ABSTRACT

GILMORE, TROY EDWARD. Groundwater Transit Times and the Fate of Aquifer Nitrate: Observations from Sampling in Stream Channels and Well Nests in an Agricultural Watershed, North Carolina, USA. (Under the direction of David Genereux, François Birgand, and R. Wayne Skaggs).

Four groundwater sampling approaches were used to study the transit times and fate of agricultural nitrate in a surficial aquifer. Three stream-based measurement approaches were applied in a 2.5 km reach of West Bear Creek: point measurements beneath the streambed, seepage blankets (novel seepage meter design), and reach mass-balance. Groundwater was also sampled from two nearby well nests. Stream-based methods gave similar mean groundwater seepage rates into the stream (0.3-0.6 m/day) over the course of two separate 3-4 day field campaigns even though stream discharge differed by a factor of ~10 between the two campaigns. At low flow, data on the flow-weighted mean nitrate concentration in groundwater discharge, $[NO_3^-]_{FWM}$, and nitrate flux from groundwater to the stream suggested lower values with increasing opportunity for channel-streambed influence as the measurement scale increased, e.g., $[NO_3^-]_{FWM}$ of 654, 561, and 451 μM for points, blankets, and reach mass-balance, respectively. At high flow the trend was reversed, likely because reach mass-balance captured greater inputs from shallow transient high-nitrate flow paths. Streambed point sampling may be better suited to estimating aquifer discharge of nitrate, and reach mass-balance better suited to estimating effective nitrate input into streams (which at high flow may be more than aquifer discharge due to transient flowpaths and at low flow may be less, due to channel retention). Modeling dissolved N_2 from point and blanket samples and from samples collected from nearby well nests suggested about 50 – 78% of groundwater nitrate was denitrified in the aquifer. Streambed sampling suggested both extent of denitrification and initial nitrate concentration in groundwater (700-1300 μM) were related to land use, and that these forms of streambed sampling for groundwater can reveal watershed spatial relations relevant to nitrate fate in the aquifer. Based on age-dating tracer data from streambed point sampling, the mean transit time (*MTT*) of groundwater (flow-weighted mean apparent groundwater age) in the surficial aquifer is about 30 years, similar to the *MTT* from well data (27 years). Blanket data, after corrections for a fraction of stream water in blanket samples, suggested similar groundwater apparent ages compared to point

samples collected adjacent to the blankets. Transit time distributions (*TTD*) for groundwater discharging to the stream were best fit by a gamma distribution model and were very consistent, despite different hydrologic conditions and sampling arrangements (closely-spaced transects in 58 m reach vs widely spaced transects in 2.5 km reach). This consistency suggests that coupled measurements of groundwater seepage rate and groundwater age may be a robust approach for determining aquifer *TTD* as well as *MTT*. An exponential-piston flow model also gave a reasonable fit to the *TTD* from point measurements, and to the age distributions observed from well nest data, which suggests significant spatial variation in recharge rates in the watershed. This has potentially important implications for future fluxes of NO_3^- to the stream, because the *TTD* for aquifer discharge to the stream is dominated by a large fraction of older groundwater (20-40 years). As a result, NO_3^- fluxes will initially respond slowly to changes in nitrogen use in the watershed before responding more rapidly during a period 20-40 years after the changes are made. Historic patterns of $[NO_3^-]$ in groundwater suggest that peak $[NO_3^-]$ may have occurred 10-20 years ago, which means that fluxes of NO_3^- from the aquifer could potentially peak in the next 1-2 decades before returning to current levels as high-nitrate groundwater is flushed from the aquifer. Collectively, the results of this study highlight the importance of directly sampling aquifer discharge to better understand historic and spatial patterns of contamination and the likely groundwater system response to nutrient management in agricultural watersheds.

Groundwater Transit Times and the Fate of Aquifer Nitrate: Observations from Sampling in Stream Channels and Well Nests in an Agricultural Watershed, North Carolina, USA

by
Troy Edward Gilmore

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Biological and Agricultural Engineering

Raleigh, North Carolina

2015

APPROVED BY:

_____	_____
Prof. David P. Genereux Committee Co-Chair	Prof. François Birgand Committee Co-Chair
_____	_____
Prof. R. Wayne Skaggs Committee Co-Chair	Prof. Helena Mitasova

DEDICATION

This work is dedicated to my wife and daughters.

BIOGRAPHY

The author received his A.S. in Mechanical Engineering Technology from the University of Akron and his B.S. in Civil Engineering from North Carolina State University.

ACKNOWLEDGMENTS

The author gratefully acknowledges the patience, support, and guidance of his research advisor Prof. David Genereux, and the encouragement and support of Prof. François Birgand. The author would also like to thank committee members Prof. Helena Mitasova and Prof. R. Wayne Skaggs for providing constructive comments. Special thanks are in order for collaborator Prof. D. Kip Solomon at the University of Utah for his guidance, data collection, and analysis of dissolved gas samples, and for John Solder at the USGS Utah Water Science Center for data collection, analysis, and helpful comments. Bryant Kimball, Vic Heilweil, and James Marlowe (all with USGS) each made important contributions during field campaigns. Matthew Burnette and Scott Becker contributed significantly to the site selection and data collection efforts. The author also appreciates the contribution of Kathleen Farrell of the North Carolina Geological Survey, who provided valuable assistance in interpreting geologic data. Financial support for this work was provided from the National Science Foundation under Award EAR-1045162 and is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

QUANTIFYING THE FATE OF AGRICULTURAL NITROGEN IN AN UNCONFINED AQUIFER: STREAM-BASED OBSERVATIONS AT THREE MEASUREMENTS SCALES

1.1. Abstract

We compared three stream-based sampling methods to study the fate of agricultural nitrate in groundwater in a coastal plain watershed: point measurements beneath the streambed, seepage blankets (novel seepage-meter design), and reach mass-balance. The three methods gave similar mean groundwater seepage rates into the stream (0.3-0.6 m/day) over the course of two separate 3-4 day field campaigns even though stream discharge differed by a factor of ~10 between the two campaigns. At low flow, data on the flow-weighted mean nitrate concentration in groundwater discharge, $[NO_3^-]_{FWM}$, and nitrate flux from groundwater to the stream showed lower values with increasing opportunity for channel-streambed influence as the measurement scale increased, e.g., $[NO_3^-]_{FWM}$ of 654, 561, and 451 μM for points, blankets, and reach mass-balance, respectively. At high flow the trend was reversed, likely because reach mass-balance captured inputs from shallow transient high-nitrate flow paths while point measurements did not. Streambed point sampling may be better suited to estimating aquifer discharge of nitrate, and reach mass-balance better suited to estimating effective nitrate input into the channel (which at high flow may be more than aquifer discharge due to transient flowpaths and at low flow may be less than aquifer discharge due to channel retention). Modeling dissolved N_2 from point and blanket samples suggested (1) about half of groundwater nitrate was denitrified prior to discharge from the aquifer, and (2) both extent of denitrification and initial nitrate concentration in groundwater (700-1300 μM) were related to land use, suggesting these forms of streambed sampling for groundwater can reveal watershed spatial relations relevant to nitrate fate in the aquifer.

1.2. Introduction

Direct measurement of nitrogen (N) fluxes from groundwater to streams is needed to assess the contribution of contaminated groundwater systems to water quality impairment in streams and estuaries, and associated algal blooms, hypoxia, and fish kills (Diaz and Rosenberg 1995; Hayashi and Rosenberry 2002; Paerl 2006; Obenour et al., 2012). Groundwater discharge to streams can account for a large fraction of total stream flow (Bachman et al., 1998; Alley et al., 1999; Clow et al., 2003; Holmes 2000; Lindsey et al., 2003; Puckett et al., 2008; Winter et al., 1998). In this paper, we address the challenge of using field measurements to quantify the two principal fates of nitrate in an unconfined aquifer: the discharge flux from the aquifer to a stream, and the denitrification occurring in the aquifer. In two companion papers (Solomon et al., 2015; Gilmore et al., 2015b) we use related field measurements to estimate groundwater transit time, an important parameter showing the time scale over which N moves through groundwater from recharge at the water table to discharge in streams. A third paper explores trends in nitrate contamination and groundwater age from streambed and well nest sampling (Gilmore et al., 2015a).

Most of our measurements were conducted in a streambed, an information-rich zone where groundwater N signals from decades of agricultural land use may be sampled as groundwater from the surficial aquifer discharges to a stream (Kennedy et al., 2009a,b; Browne and Guldan, 2005; Duff et al., 2008; Tesoriero et al., 2005, 2013; Böhlke et al., 2004, 2007; Lindsey et al., 2003; Stelzer et al., 2011a,b). Kalbus et al. (2006) reviewed the limited number of approaches available for observing N and water signals at the groundwater – surface water interface but did not make quantitative comparisons. We simultaneously applied three approaches, each with a different measurement support scale and balance of pros and cons:

- the streambed point approach, based on point-scale measurements of both streambed seepage rate and the concentrations of solutes (N species, dissolved gases) in groundwater in or just below the streambed (Kennedy et al., 2009a,b),

- the streambed blanket approach, based on collection/integration of groundwater seepage beneath an experimental " seepage blanket" deployed on top of the streambed (Solder, 2014), and

- the reach mass-balance approach, based on estimating, from tracers released to the stream, the flow-weighted mean concentrations of N species in the net groundwater input to a stream reach (McMahon & Böhlke 1996; Stolp et al., 2010).

These approaches ultimately seek integration of groundwater N discharge at the scale of a stream reach but from different starting points. The streambed point approach is based on integration of numerous groundwater point samples from within or just below the streambed (each with an effective length scale on the order of 0.1 m in terms of distance from the sampling screen, or ~0.03 m^2 in terms of streambed area). Streambed blanket sampling integrates groundwater seepage over streambed areas of 1-5 m^2. Reach mass-balance is based on the natural integration of groundwater flowpaths in a receiving stream reach at reach length scales of about 10^2-10^3 m, or streambed area scales of 10^2-10^4 m^2 for small streams 1-10 m in width.

During our two 3-4 day field campaigns, stream and groundwater samples collected at the point, seepage blanket, and reach mass-balance scales were analyzed for dissolved N species (N_2, NO_3^-, NH_4^+), noble gases (Xe, Kr, Ne, Ar, He), and other geochemical parameters. These data were combined with groundwater discharge estimates from each of the three spatial scales. The overall purpose was to compare the three field sampling scales to see if they give fundamentally different pictures of the fate of agricultural nitrate in the surficial aquifer, specifically, (1) the water and N flux out of the aquifer to the study stream, and (2) the extent of denitrification in the groundwater discharging from the aquifer to the stream.

1.3. Background

1.3.1. Overview of sampling approaches

The different spatial scales of integration inherent in the three approaches (point, blanket, and reach mass-balance) suggest contrasting general pros and cons (Table 1.1).

Table 1.1. Comparison of three sampling approaches

Issue or characteristic	Approach		
	point	blanket	RMB
shows patterns of concentrations & fluxes on streambed	yes	somewhat	no
integration by calculation or nature	calculation	both	nature
samples individual groundwater flowlines	possibly	no	no
number of samples, cost of analysis	high	moderate	low
samples solely groundwater (no mixing with stream)	yes	mostly	no
detection of seepage inputs above water line	no	no	yes
gw input << stream discharge causes high uncertainty	no	no	yes
feasible in all streambed types (e.g., silt, sand or rock)	no	no	yes
requires injected tracers in stream	no	likely	yes

"gw" = groundwater, "RMB" = reach mass-balance

1.3.2. Streambed point approach

Groundwater N fluxes into streams have previously been measured using the point approach (Kennedy et al., 2007, 2008, 2009a,b). For closely-spaced points, N concentration and groundwater age showed distinct patterns in the streambed, suggesting that point samples were representative of individual (relatively unmixed) groundwater flowlines, and making it possible to relate point N fluxes to controlling factors such as streambed permeability, groundwater age, or extent of denitrification. Other studies have combined groundwater concentrations from point sampling with groundwater inflow estimates from reach mass-balance experiments (Böhlke et al., 2004; Duff et al., 2008; Stelzer et al., 2011a) but have largely focused on in-stream cycling of N. Unlike reach mass balance, point measurements do not require injected tracers and can quantify groundwater discharge even when it is small relative to stream discharge. The point approach requires both groundwater sampling and discharge measurement at each point to calculate N flux and flow-weight mean N concentration in the groundwater, and is applicable in streambeds where groundwater probes can be installed and sampled.

1.3.3. Streambed blanket approach

Seepage meters offer an intermediate measurement scale between the point and reach mass-balance approaches and have been previously used to determine groundwater flux through lakebeds or streambeds (Bower and Rice 1968; Lee 1977; Rosenberry 2008; Solder, 2014). Individual seepage meters typically integrate groundwater signals on the scale of <1 m^2 (the streambed area covered by a meter) (Kalbus et al., 2006), although larger streambed areas can be integrated using "ganged" seepage meters (Rosenberry, 2005). Groundwater flux estimates from traditional seepage meters have been compared to other measurement methods (Cey et al., 1998; Rosenberry and Pitlick 2009; Kennedy et al., 2010), with limited agreement. While the seepage meter concept is theoretically simple (analogous to an inverted funnel placed on the streambed surface), the literature addressing seepage meter performance (e.g., Corbett et al., 2003; Murdoch and Kelly 2003), calibration (Rosenberry and Menheer

2006), and design improvements (e.g., Krupa et al., 1998) suggests that the approach is not trivial.

1.3.4. Reach mass-balance approach

Relatively few studies have used the reach mass-balance approach for the primary purpose of calculating the flux of N exiting the surficial aquifer (e.g., McMahon and Böhlke 1996; Burns 1998; Chestnut and McDowell 2000; Angier et al., 2008; Stelzer et al., 2011a), possibly because in-stream processes may be expected to alter the apparent groundwater N flux at the reach mass-balance scale (Laursen and Seitzinger, 2002, 2005; Böhlke et al., 2004, 2009; Mullholland et al., 2004; Duff et al., 2008; Gu et al., 2008; Baulch et al., 2010). Reach mass-balance measurements in streams have been used to estimate groundwater concentrations and groundwater discharges of other pollutants, such as VOCs (Kim and Hemond 1998), metals from mining activities (Kimball et al., 2002), or dissolved gases (e.g., Hemond and Duran 1989; Genereux et al., 1993; Stolp et al., 2010). The reach mass-balance approach is well suited for reaches long enough that the groundwater input is a large percentage of stream discharge. Groundwater discharge fluxes are naturally integrated with this approach so no interpolation or mathematical integration is necessary, but it is only possible to observe the spatial structure of groundwater inputs longitudinally and at a coarse scale. Groundwater N input to a stream reach is computed as the N flux out the downstream end of the reach minus the N flux in the upstream end of the reach (and any N flux from tributaries), possibly with a correction for in-channel N retention. Fluxes or concentrations of dissolved gases in groundwater require corrections for interaction with the atmosphere when using the reach mass-balance approach, which requires injected gaseous tracers (Bennett and Rathbun 1972; Kilpatrick et al., 1989; Wanninkhof et al., 1990; Genereux and Hemond 1992). Reach mass-balance generally requires fewer samples which lowers analytical costs, but this may be partially offset by the cost of injected tracers.

1.4. Study Area and Hydrologic Conditions

1.4.1. Study site

The study was conducted in West Bear Creek, in the coastal plain of North Carolina. West Bear Creek is about 7 m wide, channelized and entrenched, and located in a productive agricultural area (Fig. 1.1A). The West Bear Creek watershed (Fig. 1.1B), previously described in Kennedy et al. (2007, 2008, 2009a,b, 2010) and Genereux et al. (2008), has roughly 60% agricultural land use. The study was conducted in the stream reach between a tracer injection site and a monitoring site 2.7 km downstream (Fig. 1.1C). The furthest-upstream measurements were made 200 m downstream of the tracer injection site, thus the study reach was effectively 2.5 km in length (200 m to 2700 m). The topographically defined contributing area to the 2.5 km reach comprises about 10% of the West Bear Creek watershed and contains two concentrated animal feeding operations (at the waste lagoons in Fig. 1.1C). Land not forested is primarily used for row crops, with the exception of pasture for cattle along the north side of the stream, roughly between the 1260 m and 2530 m stream sites.

Kennedy et al. [2009a] observed high NO_3^- concentrations, $[NO_3^-]$, in groundwater seepage to West Bear Creek (mean of 438 μM, range from <7 to 1785 μM) west of the State Route 1719 bridge (Fig. 1.1C). Kennedy et al. (2009b) found that denitrification was an important process in the surficial aquifer, reducing the nitrate flux from the surficial aquifer to a 75 m reach of West Bear Creek by about half.

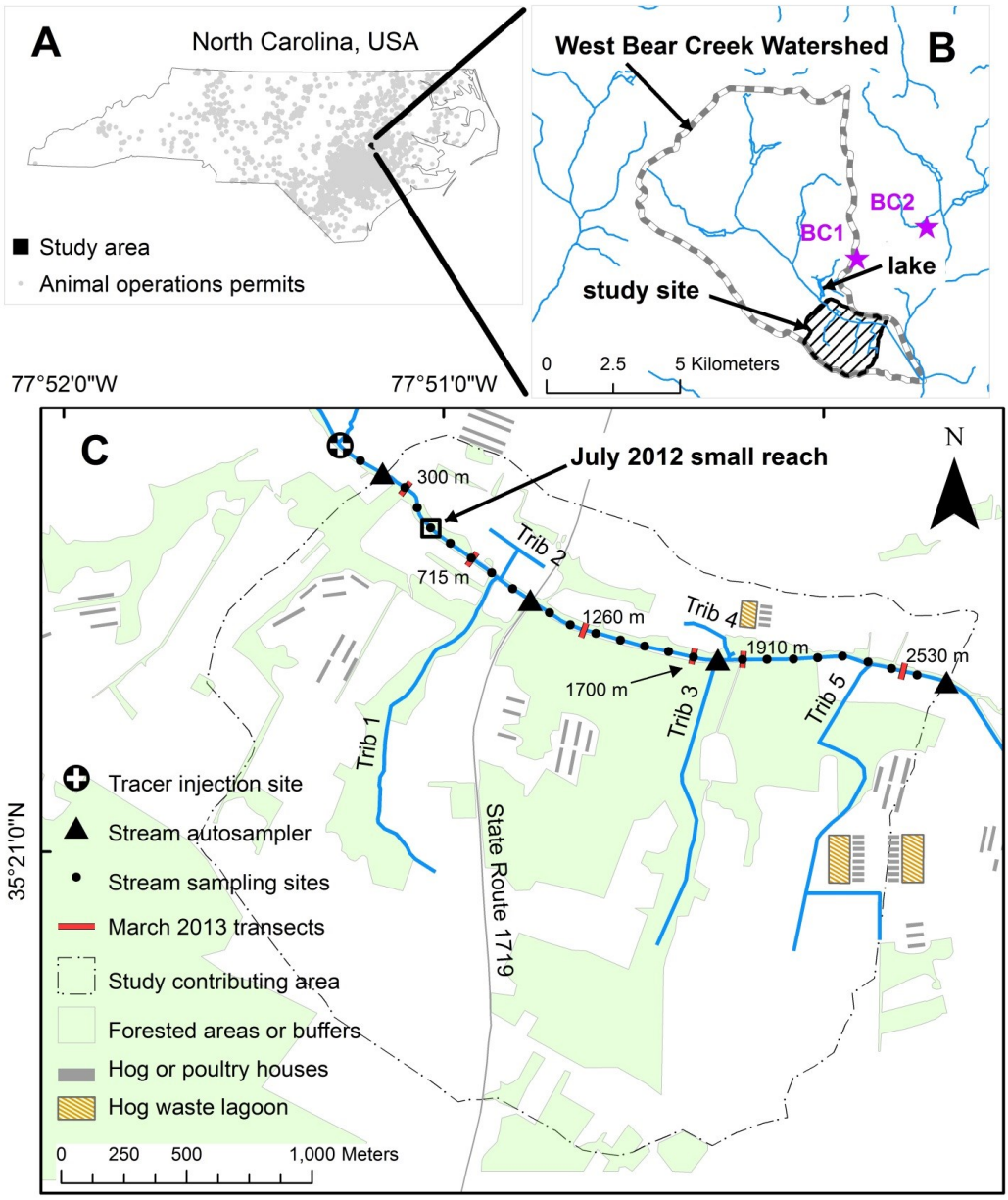

Figure 1.1. West Bear Creek study site layout for the July 2012 and March 2013 field campaigns. A: animal feeding operations in eastern North Carolina. B: the West Bear Creek watershed is outlined by the thick gray dashed line, and the topographically-defined contributing area for the 2.7 km study reach is defined by the cross-hatched area. Well nests are denoted by stars. C: the topographically-defined contributing area for the West Bear Creek study reach (outlined by thin dot-dash line), animal barns (gray rectangles), and forested areas (light green shaded areas). Stream autosamplers were 0.2-2.7 km downstream of the tracer injection site. Stream sampling sites (small black circles) were distributed at 100 meter intervals. In July 2012 all eight point transects and two blanket transects were close

together in the 58 meter "July 2012 small reach". In March 2013 six point transects were distributed throughout the 2.7 km reach. Labels for the March point transects are in meters downstream of the injection site. Blanket sampling was also conducted at the 715m transect in March. All GIS data were accessed via the NC OneMap Geospatial Portal (data.nconemap.com). Forested areas, agricultural facilities, and tributaries were defined using digital orthophotos (2010 North Carolina Statewide Digital Orthoimagery) and field observations. The contributing area for the 2.7 km reach is based on digital elevation data from the North Carolina Division of Transportation. The West Bear Creek watershed outline is from the USDA NC NRCS 12-Digit Hydrologic Units data set. The main channel of West Bear Creek and locations for animal operations permits were from North Carolina Division of Environment and Natural Resources (NCDENR) data sets.

1.4.2. Background hydrologic conditions

For July 2012 and March 2013, stream discharge at the Bear Creek USGS stream gauge (Fig. 1.2; http://waterdata.usgs.gov/usa/nwis/uv?0208925200), 10 km downstream of the downstream end of our study reach, was near the 25-year medians for July and March, respectively. In July 2012, groundwater level at NCDENR well O30J4 12 km northwest of the study site (the nearest surficial aquifer well with more than 5 years of monitoring data, http://www.ncwater.org/) was about 3 cm above the 30-year historic median for July. In March 2013 water level at the Saulston well was about 29 cm higher than in July 2012, but about 13 cm lower than the 30-year historic median for March.

Mean temperature from temperature probes at 0 m and 1000 m was used in dissolved gas modeling (explained below) for July 2012 (July 18, from 6:30 – 14:00); data at 0 m were used for March 2013 modeling (March 15, from 6:45 – 17:45). Stream water pH was about 6.6 in July 2012 but was not recorded in March 2013. Groundwater pH, conductivity and temperature were recorded during both field campaigns (Table 1.3).

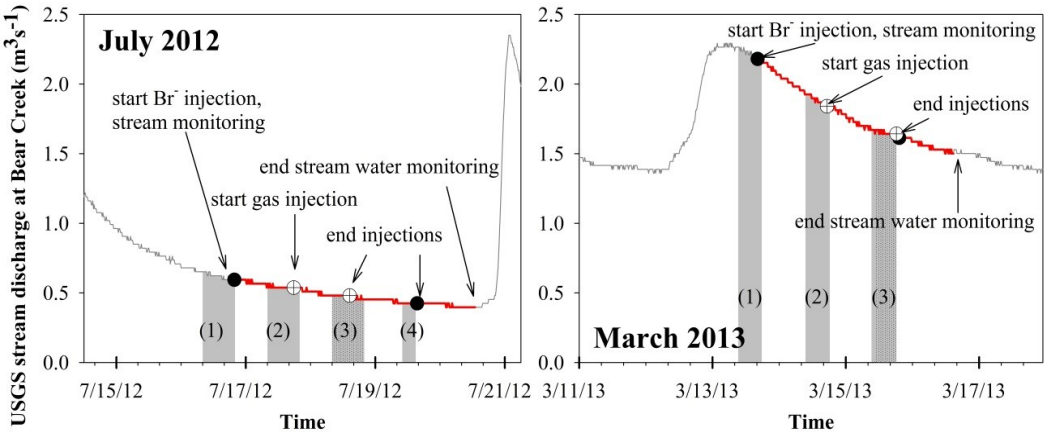

Figure 1.2. Hydrologic conditions and timelines for the July 2012 and March 2013 field campaigns. Hydrographs are from the USGS stream gauge at Mays Store, NC, downstream of the West Bear Creek study site. Shaded areas indicate times during which field work occurred and are labeled numerically (Table 1.2). Labels on the horizontal axis indicate 12am of the date shown, and intermediate ticks are located at 8-hour intervals. Start and end times for Br^- and gas injections are indicated by black circles and white circles with crosses, respectively.

