14.05	0.22	1.29E-06	N
1/8/2011	7.09	0.69	16.64	0.26	7.62E-07	N
2/11/2011	5.73	0.87	19.28	0.28	6.61E-07	N
3/16/2011	50.38	0.83	21.16	0.23	2.69E-06	N
4/9/2011	21.09	2.34	16.16	0.26	1.17E-06	N
5/14/2011	16.78	2.26	8.94	0.16	8.13E-07	N
6/11/2011	11.95	3.15	10.41	0.17	9.77E-07	N
7/7/2011	22.81	2.77	7.30	0.15	1.15E-06	N
8/3/2011	4.88	0.79	12.25	0.20	7.94E-07	N
9/17/2011	22.47	1.36	6.75	0.18	1.86E-06	N
10/23/2011	21.09	1.13	7.66	0.17	2.40E-06	N
11/16/2011	9.71	0.83	10.02	0.20	7.94E-07	N
12/21/2011	24.71	1.48	22.98	0.36	9.33E-07	N
1/21/2012	10.75	0.71	25.55	0.38	8.71E-07	N
2/11/2012	15.23	0.69	27.62	0.42	1.12E-06	N
3/12/2012	20.92	0.62	38.91	0.65	1.07E-06	N

West Kill:

Table 11: Flow rates and analyte concentrations for samples taken at West Kill.

Date	Estimated Flow (ft^3/s)	Solute Concentration DOC (mg/l)	NO_3^- (umol/l)	TN (mg/l)	H^+ (mol/l)	Duplicate Sample (Y/N)
6/29/2010	0.45	0.61	25.22	0.39	1.89E-07	Y
7/27/2010	0.73	0.74	20.24	0.32	2.00E-07	N
9/18/2010	0.42	0.53	17.26	0.25	1.02E-07	N
10/17/2010	7.64	1.28	6.66	0.14	1.10E-07	N
11/14/2010	2.98	0.75	9.06	0.15	2.82E-07	N
12/11/2010	3.26	0.61	12.65	0.19	3.63E-07	N
1/9/2011	1.98	0.72	17.50	0.16	3.63E-07	N
2/12/2011	0.97	0.52	19.01	0.19	3.24E-07	N
3/17/2011	7.97	0.68	16.12	0.20	6.17E-07	N
4/10/2011	6.97	2.16	13.25	0.21	3.47E-07	N
5/15/2011	3.26	2.74	9.86	0.18	2.82E-07	N
6/11/2011	1.42	2.92	11.47	0.21	4.37E-07	N
7/7/2011	5.39	2.71	7.29	0.15	3.89E-07	N
8/4/2011	0.63	0.65	20.32	0.32	4.90E-07	N
10/22/2011	9.31	1.08	3.79	0.10	3.63E-07	N
11/17/2011	2.30	0.87	11.23	0.20	4.90E-07	N
12/22/2011	3.97	2.18	17.40	0.28	4.17E-07	N
1/22/2012	1.42	0.71	20.68	0.33	3.55E-07	N
2/10/2011	2.02	0.50	18.45	0.29	4.68E-07	N
3/13/2012	3.97	0.74	30.97	0.53	5.50E-07	N

Appendix C: Relative Difference Between Chemistry Aggregation Methods

The following tables list the annual flux estimates calculated for each chemistry aggregation method, as well as the percent relative difference between the methods at an annual, monthly, and daily time scale.

Dissolved Organic Carbon:

Table 12: Annual flux estimates and percent relative difference for DOC at study sites.

	Annual Flux Estimate (kg/yr-ha)			% Relative Difference: Annual			Average Monthly			Average Daily		
Site	**PW**	**AA**	**SA**	**PW:AA**	**PW:SA**	**AA:SA**	**PW:AA**	**PW:SA**	**AA:SA**	**PW:AA**	**PW:SA**	**AA:SA**
Fall Brook	29.22	30.90	31.87	5.59	8.67	3.08	49.12	33.64	30.68	56.81	42.53	30.67
Hunter Brook	16.69	19.17	21.74	13.82	26.29	12.58	53.48	25.67	43.64	51.33	32.19	43.62
Rondout Creek	24.88	26.59	27.70	6.65	10.74	4.10	46.48	26.37	35.63	52.14	35.39	35.61
West Kill	21.21	24.98	28.98	16.33	30.96	14.83	64.47	28.61	52.60	62.75	36.33	52.57
			Average	10.60	19.16	8.64	53.39	28.57	40.64	55.76	36.61	40.62

Nitrate:

Table 13: Annual flux estimates and percent relative difference for NO_3^- at study sites.

	Annual Flux Estimate (mol/yr-ha)			% Relative Difference: Annual			Average Monthly			Average Daily		
Site	**PW**	**AA**	**SA**	**PW:AA**	**PW:SA**	**AA:SA**	**PW:AA**	**PW:SA**	**AA:SA**	**PW:AA**	**PW:SA**	**AA:SA**
Fall Brook	460.89	467.73	459.62	1.47	0.28	1.75	25.63	17.11	16.76	22.72	16.40	16.76
Hunter Brook	356.13	305.80	308.43	15.21	14.36	0.86	16.00	18.25	8.25	23.95	23.87	8.25
Rondout Creek	206.69	204.64	200.98	1.00	2.80	1.80	36.24	28.76	14.83	38.52	33.66	14.84
West Kill	248.41	227.23	224.35	8.90	10.18	1.28	21.94	21.49	4.10	34.82	34.51	4.10
			Average	6.65	6.90	1.42	24.95	21.40	10.99	30.00	27.11	10.99

Total Nitrogen:

Table 14: Annual flux estimates and percent relative difference for TN at study sites.

	Annual Flux Estimate (kg/yr-ha)			% Relative Difference								
				Annual			Average Monthly			Average Daily		
Site	PW	AA	SA	PW:AA	PW:SA	AA:SA	PW:AA	PW:SA	AA:SA	PW:AA	PW:SA	AA:SA
Fall Brook	6.91	7.08	6.99	2.46	1.16	1.29	17.52	11.88	12.36	15.31	11.22	12.36
Hunter Brook	5.13	4.52	4.63	12.56	10.16	2.40	14.16	15.28	7.81	19.59	18.88	7.81
Rondout Creek	3.33	3.32	3.30	0.29	0.87	0.58	20.26	18.12	4.79	22.60	21.00	4.79
West Kill	3.86	3.34	3.45	14.40	11.12	3.29	11.31	13.20	10.70	20.16	19.74	10.70
			Average	7.43	5.83	1.89	15.81	14.62	8.91	19.42	17.71	8.91

Hydrogen Ion:

Table 15: Annual flux estimates and percent relative difference for H^+ at study sites.

	Annual Flux Estimate (mol/yr-ha)			% Relative Difference								
				Annual			Average Monthly			Average Daily		
Site	PW	AA	SA	PW:AA	PW:SA	AA:SA	PW:AA	PW:SA	AA:SA	PW:AA	PW:SA	AA:SA
Fall Brook	9.38	9.52	9.70	1.49	3.34	1.85	36.91	31.28	18.03	36.40	30.62	18.02
Hunter Brook	5.88	5.86	6.01	0.31	2.11	2.43	25.18	24.06	7.88	25.43	24.29	7.89
Rondout Creek	24.19	19.39	19.35	22.07	22.24	0.18	34.24	34.26	1.46	34.45	34.68	1.46
West Kill	6.61	6.35	6.47	4.04	2.10	1.94	24.77	23.48	6.27	24.59	23.93	6.27
			Average	6.98	7.45	1.60	30.28	28.27	8.41	30.22	28.38	8.41

Appexdix D: Quality Control Results

The following tables contain information on the field blanks and duplicate sample results.

Field Blanks:

The table below shows the concentrations of the analytes of interest in the field blanks. Two blanks were analyzed each month and treated identically as stream samples.