Table 1.2. Sampling timeline for July 2012 and March 2013 field campaigns.

Time Period[a]	July 2012[b]	March 2013[b]
1	background sampling for Br^- and dissolved gases in stream water (pre-injection), point sampling	background sampling for Br^- and dissolved gases in stream water (pre-injection), point sampling
2	point and blanket sampling	point and blanket sampling
3	sampling for Br^- and dissolved gases in stream water (post-injection), point sampling	sampling for Br^- and dissolved gases in stream water (post-injection), manual stream gauging, point sampling
4	blanket sampling	n/a

[a]time periods correspond to numbers in parentheses in Fig. 1.2
[b]differences in July and March sampling approaches are described in Table 1.5

Table 1.3. Mean water quality parameters for July 2012 and March 2013 field campaigns.

Date	Groundwater[a] pH	Groundwater[a] Conductivity (μS/cm)	Groundwater[a] temperature (°C)	Stream water temperature (°C)
July 2012	4.7 (0.5)	224 (48)	20.3 (1.3)	25.2 (1.3)
March 2013	5.1 (0.6)	154 (87)	13.2 (0.9)	12.3 (1.8)

[a]Groundwater values are means from the streambed point measurements. Standard deviations (1σ) are shown in parentheses.

1.5. Methods

1.5.1. Overview

The study design and methodology integrated a number of published approaches (Table 1.4).

Table 1.4. Field methods and study design.

approach	source of methods and study design
streambed point, measurement methods	Kennedy et al. (2007), Genereux et al. (2008)
streambed point, dense sampling design	Kennedy et al. (2008, 2009a)
streambed point, distributed sampling design	Stelzer et al. (2011a)
reach mass-balance, conservative tracer injection	Kimball and Runkel (2009)
reach mass-balance, dissolved gas injection	Stolp et al. (2010), Solomon et al. (2015)
streambed blanket, design and flow meter	Solder (2014)

Two 3-4 day field campaigns were conducted, one in July 2012 and one in March 2013, to estimate the following principal groundwater variables:

- v = groundwater flux through the streambed and into the stream
- $[NO_3^-]$ = nitrate concentration in groundwater seepage to the stream (measured on individual point or blanket samples)
- $[NO_3^-]_{FWM}$ = flow-weighted mean (FWM) nitrate concentration in groundwater seepage to the stream (averaged for points or blankets)
- $[NO_3^-]^0$ = initial nitrate concentration in the groundwater at the time of recharge = what $[NO_3^-]$ would be in the absence of denitrification
- $[NO_3^-]^0_{FWM}$ = flow weighted mean initial nitrate concentration in groundwater = what $[NO_3^-]_{FWM}$ would be in the absence of denitrification
- f_{NO3} = nitrate flux through the streambed from groundwater to stream
- $f_{NO3}^{\ 0}$ = what f_{NO3} would be in the absence of denitrification
- $[N_2\text{-den}]$ = the amount of excess N_2 created in the groundwater by denitrification in the groundwater

All solute concentrations presented in this paper are groundwater concentrations, unless otherwise noted.

Table 1.5. Description of July 2012 and March 2013 field campaigns.

	July 2012	March 2013
dates: point data	7/16 – 7/18	3/13 – 3/15
dates: blanket data	7/16 – 7/19	3/13 – 3/15
dates: reach mass-balance data	7/16 (stream water background) and 7/18	3/14 (background, Br^- at injection site only) and 3/15
number/arrangement of points	40: eight closely-spaced 5-point transects across channel	30: six widely-spaced 5-point transects across channel
distance spanned by point transects	58 m	2230 m
number/arrangement of blankets	2 transects of 5 blankets (n=10)	1 transect of 5 blankets (n=5)
location of tracer injection site	0 m	0 m
extents of large/small reaches	large: 0.2-2.7km, small: 0.4-0.6km	large: 0.2-2.7km, small: 0.2-1.0km, 1.0-1.8km, 1.8-2.7km
Location of point and blanket transects[a]	466, 474, 481, 491, 499, 508, 516, and 524 meter locations	300, 715, 1260, 1700, 1910, and 2530 meter locations
Br- injection rate and duration	258 mg/s, 68 hours	514 mg/s, 41 hours
Kr injection duration	21 hours	27 hours
Xe injection duration	21 hours	n.a.[b]
locations of stream water sampling sites: Br^- and NO_3^-	100m intervals, from 200m to 2700m (no samples at 1000m, 2200m)	100m intervals, from 200m to 2700m
locations of stream water autosampler sites: Br^- and NO_3^-	200, 466, 524, and 2700m	200, 1000, 1800, and 2700m
locations of stream water sampling sites: noble gases, N_2, age-dating tracers	200, 466, 1000, 1500, 2100, and 2700 meter locations	200, 500, 800, 1000, 1200, 1500, 1800, 2100, 2400, and 2700 meter locations
locations of manual stream gauging (Sontek FlowTracker ADV)	500m, 2700m	0 m, 810 m, 860 m, 1700 m, 1840 m, 2380 m, 2420 m, 2700 m
locations of stream temp. and stage	0 m, 1000 m	0 m, 1000 m, 2700 m
variables measured in groundwater collected at streambed points	*K*, *J*, pH, T, spec. cond., DO, noble gases, tritium, N species ([TDN], $[NO_3^-]$, $[NH_4^+]$), $[Br^-]$, $[Cl^-]$, [CFC]s, $[SF_6]$, other dissolved gases	*K*, *J*, pH, T, spec. cond., DO, noble gases, tritium, N species ([TDN], $[NO_3^-]$, $[NH_4^+]$), [CFC]s, $[SF_6]$, other dissolved gases
variables calculated at points	v, f_{NO3}, [N_2-den], apparent age, $[NO_3^-]$, f_{NO3}^0	v, f_{NO3}, [N_2-den], apparent age, $[NO_3^-]$, f_{NO3}^0
depth of gw sampling in streambed	31-36 cm	31-36 cm
interval for K measurement in streambed	0-36 cm	0-36 cm

Table 1.5, Continued

variables measured in groundwater collected at blankets	DO, noble gases, tritium, N species ([TDN], $[NO_3^-]$, $[NH_4^+]$), $[Br^-]$, $[Cl^-]$, [CFC]s, $[SF_6]$, other dissolved gases	DO, noble gases, tritium, N species ([TDN], $[NO_3^-]$, $[NH_4^+]$), $[Br^-]$, $[Cl^-]$, [CFC]s, $[SF_6]$, other dissolved gases
variables calculated at blankets	v, f_{NO3}, $[N_2$-den], apparent age, $[NO_3^-]$, f_{NO3}^0	v, f_{NO3}, $[N_2$-den], apparent age, $[NO_3^-]$, f_{NO3}^0
variables measured in stream water	$[Br^-]$, noble gases, tritium, N species ([TDN], $[NO_3^-]$, $[NH_4^+]$), [CFC]s, $[SF_6]$, other dissolved gases, T, stage, conductivity	$[Br^-]$, noble gases, tritium, N species ([TDN], $[NO_3^-]$, $[NH_4^+]$), [CFC]s, $[SF_6]$, other dissolved gases, T, stage
variables calculated from stream water data	v, f_{NO3}, apparent age, $[NO_3^-]$	v, f_{NO3}, apparent age, $[NO_3^-]$
vertical profiles of streambed groundwater (porewater) chemistry	Diffusion samplers[c], 487L and 522L point locations, deployed Oct, 2012	MINIPOINT[d] and gas diffusion samplers[e] at 715m, MINIPOINT at 300m and 1910m

[a]Underlining indicates both point and blanket sampling on same transect; other transects were point sampling only.

[b]Xe data were not used in March because the tracer ran out before the stream sites could be sampled at steady state.

[c]Diffusion samplers (Hesslein 1976) were used to collect NO_3^- samples.

[d]USGS MINIPOINT (Harvey and Fuller 1998; Duff et al., 1998) was used to sample for NO^-, Br^-, and Cl^-.

[e]Gas diffusion samplers (Aeschbach-Hertig and Solomon 2013) were used to sample for Xe, Kr, Ar, Ne, He.

1.5.2. Reach mass-balance: methods

Steady Br^- tracer injections to the stream at the 0 m location (Fig. 1.1C) were used to determine groundwater input to West Bear Creek (Kilpatrick and Cobb 1985; Kimball et al., 1999; Kimball and Runkel, 2009). Gaseous tracers (Kr, Xe) were injected to determine gas transfer velocities for West Bear Creek (Bennett and Rathbun 1972; Kilpatrick et al., 1989; Wanninkhof et al., 1990; Genereux and Hemond 1992). The gas exchange information was needed for trace gas modeling described below. An automated sampler was used to collect stream water samples at the 2700 m stream monitoring site (Fig. 1.1C) for Br^- tracer analysis during both field campaigns; sampling frequency was 0.5-2 hours. In March 2013, automated sampling also occurred at 200 m, 1000 m, and 1800 m (Fig. 1.1C). In July 2012, the 200 m, 466 m, and 524 m sites were sampled manually at intermittent time intervals.

Manual stream water sampling for the Br^- tracer (100 m intervals) was carried out during both campaigns when the stream Br^- concentrations reached steady-state. Stream water samples for dissolved gas analysis (Xe, Kr, Ar, Ne, He, N_2) were also collected but from stations spaced farther apart (Table 1.5). In March 2013 NO_3^- samples were collected at the stream sites for dissolved gas sampling and from the automated samplers, while in July stream water $[NO_3^-]$ was measured in the same samples collected for Br^- analysis. In July 2012 a DH-81 sampler (Edwards and Glysson, 1999) was used to manually integrate stream water samples across the channel width (to minimize any effects of incomplete mixing of stream water). In March 2013 temporally-averaged concentrations of Br^- and NO_3^- (averaged over roughly a 4-5 hour period) from automated samplers were used for reach mass-balance calculations.

Stream discharge at stream sampling sites was calculated from Br^- data using a standard steady-state expression for chemical dilution stream gauging (stream discharge = Br^- injection rate divided by steady-state Br^- concentration in stream water (Kimball et al., 1999)). Groundwater flux based on reach mass-balance was then calculated as:

$$v = \frac{Q_{down} - Q_{up} - \sum Q_{trib}}{wL} \quad (1.1)$$

where v is groundwater flux into West Bear Creek (m/day) and the subscript RMB indicates the estimate is from reach mass-balance, Q_{down} and Q_{up} are the stream discharge at the downstream and upstream ends of the reach, respectively, Q_{trib} is the inflow from any tributary in the reach, L is the length of the reach, and w is the average width of the stream channel. Solute flux in mmol m^{-2} day^{-1} from groundwater into a stream reach was calculated for NO_3^- and Cl^- as:

$$f_C = \frac{Q_{down}C_{down} - Q_{up}C_{up} - \sum Q_{trib}C_{trib}}{wL} \quad (1.2)$$

where C is the concentration of the solute of interest at the location indicated by the subscript. Flow-weighted solute concentration in the groundwater (C_{gw}) discharging into the reach was calculated for NO_3^- and Cl^- as:

$$C_{gw} = \frac{Q_{down}C_{down} - Q_{up}C_{up} - \sum Q_{trib}C_{trib}}{Q_{down} - Q_{up} - \sum Q_{trib}} \quad (1.3)$$

1.5.3. Streambed point approach: methods

The point sampling approach involved measuring vertical hydraulic conductivity (K), vertical head gradient (J), and streambed groundwater chemistry at streambed points arranged in transects across the channel, allowing calculation of water flux ($v = KJ$) and solute flux ($f = vC$, where C = solute concentration) through the streambed at each point. In this study all the measured point fluxes were from groundwater to stream, though fluxes in the opposite direction can also be quantified. The data allowed calculation of flow-weighted mean NO_3^- concentrations (where the weighting was by v).

The July 2012 campaign used eight 5-point transects spaced roughly 6.5 meters apart between 466 m and 524 m (a 58m reach), while in March 2013 the six 5-point transects spanned a distance of over 2.2 km (Fig. 1.1C, Table 1.5). Each transect consisted of five sampling locations, labeled as right bank (RB), right (R), center (C), left (L), and left bank (LB) and defined with respect to an observer facing downstream (making left the north side of the channel). The five locations were spaced out in a roughly even pattern across the stream width (the stream surface was about 6.5 m wide on average).

1.5.4. Streambed blanket approach: methods

Each rectangular streambed blanket (71 cm x 107 cm) covered about 0.76 m^2 of streambed. Metal flanges attached to the edges of the blankets were inserted about 5 cm into the streambed to limit stream water leakage and shallow hyporheic flow into the blanket. Blankets were made of flexible rubber (Hypalon©) and lined with stainless steel foil to prevent sorption of CFCs (age-dating tracers sampled during the study) to the rubber. Each blanket had a central PVC port through which the groundwater captured by the blanket discharged (Solder, 2014). The blankets were deployed in 5-blanket transects across the channel, with each transect almost fully covering the streambed from one bank to the other. Blankets had very low profiles above the streambed (about 3 cm or less) and caused no visible disturbance to stream flow. Blankets were allowed to equilibrate at least 8 hours after installation before discharge measurement or sampling. A dilution flow meter was attached to the PVC port in each blanket and used to determine groundwater discharge from each blanket. Groundwater samples were collected through a narrow (1/4 inch) copper tube inserted into the same port used for the flow meter (Solder, 2014).

In July 2012 blankets were deployed along two transects (481 m and 516 m), which facilitated direct comparisons (n = 10) with immediately adjacent point measurements at the same transects. In March 2013 only one transect had both blanket and point sampling (715 m) but the blankets were left in place throughout the entire campaign (3 days), which allowed repeated measurement and more detailed investigation into groundwater – surface water

exchange. During the March 2013 campaign time-series Br^- samples were collected from the blanket and depth profiles of Br^- were collected from the sediments below blankets using a USGS MINIPOINT sampler (Table 1.5). In both July and March at least one blanket was not sampled for one or more analytes due to extremely low flows from the blanket.

Water samples from some blankets contained Br^- from the NaBr injection, indicating some mixing of groundwater and surface water in the streambed beneath those blankets. Given our main objective of estimating fluxes out of the groundwater system (as opposed to total fluxes, groundwater plus surface water, through the top of the streambed), a simple mixing model based on [Br^-] was used to separate groundwater from surface water in these samples and thereby estimate the NO_3^- and Cl^- concentrations in the groundwater. The fraction of groundwater in blanket outflow was determined as:

$$F_{gw} = \frac{C_{blanket} - C_{sw}}{C_{gw} - C_{sw}} \tag{1.4}$$

where F is the groundwater fraction in the water sample collected from the streambed blanket, C is the concentration of bromide, and the subscripts *sw*, *gw*, and *blanket* indicate stream water (collected nearest the location and time of blanket sampling), groundwater, and the blanket sample, respectively. Groundwater [Br^-] from 25 cm depth MINIPOINT samples (0.3 µM, collected prior to Br^- injection) was used for C_{gw} in March 2013. In July 2012, blanket corrections were much less sensitive to C_{gw} (taken as zero in July 2012 because no pre-injection groundwater samples were collected) because C_{sw} was much higher than in March 2013 (50 µM versus 11 µM). With the groundwater fraction known from equation 1.4, the groundwater concentration for a solute of interest (e.g., nitrate) was calculated as:

$$C_{gw} = \frac{C_{blanket} - \left(1 - F_{gw}\right) C_{sw}}{F_{gw}} \tag{1.5}$$

Results derived from equation 1.5 are referred to in this paper as “corrected blanket” concentrations. Corrected solute fluxes were calculated by multiplying corrected blanket C_{gw}

by corrected blanket discharge $v_{gw} = v_{blanket}F_{gw}$, where v_{gw} is the flux from the blanket attributed to groundwater and $v_{blanket}$ is the raw blanket flux, calculated as the volumetric discharge measured from the blanket divided by the streambed area covered by the blanket. Corrected solute fluxes represent solute fluxes exiting the surficial aquifer, while uncorrected solute fluxes are total fluxes (possibly a mixture of groundwater and surface water) across the groundwater – surface water interface.

1.5.5. Sample treatment and analysis

Stream water samples were filtered (0.45 μM) in the field and stored in 20 mL LDPE vials at ambient temperature. Analysis of Br^- was by ion chromatography at the Utah USGS Water Science Center in Salt Lake City, with estimated 3% analytical uncertainty. $[NO_3^-]$ was determined for July 2012 stream water samples by ion chromatography and those $[NO_3^-]$ were ultimately used in our reach mass-balance calculations for July 2012 (Appendix A). Stream water and groundwater samples from July 2012 were also analyzed for Cl^- by ion chromatography at North Carolina State University.

With the exception of July 2012 stream water samples, all TDN, NO_3^- (NO_3^- + NO_2^-), and NH_4^+ samples were filtered (0.70 μM nominal, glass fiber filter), stored in 20 mL LDPE bottles, and kept on ice in the field or refrigerated. Samples to be analyzed for NO_3^- and NH_4^+ were acidified to pH<2 using H_2SO_4. Analysis for NO_3^- and NH_4^+ was by cadmium reduction colorimetric method in the North Carolina State University Soil Science Environmental and Agricultural Testing Service laboratory, with estimated 5% analytical uncertainty. TDN analysis was by standard methods (Merriam et al., 1996) in the same laboratory.

Water samples for noble gas analysis (Xe, Kr, Ar, Ne, He) were collected in copper tubes (Aeschbach-Hertig and Solomon 2013) and analyzed by mass spectrometry at the University of Utah Noble Gas Laboratory in Salt Lake City. Estimated uncertainty due to analytical and sampling methods for July 2012 copper tube ("Cu tube") samples were about 2% for He, 3%

for Ne and Ar, and 5% for Kr and Xe. N_2 was also measured from Cu tube samples, with estimated uncertainty of about 15%. Uncertainty in concentrations for Cu tubes collected in March 2013 and from nearby wells (June 2013, see Fig. 1.1B) was about 3% for Ne, Ar, Kr, and Xe.

For a subset of the locations where Cu tube samples were collected, dissolved gas samples were also collected in glass septum bottles (no head space, sealed with rubber stoppers), stored on ice or refrigerated, and analyzed for CH_4, CO_2, N_2, O_2, and Ar at the USGS CFC Lab in Reston, VA. Reported precision for gases analyzed at the USGS CFC lab is less than 1% (reported precision divided by mean gas concentration in our July 2012 samples) for all gases except CH_4, for which it is about 2% (http://water.usgs.gov/lab/dissolved-gas/lab/analytical_procedures/).

Field measurements of groundwater temperature, dissolved oxygen (DO), pH, and conductivity were made at each streambed sampling point using a YSI ProPlus Multiparameter Meter with flow cell prior to groundwater sample collection. Groundwater temperature was also measured at each point measurement location using a bimetal thermometer inserted 36 cm into the streambed.

1.5.6. Nitrate loss by denitrification in groundwater

Excess dissolved N_2 in groundwater, defined as the difference between total N_2 and atmospherically-derived N_2, was used to estimate the amount of N_2 from denitrification, [N_2-den] (Green et al., 2008; Pucket et al., 2008; Tesoriero et al., 2000, 2005; Böhlke et al., 2002, 2007; Böhlke and Denver 1995; Smith et al., 1991; Modica et al., 1998; Vogel et al., 1981). Determination of N_2-den was based on (1) total [N_2] in groundwater, (2) concentrations of noble gases in groundwater and (3) estimated recharge temperature.

Atmospherically derived dissolved gas concentrations and the temperature of groundwater at recharge (i.e., at the water table) were modeled by fitting the Closed-Equilibrium (CE) model to observed gas concentrations from two nearby well nests (Fig.

1.1B; Appendix B) (Aeschbach-Hertig et al., 2000, 2008). The amount of atmospherically-derived N_2 in point and blanket samples (corrected and uncorrected blanket data) was then determined by fitting the CE model to measured Ar and Ne concentrations, assuming that recharge conditions matched the modeled conditions from the wells (Appendix B). Xe and Kr were not used in the determination of [N_2-den] for point and blanket samples because those gases were injected into the stream as tracers, greatly elevating their concentrations in the stream and potentially in blanket samples.

Given estimates of [N_2-den], the percent of recharge nitrate (nitrate entering the groundwater system with recharge) that was removed by denitrification prior to discharge to West Bear Creek (E_{den}) was calculated for each groundwater sample as:

$$E_{den} = \frac{2[N_2 - \text{den}]}{\left[NO_3^-\right] + 2[N_2 - \text{den}]} = 1 - \frac{\left[NO_3^-\right]}{\left[NO_3^-\right]^0} \tag{1.6}$$

where, on a molar basis, $[NO_3^-]^0 = [NO_3^-] + 2[N_2\text{-den}]$, assuming production of N_2O from denitrification was negligible compared to the production of N_2 (e.g., Fox et al., 2014) and that nitrification and ammonification in groundwater were negligible.

1.5.7. Uncertainty estimates

For reach mass-balance results, analytical uncertainties (e.g., uncertainty in Br^- or NO_3^- concentration) and measurement uncertainties (e.g., variability in stream width measurements made during July and March campaigns) were propagated through equations 1.1-1.3 using standard methods (e.g., Taylor 1997; Kirkup and Frenkel 2006) to determine uncertainties at the 95% confidence level for v, $[NO_3^-]_{FWM}$, and f_{NO3}.

Uncertainties in individual blanket estimates of groundwater flux are found in Solder (2014). We used the reported coefficient of variation (CoV) from Monte Carlo simulations (Appendix J of Solder (2014)) to estimate 95% confidence intervals (2CoV) for blanket

discharge. Uncertainties for individual point measurements were calculated according to Kennedy et al. (2007) and Genereux et al. (2008).

The variability in N concentrations and mean fluxes for point and blanket approaches were calculated as the standard error (SE = $\sigma/ n^{0.5}$) multiplied by the Student's t statistic for n-1 degrees of freedom at the 95% confidence level. These confidence intervals are effectively controlled by the range of concentrations or fluxes over the spatial area where samples were collected (i.e., they reflect mainly spatial variability, rather than the type of uncertainty from measurement error). Uncertainty in flow-weighted mean concentrations ($[NO_3^-]_{FWM}$) for points and blankets were determined by propagating error through the equation $[NO_3^-]_{FWM} = \Sigma(f_{NO3}/\Sigma v)$. The use of the mixing model (equation 1.4) was assumed to have negligible effect on corrected blanket FWM uncertainty.

Uncertainty in [N_2-den] from point and blanket data was estimated from [N_2] modeling using upper and lower limits for recharge temperature and recharge gas concentrations (additional details are found in Appendix B).

1.6. Results and Discussion

1.6.1. Overview

Streambed point, streambed blanket, and reach mass-balance results were compared for the two field campaigns (Table 1.6) to answer the question of whether the approaches gave fundamentally different pictures of the fate of agricultural nitrate in the surficial aquifer. We focus primarily on the following comparisons:

1. Streambed blanket results (corrected and uncorrected, see equations 1.4-1.5) versus results from streambed points (from all points, and the subset immediately adjacent to blankets).

2. Reach mass-balance estimates for the full 2.5 km study reach in West Bear Creek, and for smaller reaches, versus results for streambed points and blankets. In July 2012, the "small reach" was a 200 m section of West Bear Creek that contained the 58 m reach with the eight

point transects and two blanket transects. For March 2013, the 2.5 km large reach was split into three small reaches (200-1000 m, 1000-1800 m, and 1800-2700 m), each of which contained two streambed point transects; the 200-1000 m reach also contained the single March 2013 blanket transect.

Table 1.6. Summary of water and nitrate fluxes and nitrate concentrations, July 2012 and March 2013 field campaigns.

Variable[a,b]	Reach Mass-Balance [c,d]		Blankets[g]		Points		
	Small reach(es)[e]	2.5 km reach[f]	Uncor-rected	Cor-rected	All points	Points at blankets	Points in small reach(es)[e]
			July 2012				
v	0.36[h]	0.20[i]	0.1	0.07	0.35	0.63	0.35
$[NO_3^-]_{FWM}$	451[h]	209[i]	447	561	654	690	654
$[NO_3^-]$	-	-	530	619	808	818	808
f_{NO3}	163[h]	48[i]	44	40	231	434	231
			March 2013				
v	0.33[j], 0.66[k], -1.0[l]	-0.06	0.49[m]	0.23[m]	0.40	0.31	0.39, 0.61, 0.20
$[NO_3^-]_{FWM}$	430[j], 323[k], n.a[l]	n.a.	235	295[n]	142	450[o]	419, 6.2, 17
$[NO_3^-]$	-	-	238	312[n]	210	282[o]	485, 122, 24
f_{NO3}	142[j], 212[k], -734[l]	-151	145	84[n]	57	131[o]	163, 3.8, 3.3

[a]Units are m d^{-1} for v, μM for $[NO_3^-]_{FWM}$ and $[NO_3^-]$, and mmol m^{-2} d^{-1} for f_{NO3}.

[b]For v_{mean}, $[NO_3^-]_{FWM}$, $[NO_3^-]$, and f_{NO3}, n = 39 for July 2012 point samples and n= 30 for March 2013 point samples. Subscript "FWM" indicates mean value weighted by groundwater discharge, while "mean" indicates an un-weighted mean value.

[c]Stream width used for calculations was 6.5m, based on mean width at July 2012 point transects (6.4 m, n = 8, 1σ = 0.5 m) and March 2013 measurements at flow meter locations and point transects (6.5 m, n = 14, 1σ = 0.85 m).

[d]Stream discharge used in calculations was from Br^- data collected on July 18, 2012, or March 15, 2013

[e]Small reach in July 2012 was 400-600m, three small reaches in March 2013 were from 200 m to 1000 m, 1000 m to1800 m, and 1800 m to 2700 m.

[f]Large reach for July 2012 and March 2013 was from 200 m to 2700 m.

[g]"Raw" = uncorrected; "Corr." = corrected blanket; discharge and N concentrations were corrected using equations 1.4 and 1.5. Surface water [NO_3^-] collected at 481 m and 516 m on July 18, 2012, and at 715 m on March 14, 2013, were used in corrections. Surface water fractions for blankets at 516R and 481L were slightly negative (-2% and -6%, respectively) so those samples were assumed to be solely groundwater.

[h]Based on stream discharge and (where applicable) stream water [NO_3^-] at 400 m and 600 m.

[i]Based on stream discharge at 200 m and 2700 m and tributary estimates of 4.8 L/sec ([NO_3^-] = 653 μM) and 5.2 L/sec ([NO_3^-] = 425 μM) for Trib 1 and Trib 5, respectively

[j]Estimate for 200-1000 m reach based on 12.5 L sec^{-1} tributary inflow, apportioned based on field observations as 10 L sec^{-1} from Trib 1 (682 μM nitrate) and 2.5 L sec^{-1} inflow from Trib 2 (432 μM)

[k]Estimate for 1000-1800m, with downstream [NO_3^-] based on mean [NO_3^-] of two stream samples collected at 1800 m at the same time. Trib 3 discharge was estimated at 15.7 L/sec with [NO_3^-] = 31 μM.

[l]Estimate for 1800-2700 m. Calculation of groundwater [NO_3^-] was not possible because of negative groundwater flux. Trib 4 and Trib 5 estimates were 47.7 L/sec (1074 μM) and 29.3 L/sec (60 μM), respectively.

[m]Uncorrected blanket groundwater flux for March 2013 based on n = 5 blankets, corrected groundwater flux is for n = 4 corrected blankets, plus one low flow blanket that was not corrected because no Br^- sample was available.

[n]Corrected blanket values for groundwater [NO_3^-] are based on n = 4 blankets because no LB blanket samples were collected.

[o]Point means at blanket sites are based on RB through L point measurements. LB point was excluded because no blanket samples were collected.

1.6.2. Groundwater flux

In July 2012 Br^- dilution showed groundwater discharge (*v*) to the stream in the 2.5 km reach was 0.20 m/day (Table 1.6). Mean *v* from points was 0.35 m/day (58 m reach, n = 39), in good agreement with Br^- dilution data from the 200 m reach containing the point reach (0.36 m/day). July 2012 corrected blanket discharge was low (0.07 m/day, n = 10) compared to points and Br^- dilution, especially compared to points adjacent to blankets (0.63 m/day). Blanket results were in better agreement with points in March 2013, as discussed below.

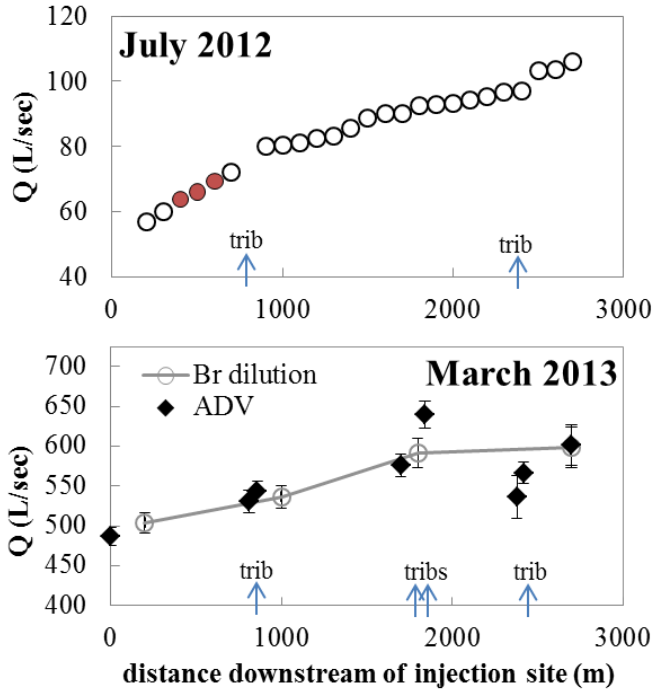

Figure 1.3. Stream discharge in July 2012 and March 2013. The locations of significant tributary inflows are indicated by arrows labeled "trib". Open circles in both figures show stream discharge calculated by Br^- dilution. Filled circles in July 2012 highlight the 200 m small reach containing the 58 m point and blanket reach. Black diamonds in March 2013 show results of ADV stream discharge measurements made the same day as Br^- sampling. Error bars show 95% confidence intervals. Uncertainties are not shown for July 2012 because the confidence interval was approximately the same size as symbols in figure.

Stream water samples for Br^- analysis were collected at 100 m resolution in both July 2012 and March 2013, but in March 2013 v was so small relative to stream discharge that $[Br^-]$ did not decrease monotonically in the downstream direction. That is, the Br^- dilution by groundwater inflow over a distance of 100 m would cause a change of about 0.1 µM in $[Br^-]$, which is smaller than the random variation of about 0.3 µM in plateau $[Br^-]$ at stream autosampler sites, indicating that 100 m was too small a spatial scale for investigation of v by Br^- dilution in March 2013 (under these conditions, point or blanket approaches are needed to look at groundwater seepage at spatial scales less than a few hundred meters). We calculated

stream discharge in March 2013 based on stream water [Br^-] from the autosampler sites, resulting in a coarser (800-900 m) spatial resolution (Fig. 1.3).

Br^- dilution in March 2013 showed an increase in stream discharge over the 200 – 2700 m reach. However, the sum of tributary inflow estimates over the 2.5 km reach was roughly equal to the change in stream discharge from Br^- dilution, which suggested no net groundwater input at this scale (200 – 2700 m). The tributary estimates in the 1800 – 2700 m section (based on the difference between ADV stream discharge estimates taken above and below tributaries) were large and likely had high uncertainty, given the disparity between the ADV measurements and the stream discharge from Br^- dilution (Fig. 1.3). There seemed to be a loss of water from the 1800-2700 m reach, perhaps in association with an agricultural withdrawal. Thus, we focus primarily on comparisons in the upper 2/3 of the 2.5 km reach, 200 m to 1800 m. Groundwater discharge to the stream in the 200-1000 m and 1000-1800 m reaches was very similar between Br^- dilution (0.33 and 0.66 m/day, respectively) and point measurements (0.39 and 0.61 m/day, respectively, 10 points in each reach; Table 1.6). March 2013 corrected blankets showed v = 0.23 m/day and were located in the reach where Br^- dilution and point measurements gave v estimates of 0.33 m/day and 0.39 m/day, respectively; the points adjacent to blankets gave v = 0.31 m/day, a closer agreement between points and blankets than in July 2012.

Stream discharge 200 m downstream of the tracer injection site was about 10x larger in March 2013 than in July 2012 (Fig. 1.3), but v was fairly similar between the two campaigns (Table 1.6; Fig. 1.4), except in the lower part of the study reach. Thus, stream discharge increased by mechanisms not directly related to groundwater flux through the bed of the main channel, such as shallow lateral discharge to small ditches and other tributaries, probably increased output from small reservoirs (Fig. 1.1B), and perhaps expanding seepage faces on the steep banks. This is relevant to understanding some of the $[NO_3^-]_{FWM}$ and f_{NO3} data.

Both point and blanket results showed higher groundwater flux toward the center of the stream and lower flux near the stream banks (Fig. 1.4), though the greater number of point than blanket observations (69 vs. 15) gives more statistical confidence in the point trend. This kind of observation (lateral variation in groundwater flux through the streambed) obviously cannot be made using the reach mass-balance approach. Uncorrected blanket discharge gives a total water flux (groundwater plus stream water) across the top of the streambed (e.g., 0.50 m/day in March 2013), while corrected blanket discharges are an estimate of groundwater flux only (e.g., 0.23 m/day in March 2013), an important distinction in studies of aquifer discharge or hyporheic zone processes.

While points and blankets both showed higher *v* near the center of the channel, *v* estimates from adjacent points and blankets were not highly correlated (for blanket *v* vs. point *v*, slope = -0.06, $R^2 = 0.04$, p = 0.49, n =15). In a similar previous comparison of *v* from point measurements and more traditional seepage meters in West Bear Creek, Kennedy et al. (2010) found slope = 0.24, $R^2 = 0.27$, with n = 53 (no p-value was reported). Kennedy et al. (2010) found the ratio of mean *v* from seepage meters to mean *v* from points was about 0.7, much higher than our blanket-to-point ratio in July 2012 (about 0.1 and 0.2 for corrected and uncorrected blankets, respectively) but the same as the blanket to point ratio for March 2013 (0.7), suggesting that the efficiency of the blankets in March was similar to traditional seepage meters. The exact reason for the closer point-blanket agreement in March 2013 compared to July 2012 is not clear (Solder 2014), but March 2013 results suggest the streambed blankets may have potential as devices for accurately measuring groundwater discharge.

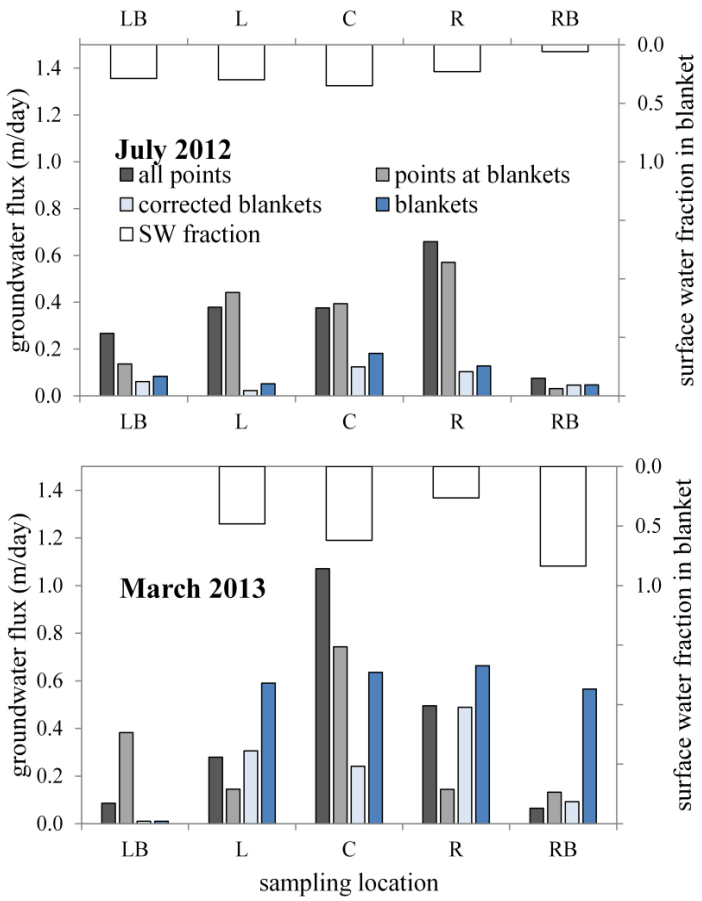

Figure 1.4. Lateral (across the channel) variability in groundwater flux from point and blanket measurements in July 2012 and March 2013. Results are arranged from left bank to right bank, from the perspective of someone looking downstream. Surface water (SW) fractions were calculated according to equation 1.4.

1.6.3. Flow-weighted mean nitrate concentration in aquifer discharge

All three sampling approaches showed groundwater $[NO_3^-]$ elevated well above the estimated 71 µM background $[NO_3^-]$ for undeveloped areas in the U.S. (Dubrovsky et al., 2010). In general, stream water $[NO_3^-]$ increased in the downstream direction, particularly upstream of about 600 m (Fig. 1.5), suggesting an influx of nitrate via groundwater during both July 2012 and March 2013. We focus here on flow-weighted nitrate concentrations in the groundwater discharging to West Bear Creek ($[NO_3^-]_{FWM}$); un-weighted groundwater $[NO_3^-]$ from points and blankets generally followed the same trends (Table 1.6). Groundwater concentrations from the reach mass-balance approach (equation 1.3) are inherently flow-weighted values.

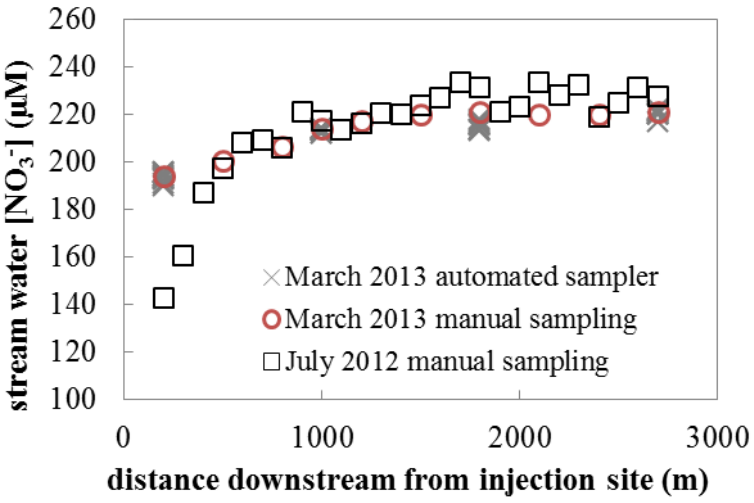

Figure 1.5. Stream water nitrate concentrations for July 2012 and March 2013 as a function of distance downstream of the tracer injection site. Automated sampler symbols show the range of concentrations over roughly a 12-hour period (n = 11-12).

Results from July 2012 showed $[NO_3^-]_{FWM}$ decreased with increasing sampling integration scale and potential for influence by channel processes: $[NO_3^-]_{FWM}$ from point measurements was highest (654 ± 18 μM, sampled at depth in the streambed), followed by blankets (561 ± 73 μM, corrected for stream water intrusion) and reach mass-balance for the 200 m reach containing the point and blanket measurements (451 ± 316 μM). The point sampling 31-36 cm deep in the streambed reached below the thin hyporheic zone (generally <10 cm deep, based on our vertical profiles in streambed, Section 1.6.4) to yield groundwater not affected by channel nitrate retention (uptake or denitrification). Samples collected from the seepage blankets represented groundwater that had interacted with the hyporheic zone and the surface of the streambed beneath the blanket, both potential sites of denitrification in biofilms. $[NO_3^-]_{FWM}$ from reach mass-balance (equation 1.3) represents groundwater that is fully mixed with the stream water and subject to the full suite of possible in-stream nitrate retention sites (hyporheic and in-channel). The trend in $[NO_3^-]_{FWM}$ suggests nitrate retention in the streambed and channel, and contrasts with the lack of a trend in the flow-weighted concentrations of the conservative ion Cl^- (315-373 μM across the three sampling approaches, Fig. 1.6).

In March 2013 $[NO_3^-]_{FWM}$ from corrected blanket data (295 μM) was lower than for adjacent points (450 μM), consistent with the trend in July 2012. March 2013 $[NO_3^-]_{FWM}$ from reach mass-balance was very similar to that from points for the 200-1000 m reach (430 versus 419 μM) and higher than that from points for the 1000-1800m reach (325 versus 6.2 μM); $[NO_3^-]_{FWM}$ could not be determined by reach mass-balance for the losing reach, 1800–2700 m (Table 1.6). Overall, July 2012 and March 2013 results suggest that the streambed point approach gives the best picture of $[NO_3^-]_{FWM}$ in groundwater discharged from the aquifer without the effects of channel or hyporheic nitrate retention. The reach mass-balance results are more complex, showing lower $[NO_3^-]_{FWM}$ during low flow (July 2012), likely due to channel nitrate retention, while during high flow (March 2013) $[NO_3^-]_{FWM}$ was equal to or greater than that from point data, likely due to increased shallow, transient lateral fluxes of

nitrate (to seepage faces on stream banks and possibly small un-sampled ephemeral tributaries) that were captured by reach mass-balance but not the streambed point measurements.

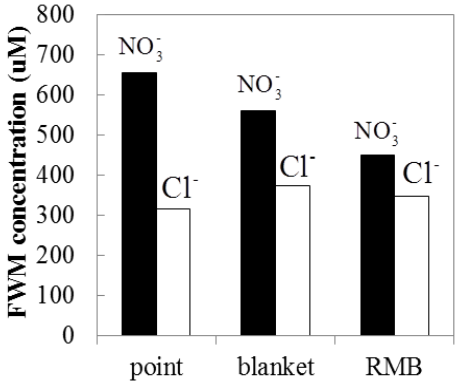

Figure 1.6. Flow-weighted mean concentrations of nitrate and chloride for three sampling approaches used in July 2012. Reach mass-balance (RMB) concentrations were based on a 200 m reach containing the 58 m reach where streambed points and blankets were sampled.

1.6.4. Nitrate flux from the aquifer

During low flow in July 2012, mean f_{NO3} of 231 mmol m^{-2} d^{-1} from point data was higher than the f_{NO3} estimate from reach mass-balance (163 mmol m^{-2} d^{-1}) in the 200 m reach that contained the points and blankets (Fig. 1.7). The difference (68 mmol m^{-2} d^{-1}) suggests nitrate retention in the stream at a rate comparable to published estimates for in-channel denitrification (Böhlke et al., 2004; Laursen and Seitzinger 2005; Birgand et al., 2007) and nitrate retention in stream sediments and channels (e.g., Stelzer et al., 2011a). Groundwater f_{NO3} from corrected blanket data was lower (40 mmol m^{-2} d^{-1}), most likely due to low estimates of v but likely also due to nitrate retention in/on streambed sediments; blanket

$[NO_3^-]$ was typically lower than point $[NO_3^-]$, even after correction for surface water in blankets (Fig. 1.8, 1.9), and some vertical profiles of streambed chemistry suggested denitrification in the streambed (Fig. 1.9). The suggestion of nitrate retention in the shallow streambed sediments is also consistent with results from March 2013 where f_{NO3} from corrected blanket data was 47 mmol m^{-2} d^{-1} lower than f_{NO3} at adjacent points (84 vs. 131 mmol m^{-2} d^{-1}, Table 1.6).

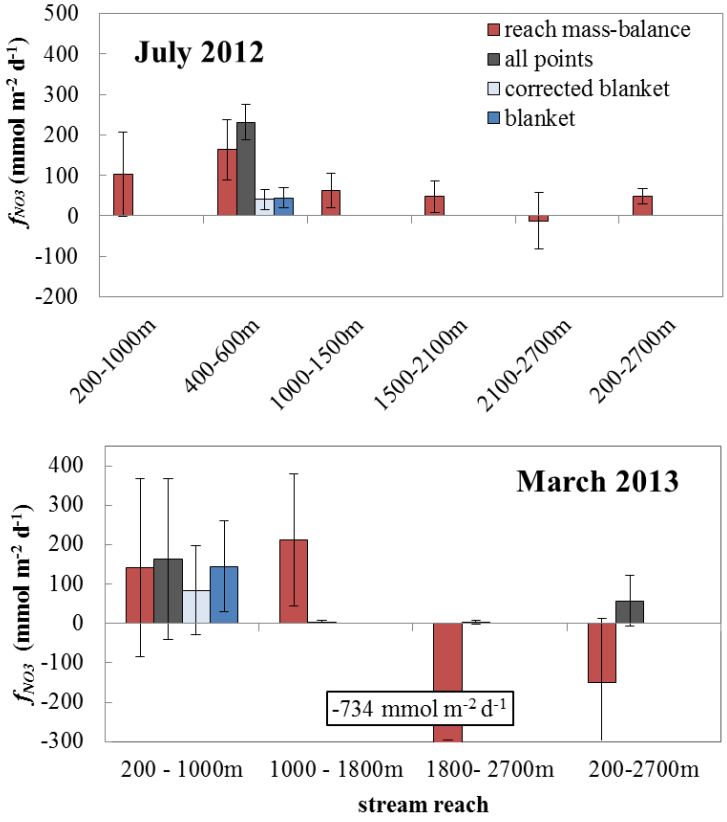

Figure 1.7. Nitrate fluxes from groundwater to West Bear Creek for (A) July 2012 and (B) March 2013. The large negative (stream-to-groundwater) reach mass-balance flux for 1800-2700m is indicated by a label and the bar is not shown to full scale.

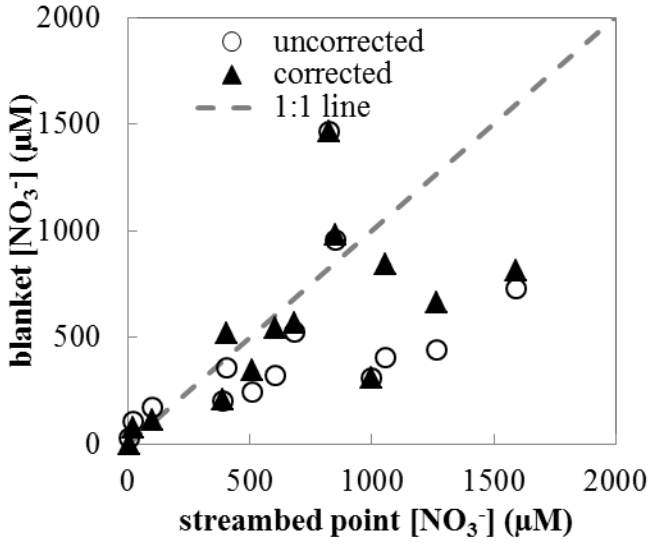

Figure 1.8. Streambed blanket $[NO_3^-]$ (uncorrected and corrected) versus $[NO_3^-]$ from streambed points.