Table 16: Analyte concentrations in the field blanks for the 2011 water year.

		Analyte Concentrations			
Month	Blank #	DOC (mg/L)	NO_3^- (umol/L)	TN (mg/L)	pH
October	1	0.212	0.034	0.007	6.06
	2	0.303	Non-detect	0.005	5.91
November	1	0.040	Non-detect	0.005	5.71
	2	0.056	Non-detect	0.014	5.61
December	1	0.077	Non-detect	0.019	5.71
	2	0.214	Non-detect	0.003	5.78
January	1	0.255	0.021	0.005	5.51
	2	0.195	0.007	0.006	5.56
February	1	0.118	0.018	0.009	5.58
	2	0.148	0.075	0.011	5.57
March	1	0.126	Non-detect	0.001	5.66
	2	0.347	0.067	0.005	5.69
April	1	0.284	0.022	0.002	5.71
	2	0.576	0.028	0.010	5.7
May	1	0.384	Non-detect	0.005	5.7
	2	0.222	Non-detect	0.003	5.81
June	1	0.579	0.042	0.008	4.81
	2	0.502	Non-detect	0.008	5.89
July	1	0.227	Non-detect	0.005	7.03
	2	0.251	Non-detect	0.015	5.95
August	1	0.135	N/A	0.004	5.75
	2	0.099	N/A	0.010	5.55
September	1	0.269	Non-detect	0.016	5.01
	2	0.324	Non-detect	0.046	5.63

Note that the field blank concentrations are very low compared to the concentrations expected from streams in forested systems.

Duplicate Samples:

The following table shows the concentrations of the analytes of interest for duplicate samples taken during the study period. This table shows all duplicates taken during the study period, even if the samples were taken from the broader list of study streams that were not used in this study. This was done because the number of duplicate samples taken from the study sites over the study period was relatively low and showing all duplicate samples should increase the confidence in the quality of the sampling and analytical methods.

Table 17: Analyte concentrations in duplicate samples.

			Analyte Concentration			
Month	Stream	Duplicate #	DOC (mg/L)	NO_3^- (umol/L)	TN (mg/L)	pH
October	Winnisook	1	2.685	4.545	0.121	4.86
		2	2.705	4.592	0.104	4.80
November	Kelly Hollow	1	0.778	28.545	0.420	6.74
		2	0.745	28.687	0.444	6.78
	Hollow Tree	1	0.807	14.018	0.208	6.75
		2	0.750	14.004	0.232	6.79
December	Pigeon	1	1.161	14.425	0.262	6.28
		2	1.150	14.241	0.234	6.36
	Black Brook	1	0.924	25.502	0.363	6.49
		2	0.926	25.704	0.364	6.52
January	Rondout	1	0.703	16.723	0.250	6.08
		2	0.674	16.563	0.261	6.16
	BWS6	1	0.559	2.480	0.067	6.46
		2	0.594	2.457	0.054	6.48
February	Willowemoc	1	0.882	21.961	0.317	6.58
		2	0.706	21.924	0.359	6.56
	BWS6	1	0.534	4.863	0.111	6.45
		2	0.507	4.837	0.104	6.47