Given the potential for nitrate retention in the streambed and channel, f_{NO3} from reach mass-balance is likely a lower limit on the actual nitrate flux leaving the aquifer, and we believe this was the case in July 2012 during low flow. During higher stream discharge conditions in March 2013, however, f_{NO3} from reach mass-balance was only slightly lower than from points for the 200-1000 m reach (142 versus 163 mmol m^{-2} d^{-1}), higher for the 1000-1800 m reach (212 versus 3.8 mmol m^{-2} d^{-1}, respectively), and much lower (negative) for the 1800-2700 m reach (-151 versus 3.3 mmol m^{-2} d^{-1}, respectively). The negative flux from reach mass-balance in the 1800-2700 m reach was due to the net loss of water from the stream. The higher f_{NO3} by reach mass-balance, compared to points, for the 1000-1800 m reach and the similarity in f_{NO3} from the two approaches in the 200-1000 m reach may be partially linked to incomplete sampling of the streambed by points (e.g., no samples collected during March 2013 in the high-nitrate streambed zone sampled by points and blankets in July 2012), but also may be a result of the reach mass-balance approach integrating shallow nitrate-contaminated bank seepage (groundwater discharge through seepage faces above the

water line, [e.g., *Böhlke et al.,* 2007]). In March 2013, $[NO_3^-]$ of 1230 μM and 135 μM was observed in samples of outflow from a seepage face (1130 m location, LB) and a macropore (1700 m location, RB), respectively, indicating that some bank seepage had $[NO_3^-]$ greater than stream water $[NO_3^-]$ (220 μM).

Streambed point sampling in March 2013 showed NO_3^- was primarily entering through the left (north) side of the stream and that 1910 m and 2530 m were the only transects where significant $[NH_4^+]$ was detected (Fig. 1.10). Although $[NO_3^-]$ was above detection (>3.6 μM) for all but the 2530 m transect, the actual f_{NO3} into the stream was negligible downstream of 715 m (Fig. 1.10). Streambed maps of nitrate concentration from point sampling in July 2012 (Fig. 1.11B) show lower nitrate concentration in the center of the stream (likely linked to older, less-contaminated groundwater beneath the center of the stream, (Kennedy et al., 2009a; Gilmore et al., 2015a); the mean for center points (C) was significantly lower than the means for points in the left bank, left, and right bank positions (LB, L, and RB, respectively), based on two-tailed t test assuming unequal variance). There were particularly high nitrate concentrations along the left bank (the mean for LB was significantly different than for C, R (right), and RB), where the tree buffer was less well developed than on the right bank. In contrast, the map of nitrate flux is dominated by 3-4 hotspots of high seepage rate (Fig. 1.11C; given the large variability only L and RB had significantly different means), illustrating the importance of flow-weighting the measured concentrations. Similarly, in March 2013, the sampling location with greatest $[NO_3^-]$ (2041 μM, Fig. 1.10A) was not the location with the greatest f_{NO3} (Fig. 1.10B). Such findings are not possible with the reach mass-balance approach.

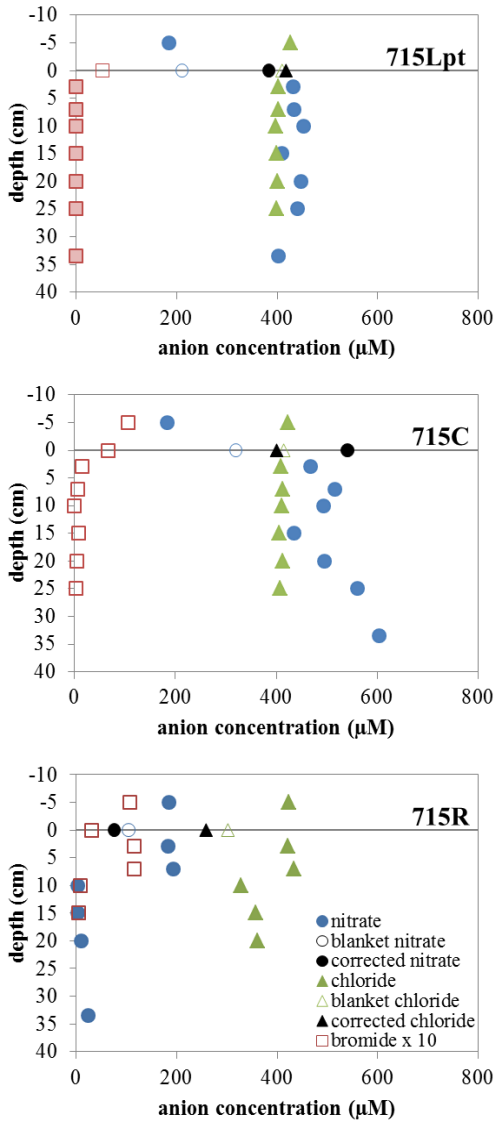

Figure 1.9. Representative vertical profiles of streambed chemistry, 715 m location in West Bear Creek. Concentrations plotted on the streambed surface (depth = 0) represent blanket samples, either uncorrected (symbols with no fill) or corrected for stream water mixing (black filled symbols). Stream water concentrations are plotted above the streambed surface (negative depth), and values plotted at 33.5cm depth are from piezomanometer samples (31-36cm screen). All other concentrations are from a USGS MINIPOINT sampler. The "pt" (pre-tracer) designation for the 715L plot indicates that the USGS MINIPOINT samples were collected prior to the Br^- stream tracer injection, so filled Br^- symbols show background [Br^-]. Note that [Br^-] has been multiplied by a factor of 10. The L, C, and R after 715 indicate left side, center, and right side of the channel.

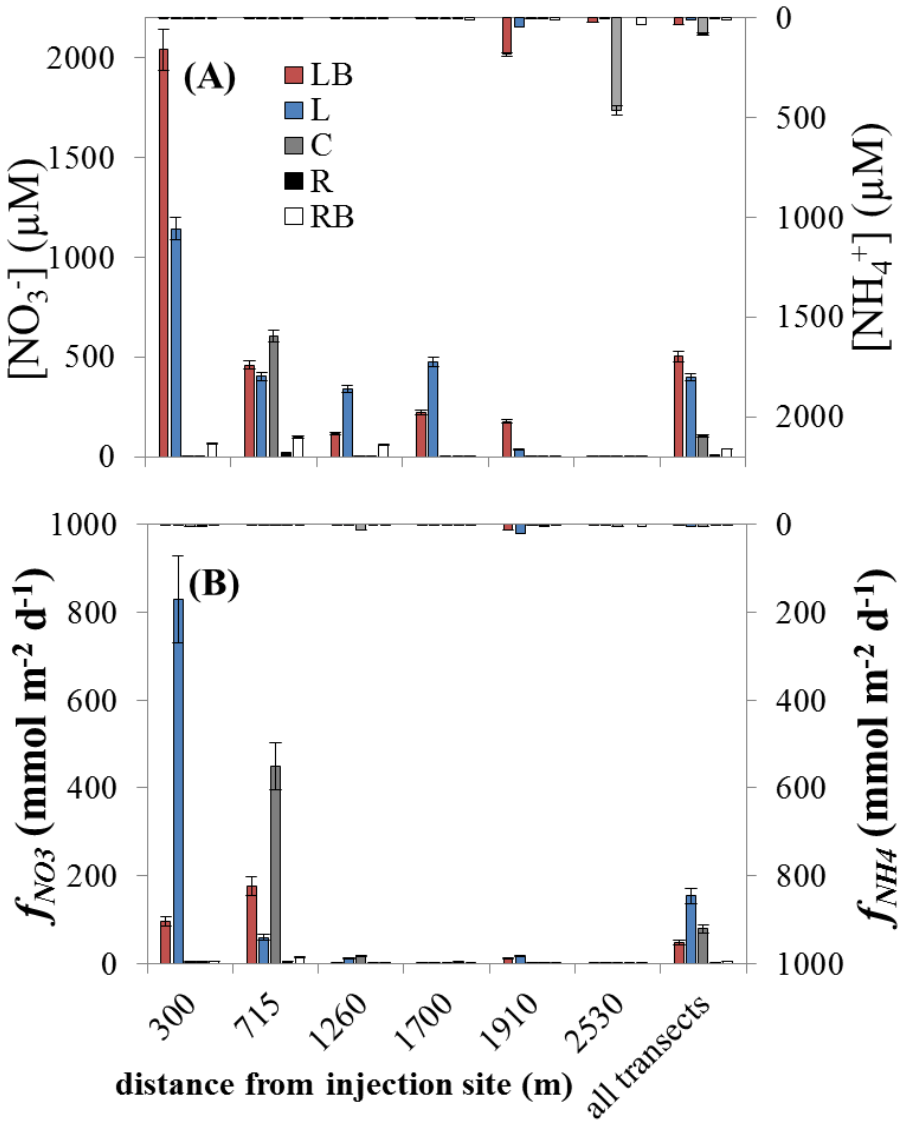

Figure 1.10. Spatial variation in groundwater nitrate and ammonium from point sampling in March 2013: (A) concentration across the stream, and (B) flux along the study reach. Error bars show 5% analytical uncertainty for concentrations and mean uncertainty (12%) for fluxes.

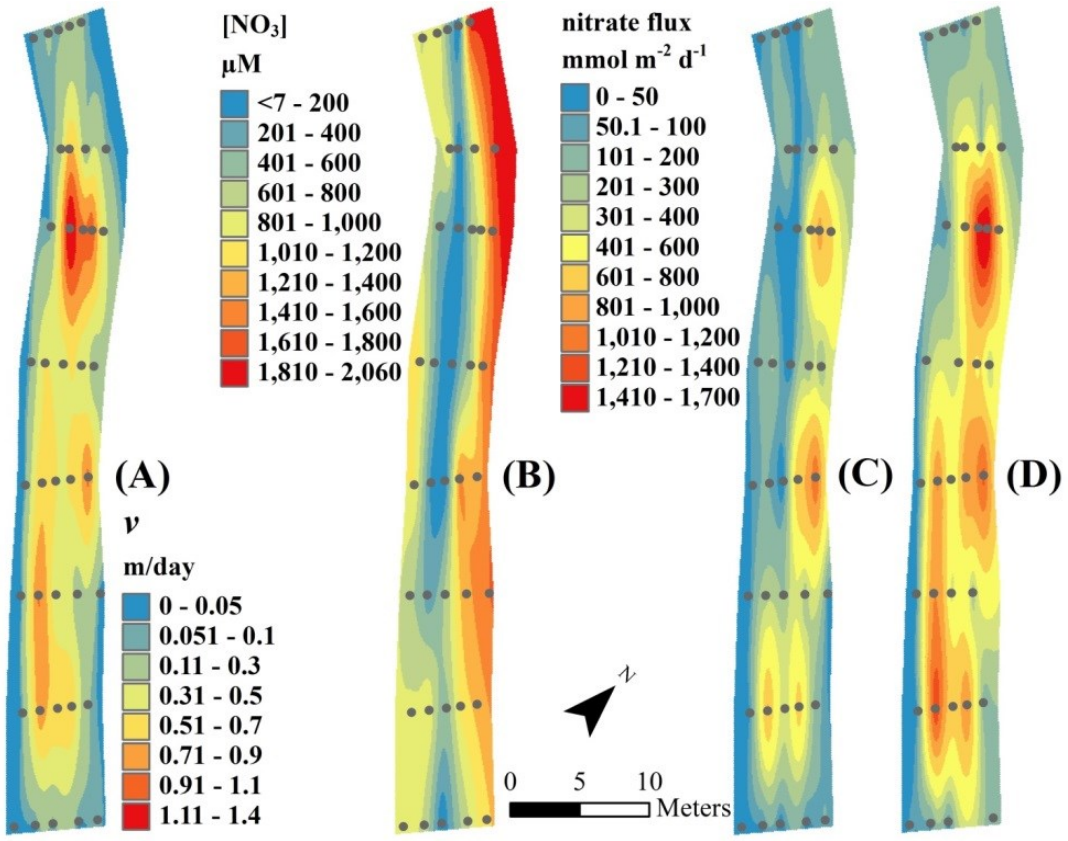

Figure 1.11. Streambed maps for 466-524 m, West Bear Creek, July 2012: (A) water flux v, (B) groundwater [NO_3^-], (C) nitrate flux f_{NO3}, and (D) f_{NO3}^0 (i.e., what nitrate flux would have been if no denitrification had occurred in the surficial aquifer). The maps were created using a multiquadric radial basis function in ESRI® ArcMap™ 10.0, with settings of 8 and 0.0 for the anisotropy and kernel parameters, respectively.

1.6.5. Amount of denitrification, and initial nitrate concentrations

1.6.5.1. Dissolved gases: degassing and excess air

[N_2-den] was estimated by modeling based on (1) measured [N_2], [Ar], and [Ne] in groundwater, and (2) an estimated recharge temperature (12.3° C) from noble gas thermometry at nearby well nests. We computed the concentrations of dissolved gases for water in equilibrium with air (WEA) at 12.3° C, which are the theoretical concentrations of

dissolved gases at recharge. In reality, many groundwater samples have gas concentrations greater than WEA, or "excess air", typically due to dissolution of trapped air bubbles near the water table (e.g., Heaton and Vogel 1981). In other cases groundwater samples have gas concentrations lower than WEA, suggesting that "degassing" has occurred. Degassing usually occurs when dissolved gases partition into bubbles in the groundwater system after production of biogenic gases such as N_2, CH_4, or CO_2 (Visser et al., 2007; Aeschbach-Hertig et al., 2008). The percent excess (or depletion) of [Ne] or [Ar] relative to WEA (denoted as ΔNe or ΔAr) were used as an indicator of the magnitude of excess air (e.g., positive ΔNe) or degassing (e.g., negative ΔNe), as discussed in the following sections.

1.6.5.2. Denitrification in the surficial aquifer: streambed-based measurements

Blanket and point samples from both July 2012 and March 2013 consistently showed that denitrification had occurred in the surficial aquifer (Table 1.7). Applying the reach mass-balance approach to stream water [N_2] showed that the method had low sensitivity to the denitrification signal associated with discharge of excess N_2 from groundwater, mainly because excess N_2 (from denitrification in the stream channel and/or in the surficial aquifer) was a very small fraction (typically <3%) of the total N_2 flux within a given stream reach and thus highly uncertain (further detail in Section 1.6.5.4 and Appendix C).

Points gave the highest estimates of the extent of denitrification (E_{den}), averaging 51% in July 2012 and 78% in March 2013. E_{den} from blanket data (36%) was lower than for all points in July 2012, but very similar to points collected adjacent to the blankets (42%). Correcting July 2012 blankets for surface water intrusion did not change the estimate of E_{den}.

Table 1.7. Summary of denitrification and recharge (initial) nitrate concentrations and fluxes from the July 2012 and March 2013 field campaigns.

Variable[a]	Blanket		Point	
	Uncorrected	Corrected[b]	All[c]	At blankets
		July 2012		
$[NO_3^-]^0_{FWM}$	802	1086	1300	1284
$[NO_3^-]^0$	908	1096	1453	1312
$f_{NO3}^{\ 0}$	79	78	477	809
$[N_2\text{-den}]$	189	238	328	247
E_{den}	36%	36%	51%	42%
		March 2013		
$[NO_3^-]^0_{FWM}$	384	660	519	783
$[NO_3^-]^0$	385	817	629	756
$f_{NO3}^{\ 0}$	237	188	235	228
$[N_2\text{-den}]$	74	253	197	237
E_{den}	39%	56%	78%	64%

[a]Units are μM for $[NO_3^-]^0_{FWM}$ and $[NO_3^-]^0$, and mmol m^{-2} d^{-1} for $f_{NO3}^{\ 0}$.

[b]Blanket discharge and N concentrations were corrected using equations 1.5 and 1.6.

[c]For July 2012 and March 2013, n = 36 and n = 26 respectively.

In March 2013 the blanket correction was more significant (largely driven by a large $[N_2$-den] estimate from the blanket near the right bank), with E_{den} equal to 39% for raw blanket results and 56% for corrected blankets; the latter was similar to E_{den} from the points adjacent to blankets (64%). With the exception of the 78% mean from March 2013 points, E_{den} was similar to the estimate of ~50% from previous work in West Bear Creek (Kennedy et al., 2009a).

Based on differences between $[NO_3^-]^0_{FWM}$ and $[NO_3^-]_{FWM}$ for July 2012, streambed point data showed nitrate declined by 646 μM in the surficial aquifer. July 2012 corrected blanket data suggested 525 μM of nitrate removed in the surficial aquifer, higher than uncorrected blankets (355 μM) and similar to the points adjacent to blankets (594 μM). In March 2013, estimates of nitrate removal were similar for all points, points at blankets, and corrected blankets (377, 333, and 365 μM, respectively), while the uncorrected blanket estimate was lower (149 μM).

1.6.5.3. Spatial patterns of denitrification

In July, most point samples (31 of 36 samples) were degassed (negative ΔNe), while in March 2013 only 14 of 29 were degassed, likely because the March sampling temperature was cooler. Lower groundwater temperatures would inhibit biogenic production of gases in the streambed and also increase gas solubility.

In July 2012, anoxic conditions were observed at the point sampling sites where degassing was most prominent, i.e., where ΔNe was most negative (Fig. 1.12A, 1.12C), suggesting biogenic production of gases (N_2, CH_4, CO_2) could have caused the degassing. Modeled $[N_{2\text{-den}}]$ for July 2012 (Fig. 1.13) followed the trends suggested by [Ne] and $[O_2]$, with more denitrification on the right side and center portion of the stream than on the left. However, $[NO_3^-]^0$ follows a different, "center-low" pattern (Fig. 1.12E), consistent with very similar recharge of nitrate into the groundwater system on both sides of the channel, an insight that could not be obtained from the reach mass-balance approach. In a companion paper we show a relationship between greater groundwater age and lower $[NO_3^-]^0$ beneath the center of the stream (Gilmore et al., 2015a), consistent with findings by Kennedy et al. (2009a). Also, elevated degassing on the right side of the stream is consistent with previous observations of lower [Ar] on the right side, interpreted in previous work (likely erroneously) as an indication of higher recharge temperature (Kennedy et al., 2009a).

Compared to streambed points, July 2012 blanket data showed similar patterns in ΔNe, $[O_2]$, and $[N_2\text{-den}]$, albeit with greater degassing (Fig. 1.12, 1.13). $[NO_3^-]^0$ varied across the channel, unlike the symmetric pattern observed with points. Corrected data in March 2013 showed degassing in two of the four blankets sampled, despite mostly oxic groundwater. One corrected estimate of $[O_2]$ was biased high (right bank, Fig. 1.12D) as the corrected concentration was roughly double the solubility equilibrium concentration at the stream temperature. Corrections to other gases (Ar, Ne, N_2) in the RB blanket gave reasonable concentrations (± 30% of the mean value for all four blankets), and the corrected estimate of $[N_2\text{-den}]$ gave a feasible value (Fig. 1.13).

Even with the much broader coverage by the point approach along West Bear Creek in March 2013, the left bank (LB) still stands out as having high $[O_2]$ and low degassing (Fig. 1.12B, D), while the lateral pattern of $[NO_3^-]^0$ (Fig. 1.12F) seems consistent with lower N use on the right (south) side of the stream, probably associated with the extensive forested areas on the right side of the stream, downstream of 1000 m (Fig. 1.1). Lateral patterns for streambed blankets in March 2013 contrast with results from points sampled throughout the 2.5 km reach (e.g., different degassing pattern, higher $[O_2]$, higher $[N_2\text{-den}]$ and $[NO_3^-]^0$ on right side compared to left side for points), but resulted in a similar pattern of denitrification ($[N_2\text{-den}]$ for RB and R > C, L, or LB) compared to points sampled adjacent to the blankets (Fig. 1.13).

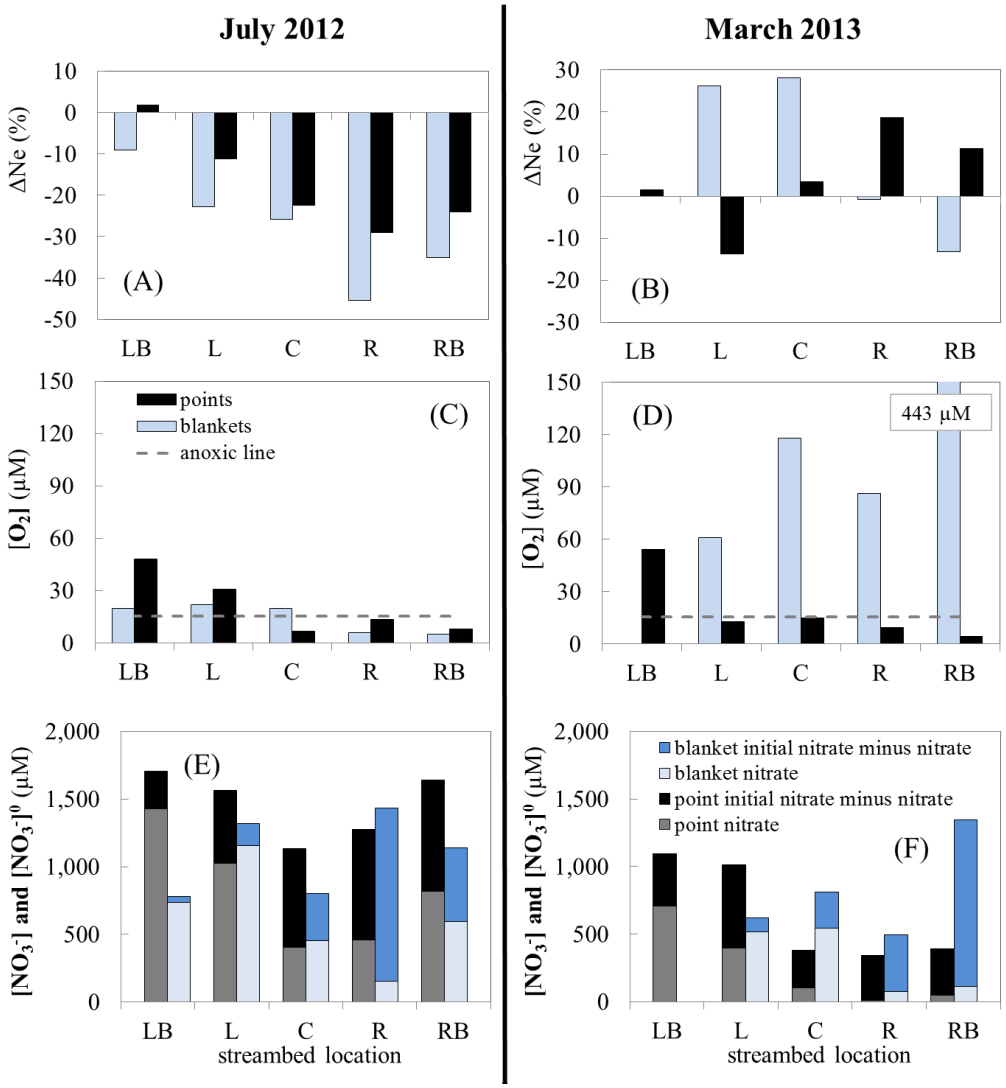

Figure 1.12. Cross-channel patterns of ΔNe, $[O_2]$, and nitrate in point and blanket samples, July 2012 and March 2013. Each black bar is the mean of 6-8 point samples (July) or 4-6 point samples (March). Grey bars in (E) and (F) show measured nitrate from point samples, and the black bars show nitrate removed in the surficial aquifer ($[NO_3^-]^0$ minus $[NO_3^-]$). Each blue bar represents 1-2 blankets (July) or one blanket (March), all corrected for surface water intrusion except for July $[O_2]$ (no stream water $[O_2]$ was measured). Light blue bars in (E) and (F) show measured nitrate from blanket samples, and the dark blue bars show nitrate removed in the surficial aquifer. The total height of each gray plus black stacked bar, and each light blue plus dark blue stacked bar, is a mean value of $[NO_3^-]^0$ at the given location (RB, R, etc.) No LB blanket samples were collected in March 2013. The dashed line in (C) and (D) shows the typical cutoff for anoxic conditions (16 µM).

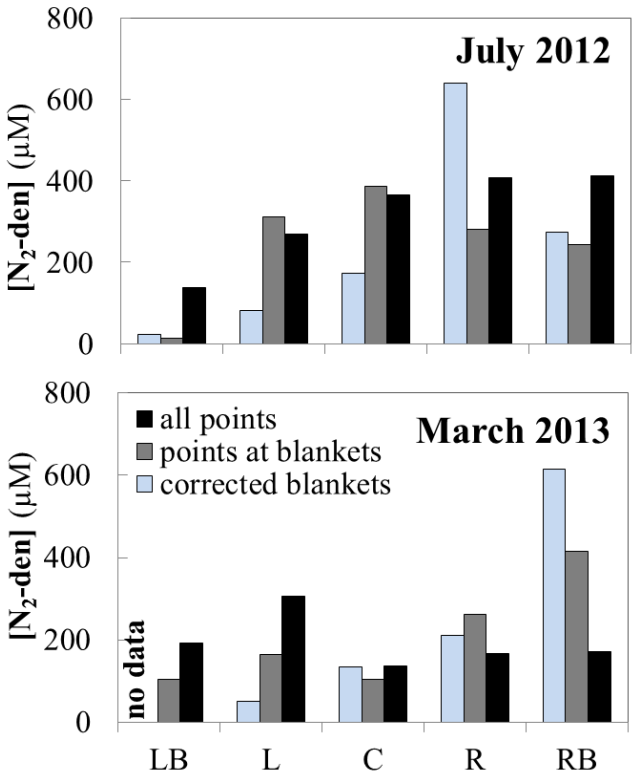

Figure 1.13. Cross-channel patterns of [N_2-den] from point and blanket samples, July 2012 and March 2013. Each black bar is the mean of 6-8 point samples (July) or 4-6 point samples (March). Grey bars show [N_2-den] from points sampled adjacent to blankets. "No data" indicates where no blanket sample was collected.

1.6.5.4. In-channel denitrification versus denitrification in the surficial aquifer

While this study focused on denitrification in the groundwater system, nitrate retention can occur as groundwater discharges through shallow streambed sediments (Sheibley et al., 2003; Gu et al., 2008; Flewelling et al., 2011; Stelzer et al., 2011a; Krause et al., 2013), and additional removal may occur in stream channels (McCutchan et al., 2003; Mullholland et al., 2004; Böhlke et al., 2004, 2009; Laursen and Seitzinger 2005; Birgand et al., 2007; Smith et al., 2008; Baulch et al., 2010). We can infer from the July 2012 and March 2013 data sets

that, relative to in-channel nitrate retention, denitrification in the surficial aquifer is an important process. Based on July 2012 point data, the 646 μM nitrate reduction in the surficial aquifer represents about half of the initial nitrate (1300 μM, Table 1.7). Roughly an additional third of the nitrate entering the streambed was removed in the streambed sediments and/or channel, based on the difference between $[NO_3^-]_{FWM}$ from streambed point (654 μM) and reach mass-balance data (451 μM, 400-600 m reach; Table 1.6). In March 2013, data showed a sequence of $[NO_3^-]^0_{FWM}$ = 783 μM and $[NO_3^-]_{FWM}$ = 450 μM from points adjacent to blankets, and $[NO_3^-]_{FWM}$ = 295 μM from blankets, suggesting almost half of the initial nitrate was removed in the surficial aquifer while roughly an additional third of the remaining nitrate was removed in the streambed.

Looking more specifically at the apparent retention of nitrate between point and adjacent blanket samples, the mean groundwater travel time from the depth of the center of the point sampling screen (33.5 cm) to the top of the streambed (33.5 cm divided by 0.35v where 0.35 is assumed streambed porosity) were in a range where a relationship between travel time and denitrification could potentially be expected (Flewelling et al., 2011). Mean travel time over this distance as determined by all point measurements was 1.3 days in July 2012, 6.9 days in March 2013; travel time for points adjacent to blankets was 0.6 day during both campaigns. However, there was no significant relationship between travel time and extent of nitrate removal between point and blanket samples, which suggests that the greater nitrate retention shown by blanket samples relative to adjacent point samples may have been mostly due to nitrate retention on the streambed surface under the streambed blankets rather than retention in the upper 33 cm or so of streambed.

At the reach mass-balance scale, excess N_2 from denitrification in either the surficial aquifer or the stream channel was generally not detected. We calculated $[N_2]$ and [Ar] in groundwater discharge to the stream (using equation 1.3 with an added term for diffusive gas exchange between the stream and the atmosphere; Appendix C). Modeling $[N_2\text{-den}]$ for reach mass-balance proved infeasible, however, given that calculated uncertainties in [Ar] were

very high (20-1073%). The uncertainty in [Ar] is critical because it is the gas used to estimate [N_2-den] (Appendix C), and typically, noble gas modeling (e.g., noble gas thermometry, [N_2-den] modeling, age-dating tracer models) is done using dissolved gas concentrations with analytical uncertainties of 5% or less. In addition, we calculated apparent production of N_2 in the stream channel (e.g., Laursen and Seitzinger (2002); McCutchan et al. (2003); Baulch et al. (2010)) using stream water [N_2] and groundwater [N_2] (from streambed point samples) in a 1-D steady-state advection-dispersion equation for the stream (Appendix C). In some cases, apparent production of N_2 by in-stream denitrification was in the range of published rates (2.4-5.2 mmol m^{-2} hr^{-1}; Baulch et al. (2010), their Appendix 1) but it was negative in other cases (-4.0 to -2.3 mmol m^{-2} hr^{-1}).

1.7. Summary and Conclusions

Collecting and using field data to assess the fate of contaminant nitrate in an aquifer is challenging, given the heterogeneity of real groundwater systems, complexity of nitrogen biogeochemistry, spatial and temporal variability in nitrate sources, site access limitations, and other factors. In the (typical) absence of a network of numerous spatially-distributed multilevel samplers or well nests, stream beds and channels (as places of groundwater discharge) may be used to access groundwater of different ages originating from different recharge locations in the watershed (e.g., left or right, upstream or downstream, and far or near, with respect to the channel).

There are relatively few field methods that can be reliably deployed for such sampling, three of which were assessed and compared in this study of an agricultural coastal plain watershed: streambed point measurements, streambed seepage blankets (a novel form of seepage meter), and reach mass-balance. The three approaches were applied during 3-4 day field campaigns in July 2012 and March 2013, under different stream flow conditions (~50 L/sec and ~500 L/sec, respectively). The variables computed to assess the fate of nitrate in the groundwater system (Section 1.5.1) also give insight into the nitrate inputs from

groundwater to surface water and initial nitrate concentration in the groundwater at the time of recharge.

Regarding groundwater flux between the aquifer and stream, rates were very similar between point sampling and the Br^- dilution work needed for reach-mass balance at low flow in July 2012: 0.35 and 0.36 m/day, respectively. At higher stream flow in March 2013, groundwater flux was again similar between point sampling and Br^- dilution for the 200-1000 m reach (0.39 and 0.33 m/day, respectively) and the 1000-1800 m reach (0.61 and 0.66 m/day, respectively), but not the more complex 1800-2700 m reach where point sampling indicated groundwater input at 0.20 m/day and Br^- dilution in conjunction with stream discharge data from an ADV suggested a losing reach (-1.0 m/day). Blanket and adjacent point data gave similar discharge in March 2013 (Table 1.6), in better agreement than in July 2012.

The comparison among the three methods was somewhat different for nitrate flux (f_{NO3}) and $[NO_3^-]_{FWM}$ than for groundwater flux. Comparison of the methods during low stream flow showed a trend related to the biogeochemistry of nitrate and high-flow comparison showed the impact of transient hydrologic flowpaths to the stream. The July 2012 trend in $[NO_3^-]_{FWM}$, points>blankets>reach mass-balance (Table 1.6), was consistent with the sampling method showing greater retention of nitrate (likely denitrification) with greater hyporheic and/or channel influence (reach mass balance sampling has the greatest engagement of hyporheic and channel processes, point sampling the least, and blankets an intermediate level). In July 2012 the situation was similar for f_{NO3} in that the estimate from point sampling was greater than that from reach mass-balance (the blanket estimate was very low because of the anomalously low groundwater discharge measured by blankets in July 2012).

However, at high flow (March 2013), the method comparisons for $[NO_3^-]_{FWM}$ and f_{NO3} were different: point sampling and reach mass-balance gave similar results for both nitrate parameters in the 200-1000 m reach, and the point estimate was much lower than the reach mass-balance estimate in the 1000-1800 m reach (Table 1.6). Why would reach mass-

balance shift from giving lower $[NO_3^-]_{FWM}$ and f_{NO3} than point sampling at low flow to giving equal or higher values at high flow? The change is likely linked to shallow lateral transient flow, discharging at high flow on banks above the waterline and into small tributaries. Such discharge is likely to be high in nitrate (water discharge at one stream bank seepage face had $[NO_3^-]$=1230 μM in March 2013). These results are also consistent with the fact that groundwater flux into the main channel between 200 m and 1000 m was nearly the same between the two sampling campaigns even though stream discharge was about 10x larger in March 2013 (transient flowpaths must explain the higher stream discharge).

Somewhat paradoxically, in the upstream part of the study area (200-1000 m), point sampling and reach mass-balance gave <u>different</u> nitrate results ($[NO_3^-]_{FWM}$ and f_{NO3}) at low flow, when they were responding to more or less the <u>same</u> inputs (groundwater flowpaths to the streambed in the main channel, no shallow lateral transient flowpaths), and <u>similar</u> results at high flow, when they were responding to <u>different</u> inputs (only groundwater flowpaths for point sampling, groundwater plus shallow lateral transient flowpaths for reach mass-balance). The different results at low flow were likely due to nitrate biogeochemistry, not simply method differences alone; these differences did not occur for Cl^- (Fig. 1.6). This underscores the importance of understanding the inherent differences in these sampling methods when investigating fluxes at the groundwater – surface water interface.

It also suggests the importance of matching the method to the question. For a study focused on nitrate flux exiting the surficial aquifer, streambed point sampling may be the best approach (least channel influence) in a streambed that allows for insertion and use of sampling probes and field permeameters (Table 1.1). Reach mass-balance may or may not give similar results to points, depending on the flow condition and potential differences between the nitrate concentration in shallow soil and $[NO_3^-]_{FWM}$. If the focus is instead on nitrate flux into the stream, point sampling in the streambed may underestimate this flux at high flow (if nitrate is high in the shallow soil) and overestimate it at low flow (if there is nitrate retention in the streambed). Reach mass-balance may represent a better approach, at

least at high flow. Our blanket sampling at low flow gave anomalously low seepage rates but in principle this method may be the best of the three for determining nitrate flux into the stream, if the flow measurement can be done more accurately as it seemed to have been in March 2013.

Of course, the complexity of N biogeochemistry is what complicates the matter; the choice among the three methods should not matter nearly as much or at all for more simple solutes (e.g., Fig. 1.6). However, the differences among the methods also present an opportunity to gather a fuller picture of the groundwater-surface water exchange of nitrate. For example, the difference in nitrate discharge rate from the groundwater system (from points) and the nitrate input rate to the stream channel (from blankets and/or reach mass-balance) was 47-68 mmol m^{-2} d^{-1} (Table 1.6), a number that falls within the published range of in-channel nitrate retention rates. Simultaneous application of point sampling with blankets, other seepage meters, or reach mass-balance may represent another approach to constraining the magnitude of channel nitrate retention rates.

Initial nitrate at recharge, $[NO_3^-]^0_{FWM}$, and percent nitrate loss by denitrification in the surficial aquifer, E_{den}, were not reliably determined at the reach mass-balance scale, but points and blankets gave similar and seemingly reliable estimates for both parameters (even when blanket nitrate flux was low because *v* was too low) (Table 1.7). About half the nitrate entering the groundwater system near West Bear Creek was lost by denitrification in the aquifer before discharge to the stream (more than half in some portions of the stream), which shows the surficial aquifer provides a major ecosystem service. Of the nitrate that reaches the stream channel, an additional third of that is lost relatively quickly in the hyporheic zone and/or after a short (~100 m) transit in the stream channel. In March 2013, the loss likely occurred in the sediments and on the streambed surface below the blankets, while in July 2012 the loss was observed over a 200 m reach (difference between points and reach mass-balance). Greater loss might be observed over greater channel lengths (a topic outside the scope of this study), but overall our results suggest about 70% of nitrate removed in total

during transit through the surficial aquifer, shallow streambed sediment, and possibly a short distance in the stream channel.

Accounting for degassing (loss of dissolved gases from groundwater due to bubble formation) was important for interpreting the amount of denitrification from both point and blanket dissolved gas data (Fig. 1.12). Potential for degassing during streambed sampling should be an important consideration in future field studies, given that streambed temperatures may be relatively warm (e.g., 20° C in July 2012), and hydrostatic pressure in the streambed is relatively low compared to pressure at the depth of well screens in most observation wells. For future studies where degassing is anticipated (e.g., in contaminated agricultural areas), it may be beneficial to collect streambed groundwater samples for dissolved gas analysis during colder weather and/or deeper in the streambed to guard against degassing.

Point and blanket approaches yield greater spatial insights than reach mass-balance, and our results suggest linkages between land use and groundwater quality (Figs. 4, 10-13). Low groundwater nitrate concentration $[NO_3^-]$ was observed in the center of the stream (Fig. 1.11), due to the presence there of older and less contaminated water likely recharged relatively far from the stream, within the watershed (Gilmore et al., 2015a). The highest $[NO_3^-]$ was on the left side of the stream where the riparian buffer was generally less well-developed, and likely reflects agricultural practices close to the stream (Fig. 1.10, 1.11). High groundwater $[NH_4^+]$ was observed only on the left or in the center of the downstream portion of the study reach, and was likely related to animal operations on the left side of that section of the stream. Low $[NO_3^-]^0$ (initial nitrate concentration in groundwater) occurred on the right side of the stream (Fig. 1.10) where forested land use dominated the recharge area (Fig. 1.1). Thus, regarding the fate of agricultural nitrate in a surficial aquifer, streambed point and potentially blanket sampling may go beyond just total rates for the two main sinks (denitrification and discharge to surface water) to allow insights with a spatial context related to land use.

A secondary question is how to arrange streambed point (or blanket) sampling, e.g., closely vs. widely spaced transects (Fig. 1.1). In West Bear Creek, using both arrangements gave both a broad picture of a range of conditions along the stream (e.g., transition from all nitrate upstream to nitrate plus ammonium downstream, different land use effects described above, etc.) and a more detailed picture of lateral variation and spatial variability in nitrate discharge and denitrification in a small area.

Overall, did the three methods (points, blankets, reach mass-balance) give a fundamentally different picture of the fate of nitrate in the surficial aquifer? Each sampling approach showed nitrate contamination in groundwater discharge, and point and blanket sampling gave very similar pictures of the amount of denitrification in the aquifer (Table 1.7), but there is a need to properly interpret differences in results among the methods based on the limitations of each method. For example, when points miss transient lateral flows above the waterline, that may be a problem if the goal is quantifying total nitrate inputs to the stream system at high flow, but a benefit if the goal is quantifying the nitrate discharge from the aquifer (it means this can be done under a wide range of flow conditions, including transient high flow after precipitation events). The three approaches are a complementary set of methods for looking at the broad picture of the fate of nitrate in the surficial aquifer (the nitrate discharge to surface water and denitrification in the aquifer), and linked surface water processes (the nitrate flux from groundwater to surface water, and nitrate retention in the channel and hyporheic zone).

REFERENCES

Aeschbach-Hertig, W., El-Gamal, H., Wieser, M., & Palcsu, L. (2008). Modeling excess air and degassing in groundwater by equilibrium partitioning with a gas phase. *Water Resour. Res.*, *44*(8), W08449. doi:10.1029/2007WR006454

Aeschbach-Hertig, W., Peeters, F., Beyerle, U., & Kipfer, R. (2000). Palaeotemperature reconstruction from noble gases in ground water taking into account equilibration with entrapped air. *Nature*, *405*(6790), 1040–1044. doi:10.1038/35016542

Aeschbach-Hertig, W., and D. K. Solomon (2013), Noble gas thermometry in groundwater hydrology, in *The Noble Gases as Geochemical Tracers*, edited by P. Burnard, pp. 81–122, Springer, Heidelberg.

Alley, W. M., Reilly, T. E., & Franke, L. O. (1999). Sustainability of Ground-Water Resources, *U.S. Geol. Surv., Circular 1186,* U.S. Geol. Surv., Reston, VA. [available at: from http://pubs.usgs.gov/circ/circ1186/.]

Angier, J. T., & McCarty, G. W. (2008). Variations in base-flow nitrate flux in a first-order stream and riparian zone1. *J. Am. Water Resour. Assoc.*, *44*(2), 367–380. doi:10.1111/j.1752-1688.2007.00153.x

Baulch, H. M., Venkiteswaran, J. J., Dillon, P. J., & Maranger, R. (2010). Revisiting the application of open-channel estimates of denitrification. *Limnol. Oceanogr.-Methods*, *8*, 202–215. doi:10.4319/lom.2010.8.202

Bennett, J. P., & Rathbun, R. E. (1972). Reaeration in open-channel flow, *PP-737,* U.S. Geol. Surv., Reston, Va. [available at: http://pubs.er.usgs.gov/publication/pp737.]

Birgand, F., Skaggs, R. W., Chescheir, G. M., & Gilliam, J. W. (2007). Nitrogen removal in streams of agricultural catchments—A literature review. *Critical Reviews in Environ. Sci. Technol.*, *37*(5), 381–487. doi:10.1080/10643380600966426

Böhlke, J. K. (2002). Denitrification in the recharge area and discharge area of a transient agricultural nitrate plume in a glacial outwash sand aquifer, Minnesota. *Water Resour. Res.*, *38*. doi:10.1029/2001WR000663

Böhlke, J. K., Antweiler, R. C., Harvey, J. W., Laursen, A. E., Smith, L. K., Smith, R. L., & Voytek, M. A. (2009). Multi-scale measurements and modeling of denitrification in streams with varying flow and nitrate concentration in the upper Mississippi River basin, USA. *Biogeochem.*, *93*(1-2), 117–141. doi:10.1007/s10533-008-9282-8

Böhlke, J. K., & Denver, J. M. (1995). Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic Coastal Plain, Maryland. *Water Resour. Res.*, *31*(9), 2319. doi:10.1029/95WR01584

Böhlke, J. K., Judson W. Harvey, & Voytek, M. A. (2004). Reach-scale isotope tracer experiment to quantify denitrification and related processes in a nitrate-rich stream, midcontinent United States. *Limnol. Oceanogr.*, *49*(3), 821–838.

Böhlke, J. K., O'Connell, M. E., & Prestegaard, K. L. (2007). Ground water stratification and delivery of nitrate to an incised stream under varying flow conditions. *J. Environ. Qual.*, *36*(3), 664. doi:10.2134/jeq2006.0084

Browne, B. A., and N. M. Guldan (2005), Understanding long-term baseflow water quality trends using a synoptic survey of the ground water-surface water interface, Central Wisconsin, *J. Environ. Qual.*, *34*(3), 825–35.

Burns, D. A. (1998). Retention of NO_3^- in an upland stream environment: A mass balance approach. *Biogeochem.*, *40*(1), 73–96. doi:10.1023/A:1005916102026

Cey, E. E., Rudolph, D. L., Parkin, G. W., & Aravena, R. (1998). Quantifying groundwater discharge to a small perennial stream in southern Ontario, Canada. *J. Hydrology*, *210*(1–4), 21–37. doi:10.1016/S0022-1694(98)00172-3

Chestnut, T. J., & McDowell, W. H. (2000). C and N dynamics in the riparian and hyporheic zones of a tropical stream, Luquillo Mountains, Puerto Rico. *J. of the N. Am. Benthol. Soc.*, *19*(2), 199–214. doi:10.2307/1468065

Clow, D. W., Schrott, L., Webb, R., Campbell, D. H., Torizzo, A., & Dornblaser, M. (2003). Ground water occurrence and contributions to streamflow in an alpine catchment, Colorado front range. *Ground Water*, *41*(7), 937–950. doi:10.1111/j.1745-6584.2003.tb02436.x

Corbett, D. R., Cable, J. E., Shinn, E. A., Reich, C. D., & Hickey, T. D. (2003). Seepage meters and advective transport in coastal environments: Comments on "Seepage Meters and Bernoulli's Revenge" by E.A. Shinn, C.D. Reich, and T.D. Hickey. 2002. "Estuaries" 25:126-132. *Estuaries*, *26*(5), 1383–1387.

Diaz, R. J., & Rosenberg, R. (1995). Marine benthic hypoxia: A review of its ecological effects and the behavioural responses of benthic macrofauna. *Oceanography and Marine Biology Annual Review*, *33*, 245–303.

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

Spahr, N. E., Sprague, L. A. and W. G. Wilber (2010). The quality of our nation's water—nutrients in the nation's streams and groundwater, 1992–200, *Circ. 1350,* U.S. Geol. Surv., Reston, VA. [available at: http://water.usgs.gov/nawqa/nutrients/pubs/circ1350/.]

Duff, J. H., Tesoriero, A. J., Richardson, W. B., Strauss, E. A., & Munn, M. D. (2008). Whole--stream response to nitrate loading in three streams draining agricultural landscapes. *J. Environ. Qual.*, *37*(3), 1133. doi:10.2134/jeq2007.0187

Edwards, T. K., & Glysson, G. D. (1999). Field methods for measurement of fluvial sediment, *Tech. of Water Resour. Investig., TWRI3-C2*, U.S. Geol. Surv., Reston, VA. [available at: from http://pubs.usgs.gov/twri/twri3-c2/pdf/TWRI_3-C2.pdf.]

Elkins, B. (2007), Nutrient transport between the riparian and hyporheic zones in a small agricultural watershed in the North Carolina coastal plain, M.S. thesis, 90 pp., Dep. of Geol. Sci., East Carolina Univ., Greenville, N.C.

Flewelling, S. A., Herman, J. S., Hornberger, G. M., & Mills, A. L. (2011). Travel time controls the magnitude of nitrate discharge in groundwater bypassing the riparian zone to a stream on Virginia's coastal plain. *Hydrol. Processes*, *26*(8), 1242–1253. doi:10.1002/hyp.8219

Fox, R. J., Fisher, T. R., Gustafson, A. B., Jordan, T. E., Kana, T. M., & Lang, M. W. (2014). Searching for the missing nitrogen: biogenic nitrogen gases in groundwater and streams. *The J. of Agric. Sci.*, *FirstView*, 1–11. doi:10.1017/S0021859614000070

Genereux, D. P., & Hemond, H. F. (1992). Determination of gas exchange rate constants for a small stream on Walker Branch Watershed, Tennessee. *Water Resour. Res.*, *28*(9), 2365–2374. doi:0043-1397/92/92WR-01083

Genereux, D. P., Hemond, H. F., & Mulholland, P. J. (1993). Use of radon-222 and calcium as tracers in a three-end-member mixing model for streamflow generation on the West Fork of Walker Branch Watershed. *J. Hydrology*, *142*(1–4), 167–211. doi:10.1016/0022-1694(93)90010-7

Genereux, D. P., Leahy, S., Mitasova, H., Kennedy, C. D., & Corbett, D. R. (2008). Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA. *J. Hydrology*, *358*(3–4), 332–353. doi:10.1016/j.jhydrol.2008.06.017

Gilmore, T. E., D. P. Genereux, D. K. Solomon, K. M. Farrell, and H. Mitasova (2015a), Aquifer nitrate legacy and dynamics observed from groundwater sampling in streambeds and well nests, *in prep*.