			Analyte Concentration			
Month	Stream	Duplicate #	DOC (mg/L)	NO_3^- (umol/L)	TN (mg/L)	pH
March	Fall Brook	1	0.909	36.800	0.445	6.19
		2	0.804	36.789	0.466	6.28
	BWS6	1	0.747	0.558	0.044	6.30
		2	0.651	0.540	0.039	6.21
April	Myrtle	1	1.482	N/A	0.070	6.75
		2	1.420	1.453	0.055	6.76
	Hunter Brook	1	1.721	17.042	0.258	6.65
		2	1.211	17.304	0.252	6.67
May	Winnisook	1	3.770	6.862	0.137	4.75
		2	3.755	6.709	0.148	4.81
	Styles Brook	1	2.916	12.885	0.205	6.57
		2	2.920	13.108	0.223	6.67
June	Colgate Lake	1	3.920	10.917	0.211	6.00
		2	4.007	10.930	N/A	6.40
	Silver Spring	1	2.810	11.262	0.190	6.41
		2	2.905	11.142	N/A	6.52
July	Hollow Tree	1	2.308	15.456	0.250	6.33
		2	2.248	15.573	0.256	6.49
	Hunter Brook	1	2.081	11.837	0.212	6.32
		2	2.019	11.607	0.191	6.40
August	Willowemoc	1	0.668	15.270	0.244	6.48
		2	0.672	15.370	0.253	6.42
	Mill Brook	1	0.998	27.972	0.404	6.45
		2	0.946	27.921	0.482	6.39
September	Wase Road	1	1.151	10.361	0.214	6.06
		2	1.132	10.642	0.215	5.88
	Black Brook	1	0.798	17.540	0.321	6.39
		2	1.000	17.308	0.457	6.25

Continuing Calibration Verification Standards

All analytical data used in the flux estimates of this project were obtained from stream samples positioned between passing CCVs. If CCVs did not pass the corresponding samples were reanalyzed or rejected. Minimum detection limits and criteria for passing CCVs are shown in the following table.

Table 18: Miniumum detection limits and continuing calibration verification concentrations for analytes.

Analyte	MDL	Concentration of CCV
DOC (mg/l)	0.058	5
NO_3^- (umol/l)	0.25	20
TN (mg/l)	0.011	0.5
pH	N/A	4 and 7

Criteria for passing CCVs were that the measured concentration be ± 10% of the expected CCV concentration.

References

Anderson, S.P., Dietrich, W.E., Torres, R. (1997). Concentration-Discharge Relationships in runoff from a steep, unchanneled catchment. *Water Resources Research, 33*(1), 211-225.

Aulenbach, B.T., and Hooper, R.P. (2006), The Composite Method: An Improved Method for Stream-Water Solute Load Estimation. *Hydrological Processes, 20,* 3029-3047.

Archfield, S.A., and Vogel, R.M. (2010). Map Correlation Method: Selection of a Reference Streamgage to Estimate Streamflow at Ungaged Catchments. *Water Resources Research, 46.*

Beven, K. (2012). *Rainfall-Runoff Modelling: The Primer*. 2nd Edition. Hoboken, NJ: John Wiley and Sons.

Castellarin, A., Galeati, G., Brandimarte, L., Montanari, A., and Brath, A. (2004). Regional Flow-Duration Curves: Reliability for Ungauged Basins. *Advances in Water Resources, 27*(10), 953-965.

CESE-IC. Center for Environmental Systems and Engineering. (2011). Ion Chromatography with Chemical Suppression of Eluent Conductivity. *Standard Operating Procedure.*

CESE-TN. Center for Environmental Systems and Engineering. (2011). Total Nitrogen (With Teledyne Tekmar Apollo 9000). *Standard Operating Procedure.*

Clark, J.M., Lane, S.N., Chapman, P.J., and Adamson, J.K. (2007). Export of Dissolved Organic Carbon from an Upland Peatland during Storm Events: Implications for Flux Estimates. *Journal of Hydrology, 347*(3), 438-447.

Dai, A., Qian, T., Trenberth, K., and Milliman, J. (2008). Changes in Continental Freshwater Discharge from 1948 to 2004. *Journal of Climate, (22)*, 2773-2792.

Dann, M.S., Lynch, J.A., and Corbett, E.S. (1986). Comparison of Methods for Estimating Sulfate Export from a Forested Watershed. *Journal of Environmental Quality, 15*(2), 140-145.

Eimers, C.M., Buttle, J., and Watmough, S.A. (2008). Influences of Seasonal Changes in Runoff and Extreme Events on Dissolved Organic Carbon in Wetland- and Upland-Draining Streams. *Canadian Journal of Fisheries and Aquatic Sciences, 65*(5), 796-808.