Gilmore, T. E., D. P. Genereux, D. K. Solomon, and J. E. Solder (2015b), Groundwater transit time distribution and mean from streambed sampling in an agricultural coastal plain watershed, North Carolina, USA, *in prep*.

Green, C. T., Puckett, L. J., Böhlke, J. K., Bekins, B. A., Phillips, S. P., Kauffman, L. J., Denver, J. M., and H. M. Johnson (2008). Limited occurrence of denitrification in four shallow aquifers in agricultural areas of the United States. *J. Environ. Qual.*, *37*(3), 994. doi:10.2134/jeq2006.0419

Gu, C., Hornberger, G. M., Herman, J. S., & Mills, A. L. (2008). Influence of stream-groundwater interactions in the streambed sediments on NO_3^- flux to a low-relief coastal stream. *Water Resour. Res.*, *44*, 13 PP. doi:200810.1029/2007WR006739

Harvey, J. W., & Fuller, C. C. (1998). Effect of enhanced manganese oxidation in the hyporheic zone on basin-scale geochemical mass balance. *Water Resour. Res.*, *34*(4), 623–636. doi:10.1029/97WR03606

Hayashi, M., & Rosenberry, D. O. (2002). Effects of ground water exchange on the hydrology and ecology of surface water. *Ground Water*, *40*(3), 309–316. doi:10.1111/j.1745-6584.2002.tb02659.x

Heaton, T. H. E., & Vogel, J. C. (1981). "Excess air" in groundwater. *J. Hydrology*, *50*, 201–216. doi:10.1016/0022-1694(81)90070-6

Hemond, H. F., & Duran, A. P. (1989). Fluxes of N_2O at the sediment-water and water-atmosphere boundaries of a nitrogen-rich river. *Water Resour. Res.*, *25*(5), 839–846. doi:10.1029/WR025i005p00839

Holmes, R. M. (2000). 5 - The importance of ground water to stream ecosystem function. In Jeremy B. Jones & Patrick J. Mulholland (Eds.), Streams and Ground Waters (pp. 137–148). San Diego: Academic Press.

Kalbus, E., Reinstorf, F., & Schirmer, M. (2006). Measuring methods for groundwater -- surface water interactions: a review. *Hydrol.& Earth Syst. Sci.*, *10*(6), 873–887.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2007), Design of a light-oil piezomanometer for measurement of hydraulic head differences and collection of groundwater samples, *Water Resour. Res., 43*, W09501, doi:10.1029/2007WR005904.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2009a), Relationships among groundwater age, denitrification, and the coupled groundwater and nitrogen fluxes through a streambed, *Water Resour. Res., 45*, W09402, doi:10.1029/2008WR007400.

Kennedy, C. D., Genereux, D. P., Corbett, D. R., & Mitasova, H. (2009b), Spatial and temporal dynamics of coupled groundwater and nitrogen fluxes through a streambed in an agricultural watershed, *Water Resour. Res., 45*, W090401, doi:10.1029/2008WR007397.

Kennedy, C. D., Genereux, D. P., Mitasova, H., Corbett, D. R., & Leahy, S. (2008). Effect of sampling density and design on estimation of streambed attributes. *J. Hydrology*, *355*, 164–180. doi:10.1016/j.jhydrol.2008.03.018

Kennedy, C. D., Murdoch, L. C., Genereux, D. P., Corbett, D. R., Stone, K., Pham, P., & Mitasova, H. (2010). Comparison of Darcian flux calculations and seepage meter measurements in a sandy streambed in North Carolina, United States. *Water Resour. Res.*, *46*. doi:10.1029/2009WR008342

Kilpatrick, F. A., & Cobb, E. D. (1985). Measurement of discharge using tracers. U.S. Geol. Surv.Tech. of Water Resour. Investig., *TWRI3-A16*, U.S. Geol. Surv., Reston, VA. [available at: http://pubs.usgs.gov/twri/twri3-a16/.]

Kilpatrick, F. A., Rathbun, R. E., Yotsukura, N., Parker, G. W., & Delong, L. L. (1989). Determination of stream reaeration coefficients by use of tracers*, Tech. of Water Resour. Investig., TWRI3-A18*, U.S. Geol. Surv., Reston, VA. [available at: http://pubs.usgs.gov/twri/twri3-a18/.]

Kimball, B. A., Nimick, D. A., Gerner, L. J., & Runkel, R. L. (1999). Quantification of metal loading in Fisher Creek by tracer injection and synoptic sampling, Park County, Montana, August 1997*, Water Resour. Investig. Rep. 99-4119*, U.S. Geol. Surv., Reston, VA. [available at: http://pubs.er.usgs.gov/publication/wri994119.]

Kimball, B. A., & Runkel, R. L. (2009). Spatially detailed quantification of metal loading for decision making: Metal mass loading to American Fork and Mary Ellen Gulch, Utah. *Mine Water & the Environ.*, *28*(4), 274–290. doi:10.1007/s10230-009-0085-5

Kimball, B. A., Runkel, R. L., Walton-Day, K., & Bencala, K. E. (2002). Assessment of metal loads in watersheds affected by acid mine drainage by using tracer injection and synoptic sampling: Cement Creek, Colorado, USA. *Appl. Geochem.*, *17*(9), 1183–1207. doi:10.1016/S0883-2927(02)00017-3

Kim, H., & Hemond, H. (1998). Natural discharge of volatile organic compounds from contaminated aquifer to surface waters. *J. of Environ. Eng.*, *124*(8), 744–751. doi:10.1061/(ASCE)0733-9372(1998)124:8(744)

Kirkup, L., & Frenkel, B. (2006). An Introduction to Uncertainty in Measurement. Cambridge University Press, New York.

Krause, S., & Blume, T. (2013). Impact of seasonal variability and monitoring mode on the adequacy of fiber-optic distributed temperature sensing at aquifer-river interfaces. *Water Resour. Res.*, *49*(5), 2408–2423. doi:10.1002/wrcr.20232

Krupa, S. L., Belanger, T. V., Heck, H. H., Brock, J. T., & Jones, B. J. (1998). Krupaseep—The next generation seepage meter. *Journal of Coast. Res.*, 210–213.

Laursen, A. E., & Seitzinger, S. P. (2002). Measurement of denitrification in rivers: an integrated, whole reach approach. *Hydrobiologia*, *485*(1-3), 67–81. doi:10.1023/A:1021398431995

Laursen, A., & Seitzinger, S. (2005). Limitations to measuring riverine denitrification at the whole reach scale: effects of channel geometry, wind velocity, sampling interval, and temperature inputs of N_2-enriched groundwater. *Hydrobiologia*, *545*(1), 225–236. doi:10.1007/s10750-005-2743-3

Lee, D. R. (1977). A device for measuring seepage flux in lakes and estuaries. *Limnol. Oceanogr.*, *22*(1), 140–147.

Lindsey, B. D., Phillips, S. W., Donnelly, C. A., Speiran, G. K., Plummer, L. N., Böhlke, J. K., M.J. Focazio, W.C. Burton, and E. Busenberg (2003). Residence times and nitrate transport in ground water discharging to streams in the Chesapeake Bay watershed*, Water Resour. Investig. Rep. 03-4035*, U.S. Geol. Surv., Reston, VA.

McCutchan, J. H. J., III, J. F. S., Pribyl, A. L., & Lewis, W. M. (2003). Open-channel estimation of denitrification. *Limnol. Oceanogr. Methods*, (1), 74–81. doi:10.4319/lom.2003.1.74

McMahon, P. B., & Böhlke, J. K. (1996). Denitrification and mixing in a stream—aquifer system: effects on nitrate loading to surface water. *J. Hydrology*, *186*(1–4), 105–128. doi:10.1016/S0022-1694(96)03037-5

Modica, E., Buxton, H. T., & Plummer, L. N. (1998). Evaluating the source and residence times of groundwater seepage to streams, New Jersey Coastal Plain. *Water Resour. Res.*, *34*, 2797. doi:10.1029/98WR02472

Mulholland, P. J., Valett, H. M., Webster, J. R., Thomas, S. A., Cooper, L. W., Hamilton, S. K., & Peterson, B. J. (2004). Stream denitrification and total nitrate uptake rates measured using a field 15N tracer addition approach. *Limnol. Oceanogr.*, *49*(3), 809–820. doi:10.4319/lo.2004.49.3.0809

Murdoch, L. C., & Kelly, S. E. (2003). Factors affecting the performance of conventional seepage meters. *Water Resour. Res.*, *39*, 10 PP. doi:200310.1029/2002WR001347

Obenour, D. R., Michalak, A. M., Zhou, Y., & Scavia, D. (2012). Quantifying the impacts of stratification and nutrient loading on hypoxia in the northern Gulf of Mexico. *Environ. Sci. Technol.*, *46*(10), 5489–5496. doi:10.1021/es204481a

Paerl, H., Valdes, L., Piehler, M., & Stow, C. (2006). Assessing the effects of nutrient management in an estuary experiencing climatic change: The Neuse River estuary, North Carolina. *Environmental Management*, *37*(3), 422–436. doi:10.1007/s00267-004-0034-9

Puckett, L. J., Zamora, C., Essaid, H., Wilson, J. T., Johnson, H. M., Brayton, M. J., & Vogel, J. R. (2008). Transport and fate of nitrate at the ground-water/surface-water interface. *J. Environ. Qual.*, *37*(3), 1034. doi:10.2134/jeq2006.0550

Rosenberry, D. O. (2008). A seepage meter designed for use in flowing water. *J. Hydrology*, *359*(1–2), 118–130. doi:10.1016/j.jhydrol.2008.06.029

Rosenberry, D. O., & Menheer, M. A. (2006). A system for calibrating seepage meters used to measure flow between ground water and surface water, *Water Resour. Investig. Rep., 2006-5053,* U.S. Geol. Surv., Reston, VA. [available at: http://pubs.usgs.gov/sir/2006/5053/.]

Rosenberry, D. O., & Pitlick, J. (2009). Local-scale variability of seepage and hydraulic conductivity in a shallow gravel-bed river. *Hydrol. Processes*, *23*(23), 3306–3318. doi:10.1002/hyp.7433

Sheibley, R. W., Duff, J. H., Jackman, A. P., & Triska, F. J. (2003). Inorganic nitrogen transformations in the bed of the Shingobee River, Minnesota: Integrating hydrologic and biological processes using sediment perfusion cores. *Limnol. Oceanogr.*, *48*(3), 1129–1140.

Sholkovitz, E., Herbold, C., & Charette, M. (2003). An automated dye-dilution based seepage meter for the time-series measurement of submarine groundwater discharge. *Limnol. Oceanogr.: Methods*, *1*, 16.

Smith, R. L., Howes, B. L., & Duff, J. H. (1991). Denitrification in nitrate-contaminated groundwater: Occurrence in steep vertical geochemical gradients. *Geochim. Chosmochim. Acta.*, *55*(7), 1815–1825. doi:10.1016/0016-7037(91)90026-2

Smith, T. E., Laursen, A. E., & Deacon, J. R. (2008). Nitrogen attenuation in the Connecticut River, northeastern USA; a comparison of mass balance and N2 production modeling approaches. *Biogeochem.*, *87*(3), 311–323. doi:10.1007/s10533-008-9186-7

Solder, J. E. E. (2014, August). Quantifying groundwater-surface water exchange: development and testing of shelby tubes and seepage blankets as discharge measurement and sample collections devices. M.S. thesis, 152 pp., Dep. of Geol. and Geophysics, Univ. of

Utah, Salt Lake City, UT. [available at: http://content.lib.utah.edu/cdm/singleitem/collection/etd3/id/3170/rec/1.]

Stelzer, R., Drover, D., Eggert, S., & Muldoon, M. (2011). Nitrate retention in a sand plains stream and the importance of groundwater discharge. *Biogeochem.*, *103*(1), 91–107. doi:10.1007/s10533-010-9449-y

Stelzer, R. S., Bartsch, L. A., Richardson, W. B., & Strauss, E. A. (2011). The dark side of the hyporheic zone: depth profiles of nitrogen and its processing in stream sediments. *Freshwater Biology*, *56*(10), 2021–2033. doi:10.1111/j.1365-2427.2011.02632.x

Stolp, B. J., D. K. Solomon, A. Suckow, T. Vitvar, D. Rank, P. K. Aggarwal, and L. F. Han (2010), Age dating base flow at springs and gaining streams using helium-3 and tritium: Fischa-Dagnitz system, southern Vienna Basin, Austria, *Water Resour. Res.*, *46*, W07503, doi:201010.1029/2009WR008006.

Taylor, J. R. (1997). An Introduction to Error Analysis (2nd ed.). USA: University Science Books.

Tesoriero, A. J. (2005). Nitrogen transport and transformations in a coastal plain watershed: Influence of geomorphology on flow paths and residence times. *Water Resour. Res.*, *41*(2). doi:10.1029/2003WR002953

Tesoriero, A. J., Duff, J. H., Saad, D. A., Spahr, N. E., & Wolock, D. M. (2013). Vulnerability of streams to legacy nitrate sources. *Environ. Science & Technol.*, *47*(8), 3623–3629. doi:10.1021/es305026x

Visser, A., Broers, H. P., & Bierkens, M. F. P. (2007). Dating degassed groundwater with $^{3}H/^{3}He$. *Water Resour. Res.*, *43*(10), W10434. doi:10.1029/2006WR005847

Vogel, J. C., Talma, A. S., & Heaton, T. H. E. (1981). Gaseous nitrogen as evidence for denitrification in groundwater. *J. Hydrology*, *50*, 191–200. doi:10.1016/0022-1694(81)90069-X

Wanninkhof, R., Mulholland, P. J., & Elwood, J. W. (1990). Gas exchange rates for a first-order stream determined with deliberate and natural tracers. *Water Resour. Res.*, *26*(7), 1621–1630. doi:10.1029/WR026i007p01621

Winter, T. C., Harvey, J. W., Franke, L. O., & Alley, W. M. (1998). Ground Water and Surface Water A Single Resource, *Circ. 1139*, [available at: http://pubs.usgs.gov/circ/circ1139/.]

CHAPTER 2

GROUNDWATER TRANSIT TIME DISTRIBUTION AND MEAN FROM STREAMBED SAMPLING IN AN AGRICULTURAL COASTAL PLAIN WATERSHED, NORTH CAROLINA, USA

2.1. Abstract

Groundwater mean transit time (*MTT*) and transit time distribution (*TTD*) are important variables for determining contaminant discharge chronologies of groundwater systems. We measured groundwater age and seepage rate in a sandy streambed using point-scale sampling (pushable probe with 5 cm screen) and seepage blankets (a novel seepage meter design). *MTT* determined from streambed point sampling along eight closely-spaced transects in a 58 m reach (29 years, n = 35) was very similar to the *MTT* estimate from sampling along six widely spaced transects within a 2.5 km reach (31 years, n = 23). The *TTD* for groundwater discharging to the stream was best fit by a gamma distribution model and was very consistent based on streambed point sampling in both the 58 m and 2.5 km reaches. An exponential-piston flow model also gave a reasonable fit. Overall, coupled streambed point measurements of groundwater age and groundwater seepage rate represent a novel and effective approach to estimating aquifer *TTD* as well as *MTT*. Apparent groundwater ages for samples collected from seepage blankets (n = 4 – 10) were generally younger compared to samples collected from adjacent streambed point locations, largely because blanket samples contained a fraction of "young" stream water. Correcting blanket data for the stream water fraction (using a tracer that was steadily injected into the stream) brought apparent ages for blanket samples closer to apparent ages for adjacent streambed point samples, although a few correction calculations yielded improbable or negative age-dating concentrations. The *MTT* from corrected blanket data were in good agreement with the *MTT* from sampling streambed points adjacent to the blankets. Comparing groundwater concentrations from multiple age-

dating tracers suggested that mixing of a wide range of groundwater ages was not prevalent in the streambed and in most cases piston flow apparent ages were reasonable estimates of the transit time of groundwater.

2.2. Introduction

Groundwater age has been used to assess historic trends in non-point source contamination by linking groundwater contaminant concentrations with apparent recharge years [*Böhlke and Denver*, 1995; *Modica et al.*, 1998; *Böhlk*e, 2002; *Böhlk*e *et al.*, 2002; *Puckett and Cowdery*, 2002; *Tesoriero et al.*, 2005; *Koh et al*, 2006; *Puckett et al.*, 2011; *Welch et al.*, 2011]. Groundwater transport of legacy contaminants (e.g., excess nutrients in agricultural watersheds) into streams and rivers is a likely contributor to the lag in water quality improvement in response to nutrient management initiatives [*Meals et al.,* 2010; *Sanford and Pope,* 2013] and ecological restoration [*Puckett*, 2004; *Hamilton et al.*, 2012]. This lag is likely linked to both the groundwater mean transit time (*MTT*) and the full groundwater transit time distribution (*TTD*). Here we use standard definitions in the groundwater age-dating field: groundwater age is the travel time through the aquifer between recharge at the water table and sample collection, and groundwater transit time is the travel time through the aquifer between recharge at the water table and discharge at a stream or other surface water body. Thus, transit time ≥ age, the two being equal when the groundwater sample is collected in/below a streambed, just prior to discharge.

Groundwater sampling in streambed discharge areas has been used to estimate the age of groundwater discharging into streams [e.g., *Böhlke and Denver* 1995, *Lindsey et al.,* 2003*, Tesoriero et al.,* 2005]. *Modica et al.* [1998] showed good agreement between ages estimated by particle tracking in a groundwater flow model and tracer-based age dating of samples collected from beneath a gaining stream. Stream water sampling has also been used to determine flow-weighted concentrations of age-dating tracers in groundwater discharge to streams [*Stolp et al.*, 2010; *Solomon et al.*, 2015]. Other groundwater modeling studies have suggested that apparent groundwater age from age-dating tracers may be useful for model

calibration [*Solomon and Sudicky*, 1991; *Reilly et al.,* 1994; *Portniaguine and Solomon*, 1998; *Sanford*, 2011], especially if sampling is conducted in discharge zones [*Molenat et al.,* 2013].

Traditionally, groundwater *MTT* and age distributions have been evaluated by analysis of age-dating tracers in groundwater samples collected from well nests in the recharge areas of unconfined aquifers [*Solomon et al.*, 2006]. Only *Browne and Guldan* [2005] and *Kennedy et al.* [2009a] have combined age estimates with groundwater flux rates at numerous points in a streambed to estimate *MTT* and *TTD* of the groundwater discharging from an aquifer to a stream. Mean transit time (*MTT*) was calculated as the flow-weighted mean apparent age of groundwater seeping through a streambed: $MTT = \Sigma v\tau / \Sigma v$, where v is groundwater seepage rate and τ is apparent groundwater age, both measured at the same location in the streambed. *TTD* were evaluated by plotting apparent ages versus the fraction of groundwater discharge for each sampling location. Based on streambed sampling, *Browne and Guldan* [2005] estimated an *MTT* value of 24 years for groundwater in a sand and gravel aquifer 12 – 30 m thick [*Weeks et al.*, 1965] in central Wisconsin. In the same coastal plain streambed in which we worked, *Kennedy et al.* [2009a] calculated *MTT* of 30 years. Large lateral gradients in apparent groundwater age were observed by *Kennedy et al.* [2009a], and CFC-12 and CFC-113 concentrations from many samples were consistent with the piston flow curve on two-tracer plots, suggesting that groundwater age information in individual flowlines is not "mixed away" by dispersion as groundwater flowlines converge on the stream. The work of *Kennedy et al.* [2009a] suggests that individual point samples of groundwater, collected through a very short screen in/below a streambed, are analogous to samples from short screens in the surficial aquifer. The flow-weighted concentrations derived from stream water sampling may be analogous to sampling a fully screened well. To our knowledge, groundwater age and *MTT* have not been assessed using groundwater collected by seepage meters or related streambed devices.

In this study, we further explored streambed sampling for groundwater age, *MTT*, and *TTD*. We collected groundwater in a coastal plain streambed using short-screened probes at

points in the streambed and analyzed the groundwater for multiple age-dating tracers (^{3}H, ^{3}He, SF_6, and CFCs) to estimate the apparent age of individual groundwater samples, and the *MTT* and *TTD* of the surficial unconfined aquifer. We used different sampling designs for point sampling to observe spatial variability of groundwater age in the streambed and sampled closer to stream banks than in previous work in an effort sample young groundwater and thus more fully capture the full *TTD* for the aquifer. Samples were also collected from seepage "blankets" (a novel seepage meter design) deployed near a subset of point transects, an approach that could require fewer samples (i.e., lower analytical costs) because each blanket integrates more streambed area compared to the short-screened probes. Related papers based on the same study explore groundwater *MTT* at the reach mass-balance scale based on surface water sampling [*Solomon et al.*, 2015], the fate of nitrate in the surficial aquifer [*Gilmore et al.*, 2015b], and recharge rates and trends in nitrate contamination based on streambed point sampling and well sampling [*Gilmore et al.*, 2015a].

2.3. Study site and hydrologic conditions

Our study was conducted in West Bear Creek, within a 2.7 km reach defined by *Gilmore et al.* [2015a,b] that contained reaches previously described by *Kennedy et al.* [2007, 2008, 2009a, 2009b], *Genereux et al.* [2008], and *Solder* [2014]. The stream is deeply channelized and the sandy streambed is about 6.5 m wide. During our first sampling campaign in July 2012, stream discharge at the 200 m station (Fig. 2.1C) was about 57 L/s, roughly an order of magnitude lower than discharge during our second field campaign in March 2013 (~500 L/s) [*Gilmore et al.*, 2015b]. All stream locations are named by their distance in meters downstream of a tracer injection site defined as 0 m (Fig. 2.1C).

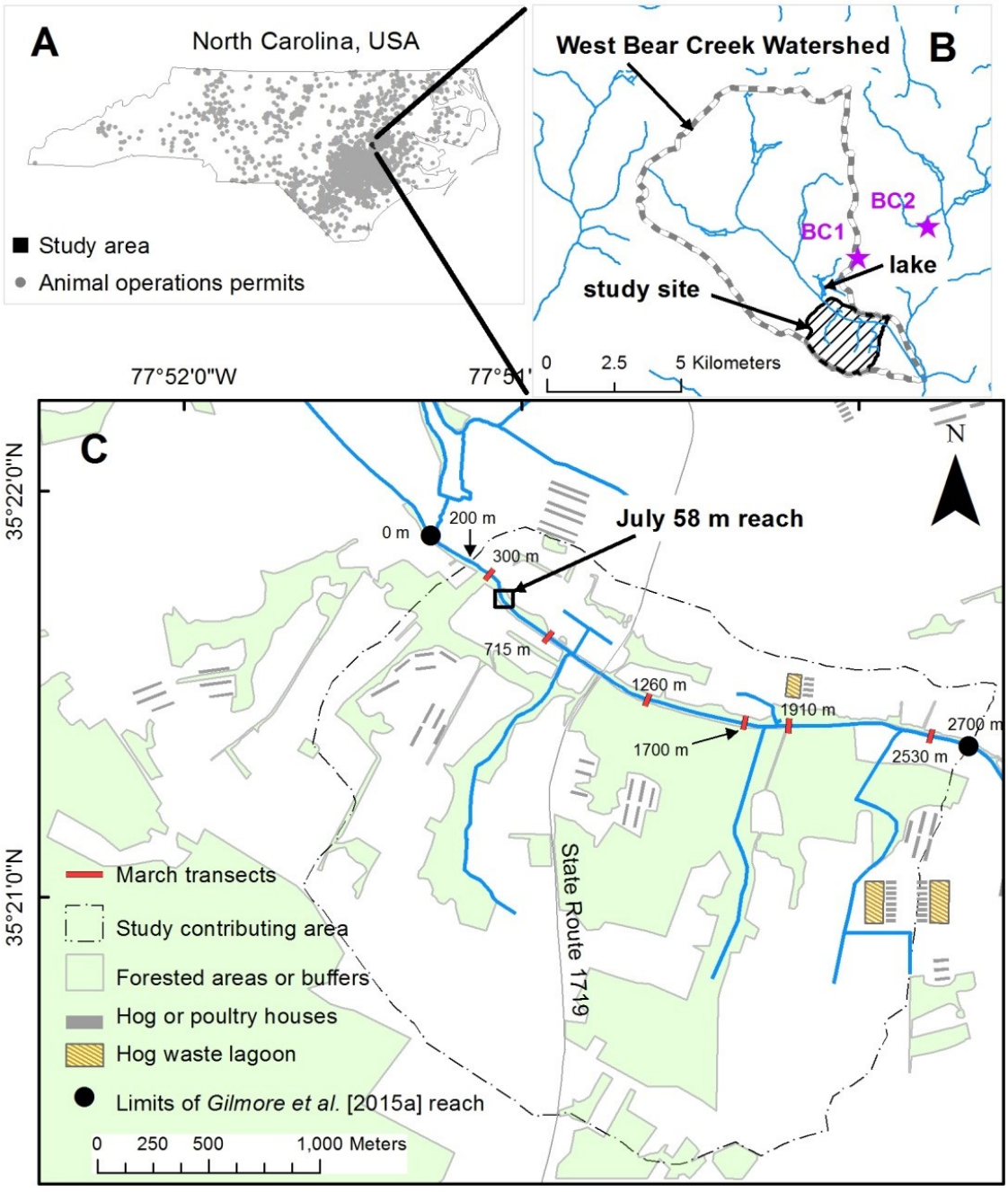

Figure 2.1. Study site and topographically defined contributing area for July 2012 and March 2013 field campaigns. A: study area location in eastern North Carolina. B: West Bear Creek watershed is outlined by the dashed line, and the topographically-defined contributing area for the 2.5 km study reach is defined by the cross-hatched area. Well nests are denoted by stars. C: the West Bear Creek study site contributing area and sampling locations. All sampling occurred within a roughly 2.5 km section of the reach (200 m to 2700 m). In July

2012 all point and blanket sampling was conducted in the "July 58 m reach". In March 2013, six point transects were distributed throughout the 2.5 km reach. Seepage blanket sampling was also conducted at the 715m transect in March. All GIS data were accessed via the NC OneMap Geospatial Portal (data.nconemap.com). Forested areas, agricultural facilities, and tributaries were defined using digital orthophotos (2010 North Carolina Statewide Digital Orthoimagery) and field observations. The contributing area for the 2.5 km reach is based on digital elevation data from the North Carolina Division of Transportation. The West Bear Creek watershed outline is from the USDA NC NRCS 12-Digit Hydrologic Units data set. The main channel of West Bear Creek and locations for animal operations permits were from North Carolina Division of Environment and Natural Resources data sets.

The surficial aquifer is underlain by the Black Creek confining unit [*Winner and Coble*, 1996], the top of which is 18 meters below ground surface at well nest BC1 near West Bear Creek (Fig. 2.1). In a separate borehole along the left bank of West Bear Creek (about 75 m upstream of the 715 m transect in Fig. 2.1C), clayey material at a depth of about 10.6 meters was interpreted as the base of the surficial aquifer [*Kennedy et al.,* 2009b], possibly the Black Creek confining unit.

2.4. Methods

Streambed point and blanket groundwater sampling was done during 3-4 day campaigns in July 2012 and March 2013 (additional details in *Gilmore et al*. [2015b] and *Solder* [2014]). Samples from 35 point locations in July 2012 and 23 point locations in March 2013 were analyzed for ^{3}H and ^{3}He. Five-point transects were closely-spaced in July 2012 (8 transects within a 58 m reach of West Bear Creek), and widely-spaced in March 2013 (6 transects spaced out over >2 km) (Fig. 2.1). The March reach is referred to as the "2.5 km reach" throughout this paper [*Gilmore et al.*, 2015b]. The positions of points in each transect were denoted as right-bank (RB), right (R), center (C), left (L), and left bank (LB). RB and LB points were located as close to the edge of the stream as feasible; fine-grained sediment derived from bank erosion made it difficult to sample within < 1 meter of the water line. In

each field campaign, two transects were also sampled for chlorofluoromethane (CFC-11), dichlorodifluoromethane (CFC-12), trichlorotrifluoroethane (CFC-113), and sulphur hexafluoride (SF_6) in groundwater (Table 2.1).

At each sampling location a piezomanometer [*Kennedy et al.*, 2007] was inserted into the streambed and purged before measuring vertical hydraulic head gradient (*J*). Groundwater samples were then collected from the piezomanometer using a syringe for nitrate or major ion samples [*Gilmore et al.*, 2015b] or peristaltic pump for CFC, SF_6, and tritium samples (Table 2.1). Piezomanometers used at CFC sampling locations were constructed of stainless steel and refrigeration grade copper, with about 30 cm of Viton® tubing in the peristaltic pump head. After sampling from the piezomanometer, a second probe (piezometer, Fig. 2.2) was inserted into the streambed roughly 10 cm from the piezomanometer. A noble gas sample was collected from the piezometer using an inertial pump (Waterra® check-valve installed on the end of the copper tube sample container) to minimize degassing of samples during collection. Once sampling was complete a permeameter was inserted within 10 cm to the right or left of the piezomanometer (and also about 10 cm from the location where the noble gas piezometer was inserted; Fig. 2.2) and vertical hydraulic conductivity (*K*) was measured in the streambed [*Genereux et al.,* 2008].

Streambed blankets were installed at two transects in July 2012 (10 blankets total) and one transect in March 2013 (5 blankets) (Table 2.1). Streambed blankets are low-profile rectangular (71 cm x 107 cm) seepage meters constructed of flexible rubber material and lined with stainless steel foil to avoid sorption of CFCs to the rubber material [*Solder*, 2014]. Each blanket was equipped with a dilution flow meter to measure groundwater seepage rate [*Solder*, 2014]. A complete 5-blanket transect covered the streambed almost fully from waterline to waterline. Groundwater samples were collected from the blankets after measuring blanket discharge. Samples were pumped from each blanket at a flow rate that was lower than the field estimate of ambient groundwater discharge from the blanket. In some cases, blanket discharge was very low and no sample was collected (715LB in March 2013) or only a subset of samples were collected (522L in July 2012).

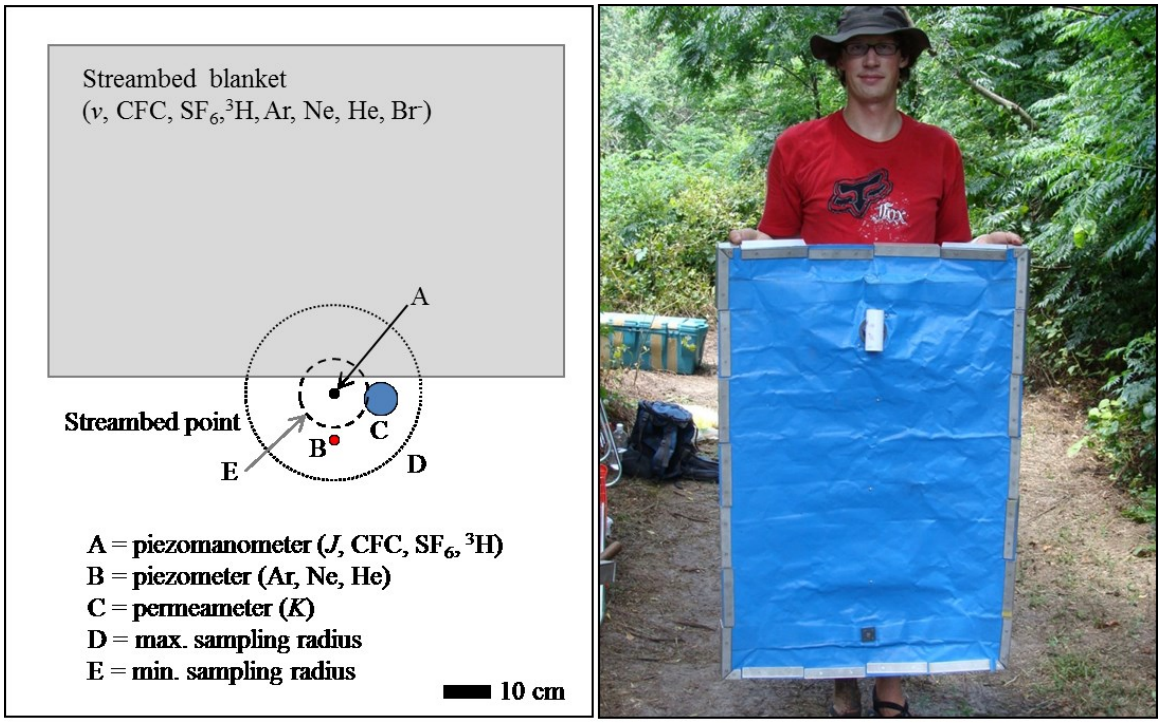

Figure 2.2. Map view of typical streambed sampling location and image of streambed blanket. Each transect consisted of five such locations distributed laterally across the channel. The piezomanometer sampling radius (11-19 cm) is a rough estimate based on minimum and maximum sampling volumes for points. The point measurements were roughly centered in the lateral direction (left to right in the figure) at each blanket location. Point measurements were aligned with the downstream edge of the blanket in March 2013 (as shown in figure), while in July 2012 point measurements were made at either the upstream or downstream edge of each blanket.

Table 2.1. Sampling locations and methods for July 2012 and March 2013 field campaigns.

sampling approach	noble gas sampling locations (m)	noble gas sampling technique; container	CFC, SF_6, and USGS diss. gas sampling locations (m)	CFC, SF_6, and USGS diss. gas sampling technique; container
July 2012 points	466, 474, 481, 491, 499, 508, 516, 524	inertial pump; copper tube	481, 516	peristaltic pump; glass bottles
July 2012 blankets	481, 516	peristaltic pump; copper tube	481, 516	peristaltic pump; glass bottles
March 2013 points	300, 715, 1260, 1700, 1910, 2530	inertial pump; copper tube	715, 1260	peristaltic pump; glass bottles
March 2013 blankets	715	peristaltic pump; copper tube	715	peristaltic pump; glass bottles

Noble gas samples (Xe, Kr, Ar, Ne, He) were collected in copper tubes sealed with steel pinch clamps [*Aeschbach-Hertig and Solomon,* 2013]. Tritium samples were collected in 500 mL HDPE bottles. Tritium and noble gas samples were analyzed at the Dissolved and Noble Gas Laboratory at the University of Utah in Salt Lake City, UT. Groundwater samples for SF_6 and CFC analysis were collected using peristaltic pumps, according to USGS methods (http://water.usgs.gov/lab/) and analyzed at the USGS CFC Lab in Reston, VA.

2.5. Modeling

2.5.1. Groundwater flux

Vertical groundwater flux (seepage rate) was determined at each point sampling location as $v = KJ$, where K (meter/day) is hydraulic conductivity and J (meter/meter) is hydraulic head gradient. Volumetric discharge from each streambed blanket was measured using a flow dilution meter [*Solder*, 2014], then divided by the streambed area covered by the blanket (0.76 m^2) to determine v. Two water fluxes were calculated for each blanket measurement:

"uncorrected" and "corrected". Uncorrected flux was the total water discharge from the blanket, which could include groundwater plus any stream water that entered the blanket through hyphoreic flowpaths. Corrected blanket flux was an estimate of just the groundwater flux from the blanket; to obtain the corrected flux, the streamwater component of the uncorrected flux was estimated from a chemical mixing model and then subtracted from the uncorrected flux. Br^- was injected into the stream as part of a reach mass-balance experiment that was concurrent with the streambed blanket and point sampling [*Gilmore et al.*, 2015a], giving the streamwater component of blanket output a known Br^- concentration. The fraction of blanket flux derived from groundwater was calculated as:

$$F_{gw} = \frac{[\mathrm{Br}^-]_{blanket} - [\mathrm{Br}^-]_{sw}}{[\mathrm{Br}^-]_{gw} - [\mathrm{Br}^-]_{sw}} \tag{2.1}$$

where F_{gw} is the fraction of blanket discharge that was groundwater, and the subscripts "blanket", "sw", and "gw" represent the blanket discharge, stream water, and groundwater, respectively. With F_{gw} known, corrected blanket flux was calculated as $v_{gw} = v_{blanket}F_{gw}$, where v_{gw} and $v_{blanket}$ are the groundwater flux and the total water flux from the blanket, respectively.

Corrected and uncorrected concentrations for dissolved gases (He, Ne, Ar, SF_6, CFCs) and ^3H were also calculated for the water samples from blankets:

$$C_{gw} = \frac{C_{blanket} - \left(1 - F_{gw}\right) C_{sw}}{F_{gw}} \tag{2.2}$$

where C is the solute concentration and the subscript "gw" indicates a corrected blanket value (i.e., a groundwater value). Uncorrected and corrected concentrations were then used to model uncorrected and corrected apparent groundwater ages.

2.5.2. Apparent groundwater age and mean transit time

The closed equilibrium (CE) model [*Aeschbach-Hertig et al.,* 2008] was fit to noble gas (Ar, Ne) data to model He, SF_6, and CFC concentrations, which were then used to determine apparent groundwater age. Groundwater apparent age is an estimate of groundwater age based on the concentrations of age-dating tracers in a groundwater sample, which are assumed to have been transported with that groundwater from recharge until sampling, unaffected by processes such as mixing, dispersion, matrix diffusion, and degradation [*Plummer et al.,* 2006]. The model formulation we used was:

$$C_{i-\mathrm{mod}} = \frac{C_i^{eq}(1 + AH_{i-rech})}{(1 + BH_{i-sam})} \tag{2.3}$$

$C_{i\text{-}mod}$ is the modeled concentration of gas *i* (Ne, Ar), C_i^{eq} is the solubility equilibrium concentration of gas *i* at recharge conditions (recharge temperature, atmospheric pressure and recharge salinity), $H_{i\text{-}rech}$ and $H_{i\text{-}sam}$ are the Henry's Law constants for gas *i* at recharge conditions and sampling (discharge) conditions, respectively, and A and B are gas to water volume ratios in pore space at recharge and discharge, respectively [*Aeschbach-Hertig et al.,* 2008].

Equation 2.3 was applied according to *Gilmore et al.* [2015b]. Samples were categorized as having "excess air" if dissolved [Ne] (or [Ar], if [Ne] was not available) was greater than solubility equilibrium at the recharge temperature, i.e., if ΔNe was positive, where ΔNe = 1 – $[Ne^{eq}_i]$/ $[Ne_{meas}]$, and "meas" = measured concentration in the groundwater sample. Alternately, the sample was considered "degassed" if [Ne] was lower than solubility equilibrium (negative ΔNe) (30 of 35 samples were degassed in July, 10 of 23 samples in March 2013; mean ΔNe was –18% and +4%, respectively). For samples with excess air we set B = 0 and calculated the value of parameter A that gave the best fit to measured [Ar] and [Ne] in the sample. The assumption of B = 0 simplified equation 2.3 to an unfractionated excess air (UA) model (special case of the more general CE model). For degassed samples A

was set equal to 2.3 mL air/L water and the parameter B was determined. The value for the A parameter used in the CE model for degassed streambed groundwater samples (2.3 mL/L) was the mean A value from modeling the dissolved gas data (Xe, Kr, Ar, Ne) from the nearby well nests (Fig. 2.1B). The recharge temperature (12.8 C) used to calculate C_i^{eq} was based on noble gas thermometry [e.g., *Aeschbach-Hertig and Solomon*, 2013] that involved fitting the dissolved gas data from well samples to the UA model (details in *Gilmore et al.* [2015a,b]). Only [Ar] and [Ne] were used in the model because Xe and Kr were injected into the stream as part of a reach mass-balance experiment that coincided with the point and blanket sampling [*Gilmore et al.*, 2015]. Solubility equilibrium concentrations were calculated using solubility equations for Ne [*Weiss*, 1971] and Ar [*Weiss*, 1970].

Building on He-^{3}H relations shown previously [*Schlosser et al.*, 1988; *Solomon et al.*, 1993], tritiogenic helium, $\left[{}^3He_{trit}\right]$, was calculated as:

$$\left[{}^3He_{trit}\right] = \left(\left[{}^4He_{meas}\right]R_{meas} - \alpha\left[{}^4He_{mod}\right]R_{atm} - \left[{}^4He_{terr}\right]R_{terr}\right)\left(1 + BH_{sam}\right) \qquad (2.4)$$

where subscripts "*meas*", "*terr*", "*atm*", and "*mod"* represent measured, terrigenic, atmospheric, and modeled concentrations, respectively, R is the [^{3}He]/[^{4}He] ratio, and α is the isotope fractionation factor (0.983) for ^{3}He and ^{4}He (R_{gas}/R_{water}). [^{4}He$_{terr}$] was calculated as [^{4}He$_{meas}$] – [^{4}He$_{mod}$]. Equation 2.4 assumes that degassing occurred in the discharge area (degassing was not detected in dissolved gas data from recharge area wells, suggesting degassing near the stream was more likely [*Gilmore et al.*, 2015a,b]). The factor (1+BH) is a correction factor for degassing [*Aeschbach-Hertig et al.*, 2008]. R_{terr} was assumed to be 2.0x10^{-8} [*Schlosser et al.*, 1988; *Solomon et al.*, 1993]. Solubility of He was calculated according to Weiss [1971]. Given [^{3}He$_{trit}$] and [^{3}H$_{meas}$], apparent age (τ) was calculated according to the standard ^{3}H-^{3}He age equation [e.g., *Poreda et al.*, 1988].

A and B parameters were used to convert the measured CFC or SF_6 concentrations to an atmospheric mixing ratio (x_i, pptv) as follows [*Friedrich et al.,* 2013]:

$$x_i = \frac{C_{i-sam} K_{i-rech} (1 + BH_{i-sam})}{\left(P_a - p_{H_2O}\right)(1 + AH_{i-rech})} \tag{2.5}$$

where $C_{i\text{-}sam}$ is the concentration of gas *i* in groundwater, $K_{i\text{-}rech}$ is the Henry's Law constant at the recharge temperature and salinity in units of kg-atm/mol, P_a is atmospheric pressure at the recharge elevation and temperature (atm), and p_{H20} is water vapor pressure at the recharge temperature and salinity. Mixing ratios calculated from Equation 2.5 were matched to historic records of CFCs and SF_6 in the atmosphere to determine recharge year, and apparent groundwater age was calculated by subtracting the recharge year from the sampling date (taken as 2012.5 for July 2012 and 2013.0 for March 2013 sampling). SF_6 and CFC solubilities were calculated according to *Bullister et al.* [2002] and *Plummer and Busenberg* [2000], respectively.

Uncertainty in apparent groundwater age from equations 2.3 and 2.5 was assessed for a subset of point samples (those from the 516 m transect, where apparent groundwater age ranged from about 2 to 41 years) using a Monte Carlo approach. For each iteration, (1) input variables were randomly varied ([Ne], [Ar], [He], [SF_6], [CFC]s, [^{3}H], and recharge temperature), (2) [Ne] and [Ar] data were fit to equation 2.3 and either A or B was optimized, and (3) new apparent age estimates were calculated according to equations 2.4 and 2.5. We also evaluated the uncertainty in apparent ages as a result of using recharge temperature and excess air values derived from well data, and sensitivity of ^{3}H/^{3}He apparent age to the value used for R_{terr}. The potential for bias in ^{3}H/^{3}He apparent age if our assumption that degassing occurred at sampling was incorrect was also explored. Uncertainty in apparent age from the ^{3}H/^{3}He was on the order of 2 – 6 years (where older apparent age gave the lowest uncertainties), and uncertainty in *MTT* was estimated at about 15 – 20%. Details of uncertainty analyses are given in Appendix D.

Apparent groundwater ages were weighted by groundwater discharge (v) from points or blankets to determine groundwater MTT through the surficial aquifer ($MTT = \Sigma v\tau / \Sigma v$) [*Kennedy et al.*, 2009a].

2.6. Results

2.6.1. Apparent groundwater ages

2.6.1.1. Groundwater piston-flow ages from points and blankets

Piston-flow groundwater ages from point and blanket sampling approaches ranged from modern (<3 years) to about 70 years (Appendix D). Relative to streambed point results, uncorrected results from streambed blankets showed a bias toward younger apparent age (blanket age = 0.57x(point age) + 7.3 years, $R^2 = 0.37$, p-slope < 0.01, p-intercept = 0.04). Correcting blanket data (unmixing the stream water to isolate the groundwater in blanket samples) improved agreement between blanket and point sampling, but still gave a slope of less than one (age from corrected blanket samples = 0.74x(point age) + 6.1 years, $R^2 = 0.53$, p-slope < 0.01, p-intercept = 0.07). Slopes were closer to one for both uncorrected blanket results (0.77, $p < 0.01$) and corrected blanket results (0.92, $p < 0.01$) when the regression was forced through the origin (Fig. 2.3).

The bias of uncorrected streambed blanket results toward younger groundwater ages is largely explained by the presence of stream water (with young apparent age due to partial in-stream re-equilibration with the atmosphere) in some blanket samples. Modeling stream water dissolved gas concentrations (equations 2.3-2.5) that were used to correct blanket data suggested relatively young piston-flow ages for stream water in March 2013 ($^3H/^3He$ age = 2.2 years, SF_6 age = 4.5 years), and somewhat older ages in July 2012 ($^3H/^3He$ age = 13.6 years, SF_6 age = 8.0 years). Thus, uncorrected blanket ages are younger than corrected blanket ages because stream water is relatively "young" with respect to its age-dating tracer signatures. Details of the stream water dissolved gas concentrations are provided in the

Appendix D. Groundwater fractions in blanket samples were determined from equations 2.1 and 2.2 and ranged from 0.16 to 1.0, where a fraction of 1.0 indicates that no stream water was detected in the blanket sample (Table 2.2).

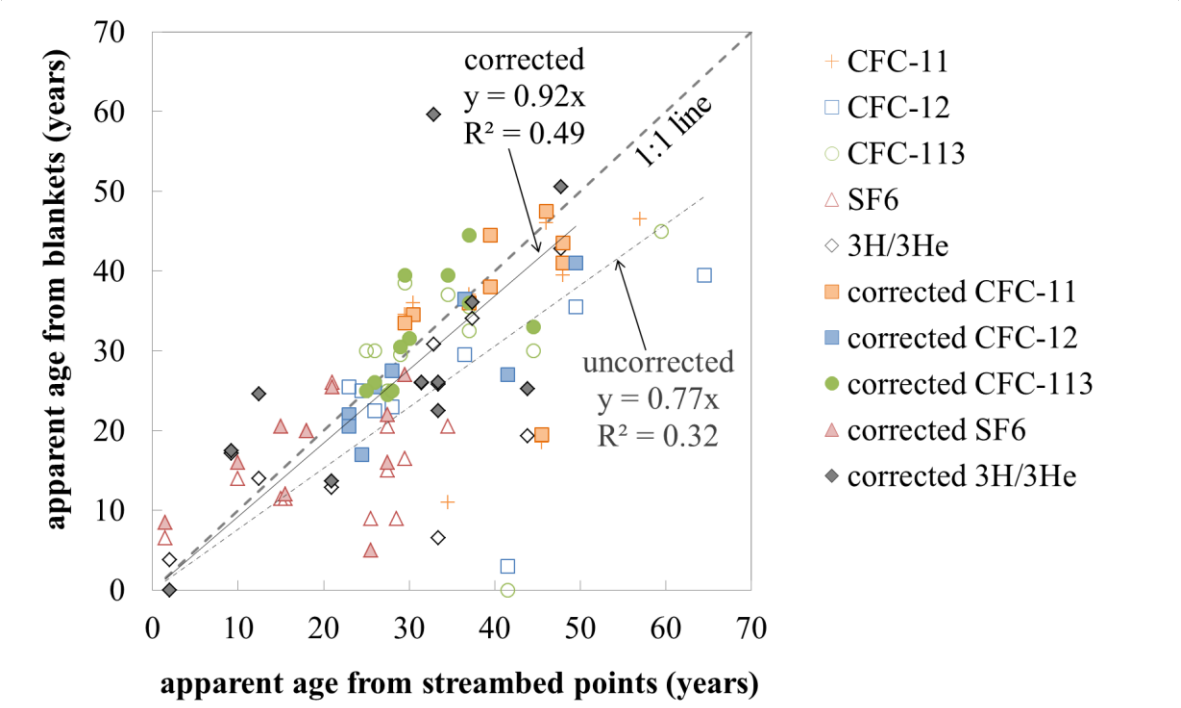

Figure 2.3. Relationship between apparent age of groundwater from streambed point sampling and from streambed seepage blankets. Apparent ages are from $^{3}H/^{3}He$, SF_6, CFC-11, CFC-12, and CFC-113 age-dating methods. "Corrected" indicates that apparent age from corrected blanket data is plotted against apparent age from points. Equations shown in figure are for regressions fit through the origin.