Godsey, S.E., Wirchner, J.W., and Clow, D.W. (2009). Concentration-Discharge Relationships Reflect Chemostatic Characteristics of US Catchments. *Hydrological Processes, 23,* 1844-1864.

Gomi, T., Sidle, R.C., and Richardson, J.S. (2002). Understanding Processes and Downstream Linkages of Headwater Systems. *Bioscience, 52*(10), 905-916.

Grayson, R. and Holden, J. (2012). Continuous Measurements of Spectrophotometric Absorbance in Peatland Streamwater in Northern England: Implications for Understanding Fluvial Carbon Fluxes. *Hydrological Processes*, *26*, 27-39.

Hinton, M.J., Schiff, S.L., and English, M.C. (1997). The Significance of Storms for the Concentration and Export of Dissolved Organic Carbon from two Precambrian Shield Catchments. *Biogeochemistry*, *36,* 67-88.

Hortness, J. E. (2006). Estimating Low Flow Frequency Statistics for Unregulated Streams in Idaho. *US Geological Survey. Scientifi. Investigation Report 2006-5035.*

Johnson, N.M., Likens, G.E., Borman, F.H., Fisher, D.W., and Pierce, R.S. (1969). A Working Model for the Variation in Stream Water Chemistry at the Hubbard Brook Experimental Forest, New Hampshire. *Water Resources Research*, *5*(6), 1353-1363.

Koltun, G.F. and Shwartz, R.R. (1987). Multiple-Regression Equations for Low Flows at Ungaged Streams in Ohio. *US Geological Survey. Water Resources Investigation Report 86-4354.*

Kudish, M. (1979). *Catskill Soils and Forest History*. The Catskill Center for Conservation and Development, Inc. Hobart, NY.

Kudish, M. (2000). *The Catskill Forest: A History.* 1st Edition. Fleischmanns, NY: Purple Mountain Press.

Lowe, W.H., and Likens, G. E. (2005). Moving Headwater Streams to the Head of the Class. *Bioscience*, *55*(3), 196-197.

Lumia, R. (1991). Regionalization of Flood Discharges of Rural, Unregulated Streams in New York, Excluding Long Island. *Water-Resources Investigations Report 90-4197*. USGS.

McDaniel, P.A., Regan, M.P., Brooks, E., Boll, J., Barndt, S., Falen, A., Young, S.K., and Hammel, J.E. (2008). Linking Fragipans, Perched Water Tables, and Catchment-Scale Hydrological Processes. *Catena, 73*, 155-173.

Mohamoud, Y.M. (2008). Prediction of Daily Flow Duration Curves and Streamflow for Ungaged Catchments using Regional Flow Duration Curves. *Hydrological Sciences Journal, 53*(4), 706-724.

Mohamoud, Y.M. and Parmar, R.S. (2006). Estimating Streamflow and Associated Hydraulic Geometry, The Mid-Atlantic Region, USA. *Journal of the American Water Resources Association*, *42*(3), 755-768.

Mulholland, P.J. (2003). Large-Scale Patterns in Dissolved Organic Carbon Concentration, Flux, and Sources. In S. Finlay and R. Sinsabaugh (Eds.), *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter.* San Diego, California: Academic Press.

Mulvihill, C.I., Baldigo, B.P., Miller, S.J., DeKoskie, D., DeBois, J. (2009). Bankfull Discharge and Channel Characteristics of Streams in New York State. *Scientific Investigations Report 2009–5144* . USGS.

Ohio EPA. (2009). Total Maximum Daily Loads for the White Oak Creek Watershed. Final Report. December, 2009.

Post, D.A. and Jakeman, A.J. (1999). Predicting the Daily Streamflow of Ungauged Catchments in S.E. Australia by Regionalising the Parameters of a Lumped Conceptual Rainfall-Runoff Model. *Ecological Modelling*, *123*, 91-104.

Ries, K.G. and Friesz, P.J. (2000). Methods for Estimating Low-Flow Statistics for Massachusetts Streams. *US Geological Survey. Water-Resources Investigations Report 00-4135.*

Riggs, H.C. (1990). Estimating Flow Characteristics at Ungauged Sites. In *Regionalisation in Hydrology, Proceedings of the Ljubljan Symposium, April 1990*, IAHS Publication No. 191.

Rusjan, S., Brilly, M., and Mikos, M. (2008). Flushing of Nitrate from a Forested Watershed: An Insight into Hydrological Nitrate Mobilization Mechanisms through Seasonal High-Frequency Stream Nitrate Dynamics. *Journal of Hydrology, 354*(2), 187-202.

Rusjan, S. and Mikos, M. (2010). Seasonal Variability of Diurnal In-Stream Nitrate Concentrate Oscilliations Under Hydrologically Stable Conditions. *Biogeochemistry, 97,* 123-140.

Seary, J.K. (1959). Flow-Duration Curves. *Manual of Hydrology: Part 2. Low Flow Techniques. US Geological Survey Water Supply Paper 1542-A.*

Semkin, R.G., Jefferies, D.S., and Clair, T.A. (1994). Hydrochemical Methods and Relationships for Study of Steam Output from Small Catchments. In B. Molden and J. Cerni (Eds.), *Biogeochemistry of Small Catchments: A Tool for Environmental Research*. (163-187). New York: John Wiley and Sons.

Smakhtin, V.U. (1999). Generation of Natural Daily Flow Time Series in Regulated Rivers using a Nonlinear Spatial Interpolation Technique. *Regulated Rivers: Research and Management, 15,* 311-323.

Sullivan, T.J., Christophersen, N., Muniz, I.P. Seip, H.M. and Sullivan, P.D. (1986). Aqueous Aluminum Chemistry Response to Episodic Increases in Discharge. *Nature, 323*, 324-327.

Tan, K.S., Chiew, F.H.S., Grayson, R.B., Scanlon, P.J., Siriwardena, L. (2005). Calibration of a Daily Rainfall-Runoff Model to Estimate High Daily Flows. In: MODSIM 2005

International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand, Melbourne, December 2005, pp. 2960-2966.

Teledyne Tekmar. (2003). TOC Combustion Analyzer. *User Manual*.

Titus, R. (1998). *The Catskills: A Geological Guide*. Fleischmanns, NY: Purple Mountain Press.

USGS. (2007). Methods used to Estimate Nutrient Fluxes. Streamflow and Nutrient Fluxes of the Mississippi-Atchafalaya River Basin and Subbasins for the Period of Record Through 2005. Open File Report.

Vogel, R.M., Wilson, I., and Daly, C. (1999). Regional Regression Models of Annual Streamflow for the United States. *Journal of Irrigation and Drainage Engineering, 125*(3), 148-157.

Wagener, T. and Wheater, H. (2006). Parameter Estimation and Regionalization for Continuous Rainfall-Runoff Models Including Uncertainty. *Journal of Hydrology, 320,* 132-154.

Walling, D.E, and Webb, B.W. (1986). Solutes in River Systems. In: *Solute Processes* (Ed. S.T. Trudgill), 251-327. Wiley and Sons: Chichester, UK.

Webster, R. (1997). Regression and Functional Relations. *European Journal of Soil Science, 48,* 557-566.

Vita

Name of Author: Chris Gianfagna

Place of Birth: Watertown, New York

Date of Birth: November 27, 1987

Undergraduate Education: SUNY Plattsburgh, Plattsburgh, New York

Graduate Education: Syracuse University, Syracuse, New York

Degrees Awarded:
Bachelor of Science in Environmental Science, SUNY Plattsburgh, 2010

Honors and Awards:
Transfer Achievement Scholarship, SUNY Plattsburgh

Professional Experience:
Teaching Assistant, Dept. of Civil and Environmental Engineering, Syracuse University, 2010
Research Assistant, Dept. of Civil and Environmental Engineering, Syracuse University, 2011