Table 2.2. Groundwater fractions, F_{gw}, in blanket samples in July 2012 and March 2013.

sample	F_{gw}
July 2012	
481RB	0.91
481R	0.52
481C	0.89
481L	1.06[a]
481LB	0.87
516RB	0.97
516R	1.02[a]
516C	0.41
516L	0.35
516LB	0.55
March 2013	
715RB	0.16
715R	0.74
715C	0.38
715L	0.52
715LB	no sample

[a] $F_{GW} > 1.0$ is physically impossible, samples were assumed to have no stream water fraction ($F_{GW} = 1.00$)

The distribution of point age minus blanket age (all age-dating tracers) seemed to suggest that differences between point and blanket ages largely were due to random variation, with perhaps a small bias toward positive values (Fig. 2.4). There were more large point-minus-blanket differences (>15 years) for uncorrected blanket ages (7) than for corrected blanket ages (4), and a larger mean difference (6.1 years for uncorrected compared to 1.3 years for corrected). Overall, the difference in the mean apparent age was statistically significant ($p <$

0.01; standard two-tailed t test) for points versus uncorrected blankets but not points versus corrected blankets (p = 0.52). It seems clear that correcting blanket data improved agreement between points and blankets, but after correcting blanket data there were fewer blanket results that could be compared with points (n = 49, versus n = 56 comparisons for uncorrected blankets), because (1) some corrections resulted in negative groundwater concentrations for SF_6 and/or CFCs (481R, 715RB) and (2) some corrected concentrations showed contamination of CFCs (715RB, 715C) (Appendix D). CFC contamination was observed in other point and uncorrected blanket data, but finding negative corrected concentrations is an indication of uncertainty (most likely in the choice of appropriate stream water Br^- and gas concentrations) in correcting the blanket tracer gas data.

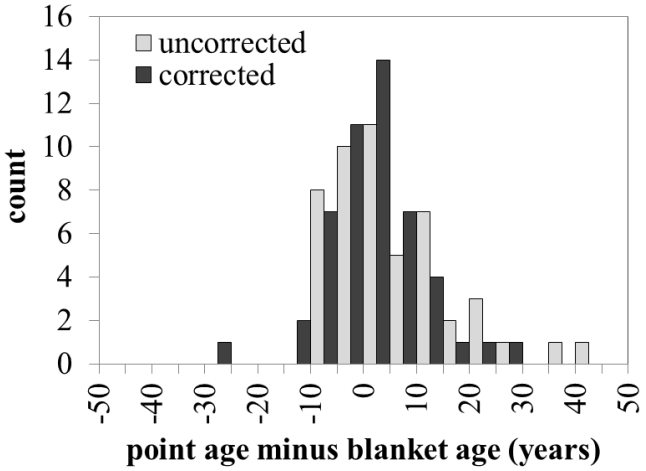

Figure 2.4. Histogram for point age minus blanket age.

2.6.2. Groundwater piston-flow ages from different age-dating tracers

CFC apparent ages for samples collected from both points and blankets suggested older groundwater compared to the other age-dating methods, including $^3H/^3He$ (Fig. 2.5), ranging roughly from 15 – 50 years. Relatively few CFC samples (one point age and one corrected blanket age in July 2012, four uncorrected blanket ages in March 2013; Appendix D) had ages that were ambiguous because of declining CFC concentrations in the atmosphere in recent years. Points or blankets with ambiguous age were assumed to have the apparent CFC age most similar to that suggested by the other age-dating tracers collected from the same point or blanket. Some CFC-12 samples were contaminated (e.g., 4 of 10 points in July 2012), and the pattern of apparent age, CFC-12 < CFC-113 < CFC-11, suggested that the CFC results may have been influenced by microbial degradation and/or sorption in the surficial aquifer (*Cook et al.* [1995]; *Newman et al.* [2010]; *Darling et al.* [2012], and references therein). Additionally, differences in when the groundwater age "clock" was set for CFCs (possibly in unsaturated zone [*Cook and Solomon*, 1995]) compared to $^3H/^3He$ (seasonally low water table [*Cook and Solomon*, 1997]) may have also contributed to the large intercepts (18-25 years) in the regressions in Fig. 2.5. As a result, we believe the SF_6 and $^3H/^3He$ age dating methods are likely the most reliable in this study and focus primarily on those results.

In March 2013, SF_6 and $^3H/^3He$ apparent age from sampling both points and blankets plotted near the 1:1 line (Fig. 2.6) and a regression on apparent ages from points gave a slope of close to one (SF_6 age = 0.85x($^3H/^3He$ age) + 2.4 years, p-slope < 0.01, $R^2 = 0.89$). July 2012 apparent ages from SF_6 and $^3H/^3He$ also showed a strong correlation ($R^2 = 0.86$), but samples generally did not fall along the 1:1 line (SF_6 age = 0.39x($^3H/^3He$ age) + 10.7 years, p-slope < 0.01). In July 2012, groundwater ages based on SF_6 in blanket samples seemed biased-low, with almost no relationship between SF_6 age and $^3H/^3He$ (slope = -0.13 and 0.09 for uncorrected and corrected blanket data, respectively; Fig. 2.6).

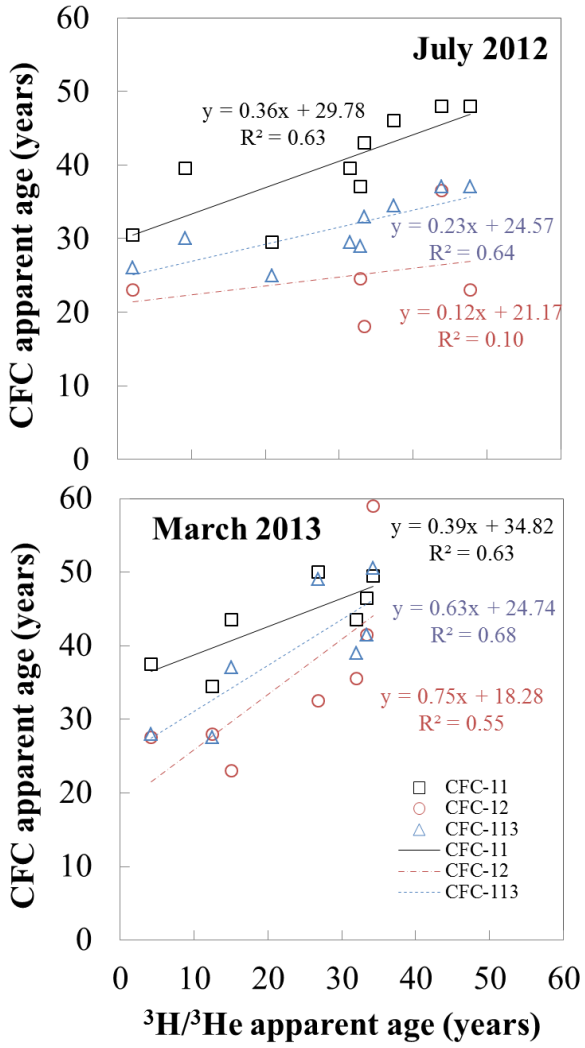

Figure 2.5. CFC apparent ages versus $^{3}H/^{3}He$ apparent ages for groundwater from streambed point sampling.

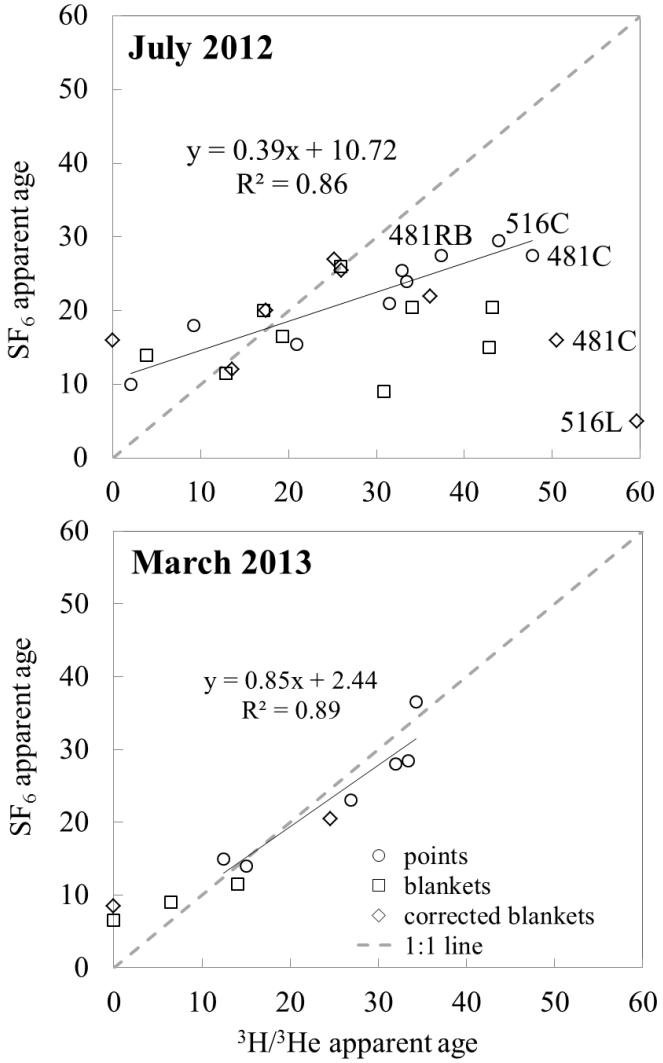

Figure 2.6. SF_6 apparent ages versus $^3H/^3He$ ages, for groundwater from streambed point and blanket sampling. Regression results are shown for point data only.

One difference between the July 2012 and March 2013 data sets was that about 85% of July point samples were degassed, while only 43% of point samples were degassed in March. *Plummer et al.* [2001] and *Visser et al.* [2007] both suggest that SF_6 apparent age may be less reliable than $^3H/^3He$ age when degassing occurs. Our data seem consistent with that assertion, particularly for older groundwater where SF_6 apparent ages are sensitive to small differences in SF_6 concentrations. In contrast, $^3H/^3He$ uncertainties are lower (on a percentage basis; Appendix D and e.g., *Visser et al.* [2014]) for older groundwater. In July 2012, there was a statistically significant correlation between ΔNe and the differences between $^3H/^3He$ and SF_6 ages ($y = -0.68x + 11.7$, $p < 0.01$, $R^2 = 0.44$, where x = difference between $^3H/^3He$ and SF_6 age) for points and blankets, suggesting that the differences between $^3H/^3He$ and SF_6 ages are related to degassing. It is possible that the solubility-based CE model is a poor predictor of $[SF_6]$ in degassed samples because of different kinetics for SF_6 versus lighter noble gases (e.g., Ar, Ne). In theory, SF_6 is highly susceptible to degassing because the solubility of SF_6 is low [e.g., *Visser et al.* 2007] (dimensionless Henry's Law constant K_H is 176 for SF_6 versus 90.9 for Ne at 25 °C), but the aqueous diffusion coefficient (D) for SF_6 is also low due to the high molecular weight of SF_6 (D $x10^5$ = 1.2 versus 4.0 cm^2 s^{-1} for Ne at 25 °C [*King and Saltzman*, 1995; *Wise and Houghton*, 1966]). Thus, if the degassing event was brief (e.g., bubble formation during sampling, or ebullition in blankets), it is possible that light noble gases reached solubility equilibrium but SF_6 did not. If so, the CE model would overestimate $[SF_6]$, which, for older groundwater, could lead to significant bias toward younger apparent age.

The same characteristics of SF_6 (low solubility and diffusion coefficient) could contribute to disagreement between $^3H/^3He$ and SF_6 ages for young groundwater, because differences in diffusion coefficients for He and SF_6 could lead to differences in when the groundwater "clock" was set at recharge: at the seasonally low water table for $^3H/^3He$ [*Cook and Solomon*, 2007], but possibly in the unsaturated zone for SF_6. Alternative explanations for differences between SF_6 and $^3H/^3He$ apparent age include (1) mixing of old and young groundwater [*Plummer et al.*, 2001], or (2) the presence of terrigenic SF_6 [*Plummer et al.*, 2001;

Busenberg and Plummer, 2000]. We believe that (1) mixing of a wide range of groundwater ages was generally not a major feature of point samples (see Fig. 2.8-2.9, Section 2.6.6), and (2) significant amounts of terrigenic SF_6 are not likely to be found in a surficial aquifer in the Atlantic coastal plain [*Busenberg and Plummer*, 2000].

Based on this comparison of apparent ages derived from different age-dating tracers, it appears that the $^3H/^3He$ apparent ages likely give the most reliable estimates of groundwater transit time.

2.6.3. Groundwater flux used in MTT and TTD calculations

Groundwater flux estimates needed for calculation of *MTT* ranged in July 2012 and March 2013 from <0.002 cm/day to 4.4 m/day, similar to the range observed in 422 measurements by *Kennedy et al.* [2009b]. Mean groundwater flux based on point measurements was 0.35 m/day (n = 39) in July 2012 and 0.40 m/day (n = 30) in March 2013. In July 2012, uncorrected and corrected water fluxes from blankets were low (v = 0.1 and 0.07, respectively) compared to adjacent point measurements (v = 0.63 m/day, n = 10), but the blanket estimates followed the same pattern across the stream (higher groundwater flux in the center) as point measurements. In March 2013, blanket fluxes (0.49 m/day uncorrected, 0.23 m/day corrected) were similar to adjacent point measurements (0.31 m/day). The reason for the closer point-blanket agreement in March 2013 is not clear but suggests further refinement of groundwater flux measurement from the seepage blanket may be possible and warranted for future studies [*Solder*, 2014; *Gilmore et al.*, 2015b].

2.6.4. Groundwater mean transit times

Groundwater sampling by the point approach was conducted during different seasons, under different stream flow conditions, at different streambed locations and using different sampling densities (points per m^2 of streambed, Table 2.3), but the *MTT* determined by $^3H/^3He$ showed close agreement between July 2012 (29 years) and March 2013 (31 years). These *MTT* values were also similar to another *MTT* estimate from previous work in West

Bear Creek: 30 years, from [*Kennedy et al.,* 2009a], based on a different age dating tracer (CFCs, mainly CFC-12) and streambed point density (Table 2.3). The three estimates of *MTT* from streambed sampling were slightly greater than the *MTT* determined from nearby well nests (Fig. 2.1), which was about 27 years based on $^3H/^3He$ and SF_6 [*Gilmore et al.*, 2015a].

Table 2.3. *MTT* results from three streambed and one well sampling campaign.

	July 2012	March 2013	April 2007[a]	June 2013[b]
MTT (years)	29	31	30	27
Age-dating method	$^3H/^3He$	$^3H/^3He$	CFC	$^3H/^3He$, SF_6
number of samples[c]	35	23	21	6
points per transect	5	5	3	n.a.
sampling density (pts/m^2)	0.09	0.002	0.04	n.a.
distance between transects (m)	8.3	800-900	12.5	n.a.
location in WBC[d] (m)	466-524	300-2530	613-688	near WBC
WBC stream discharge[e] (L/s)	57	504	low flow[f]	n.a.
USGS BC stream discharge[g] (L/s)	0.5	1.8	0.8	3.6

[a] *Kennedy et al.* [2009a]
[b] *Gilmore et al.* [2015a], two well nests (Fig. 2.1) with 3 wells in each nest
[c] in July 2012 and March 2013 the number of samples is not equal to the number of transects times points per transect because some samples were lost during analysis or gave anomalous noble gas concentrations; for wells, a total of 12 samples (3 wells x 2 well nests x 2 tracers) were collected
[d] m downstream of the "0 m" site in Fig. 2.1
[e] discharge at 200 m station in West Bear Creek (WBC)
[f] stream flow was lower than long-term median flow at nearest USGS stream gauge [*Kennedy et al.*, 2009a]
[g] Bear Creek (BC) stream discharge at USGS stream gauge at Mays Store, NC during the middle of the sampling period (http://waterdata.usgs.gov/nwis/uv?site_no=0208925200)

Table 2.4. Mean transit times determined by $^3H/^3He$, SF_6, and CFC age-dating methods.

Row	Sample type	$^3H/^3He$	SF_6	CFC-11[a]	CFC-12[b]	CFC-113[a]
	July 2012 mean transit time in years (# samples[c])					
A	Points	29.2 (35)	26.4 (10)	45.2 (10)	36.0 (6)	38.5 (10)
B	Points at blankets	35.0 (9)	26.9 (9)	45.6 (9)	41.1 (5)	39.6 (9)
C	Blankets	25.2 (9)	17.0 (9)	40.1 (9)	27.4 (5)	34.6 (9)
D	Points at corr. blankets	35.0 (9)	24.0 (8)	41.3 (8)	27.0 (4)	32.0 (8)
E	Corrected blankets	30.2 (9)	18.8 (8)	40.1 (8)	23.5 (4)	34.7 (8)
F	Points with all tracers[d]	38.1 (5)	25.5 (5)	43.7 (5)	24.2 (5)	33.8 (5)
	March 2013 mean transit time in years (# samples[c])					
G	Points	31.0 (23)	31.7 (9)	46.8 (10)	50.2 (9)	45.7 (10)
H	Points at blankets	15.6 (2)	14.9 (3)	36.3 (2)	31.9 (4)	31.3 (4)
I	Blankets	10.5 (2)	9.1 (3)	14.8 (2)	25.1 (4)	20.6 (4)
J	Points at corr. blankets	15.6 (2)	12.9 (2)	45.5 (1)	31.9 (4)	29.9 (3)
K	Corrected blankets	24.0 (2)	13.8 (2)	19.5 (1[e])	32.8 (4)	28.7 (3[e])
L	Points with all tracers[d]	30.8 (6)	32.3 (6)	47.6 (6)	52.2 (6)	47.3 (6)

[a]*MTT* from CFC-11 and CFC-113 are believed to be impacted by sorption or microbial degradation in the surficial aquifer; see Section 2.6.6.
[b]CFC-12 was contaminated in 4 of 10 samples in July 2012.
[c]Number of samples varies between "points at blankets", "corrected blankets", and "blankets" due to samples lost during transport or analysis, noble gas concentrations that were 2x – 3x different than expected based on the mean concentration for the given campaign, contamination of CFCs, or due to impossible negative groundwater concentrations from corrected blanket data.
[d]mean transit times from point sampling locations where apparent age was able to be determined from all five available age-dating tracers
[e]CFC-11 and CFC-113 suggested contamination in one or more blanket samples

MTT values from corrected blanket data were generally in good agreement with those based on point measurements adjacent to the blanket locations (< 1-5.2 years different for cases where n > 2; Table 2.4). *MTT* values from uncorrected data were consistently lower than those from both point measurements and from corrected blanket data, reflecting the

influence of "young" stream water that was present in most blanket samples. With more blanket samples collected in July 2012 than March 2013, the July campaign offers a better point-blanket comparison. Differences in *MTT* between point samples and corrected blanket samples were 0.2 – 5.2 years in July 2012 (Row D vs Row E, Table 2.4). These differences were similar and in some cases less than differences in *MTT* between different age dating tracers in July 2012 (e.g., 1.3-19.5 years in July 2012, Row F in Table 2.4), or between different sample sizes (e.g., 5-15 years based on $^{3}H/^{3}He$ from sampling all the points versus only the points at blankets in July and March; Row A vs B, and Row G vs H in Table 2.4).

The *MTT* values shown in Row F and L of Table 2.4 differed from mean apparent ages by 1 year to 15 years. The largest differences were from the March 2013 sampling where one high groundwater seepage rate measurement (4.6 m/day at location 1260C, more than 10x the mean of 0.40 m/day) strongly influenced the *MTT* values. The 4.6 m/day value had less impact on the larger set of $^{3}H/^{3}He$ ages (n = 23) for which *MTT* was 29 years and un-weighted mean apparent age was 27 years. Thus, we emphasize the need for flow-weighting apparent groundwater age, particularly for small data sets.

2.6.5. Transit time distribution from streambed point samples

When $^{3}H/^{3}He$ groundwater ages from streambed point sampling are weighted by groundwater flux [e.g., *Browne and Guldan*, 2005; *Kennedy et al.*, 2009a] the shape of the cumulative *TTD* [e.g., *Visser et al.*, 2013] are very similar for the July 2012 and March 2013 field campaigns (Fig. 2.7). Data show that about 76% of the groundwater discharging into West Bear Creek had apparent age of 20 – 40 years, with 11-12% of discharge having < 20 year apparent age, and about 13% of discharge in the 40 – 60 year range. *Kennedy et al.* [2009a] also observed a high percentage (66%) of groundwater with apparent CFC age of 20 – 40 years, based on 21 point samples collected in the streambed of West Bear Creek in April 2007 (when stream discharge was intermediate compared to July 2012 and March 2013) (Fig. 2.7). The high degree of reproducibility among the three streambed point sampling campaigns suggests that the *TTD* and *MTT* of groundwater discharging through the

streambed is a highly steady and stationary parameter at West Bear Creek, and that, given enough samples (>20 for these three cases), the streambed sampling approach is robust (i.e., outcomes are not sensitive to the spacing or locations of point transects, hydrologic conditions during sampling, and perhaps even the choice of age-dating tracer, though local contamination and falling atmospheric mixing ratios present challenges for CFC dating). Further, when age results from groundwater sampling beneath a meandering stream channel in Wisconsin [*Browne and Guldan* 2005] are plotted as a cumulative *TTD*, the data show a similar shape to that observed at West Bear Creek (although shifted toward younger ages due to a lower *MTT*; Fig. 2.7). This suggests that the *TTD* observed in this study may not be unique to West Bear Creek.

Our sampling captured somewhat more 0-10 year-old groundwater than previous streambed work by *Kennedy et al.* [2009] (or *Browne and Guldan* [2005]; Fig. 2.7), but still relatively little of this young groundwater. For example, the measured distributions have far too little young groundwater to fit the exponential model (EM, Fig. 2.7) distribution expected for a simple aquifer of uniform thickness and recharge [e.g., *Vogel*, 1967]. Instead, we found that the observed *TTD* more closely resembled distributions that could arise from spatial variability in recharge, such as the exponential-piston flow model (EPM) [e.g., *Solomon et al.*, 2006; their equation 6.12], or a gamma distribution [e.g., *Amin and Campana*, 1996; *Kirchner et al.*, 2010] (Fig. 2.7).

The EPM equation

$$\tau = \frac{L\theta}{R_{EPM}} \ln\left(\frac{L}{L-z}\right) + \frac{L\theta}{R_{EPM}}\left(\frac{x^*}{x}\right) \tag{2.6}$$

describes a flow system where groundwater is recharged in an unconfined portion of the aquifer (length = x), then flows into a confined portion of the aquifer (length = x^*). The aquifer is assumed to have uniform thickness (L) and porosity (θ), and recharge at a rate that is steady as well as uniform in the unconfined portion (R_{EPM}). The age of groundwater (τ) increases with depth (z) in the aquifer, and the minimum age in aquifer discharge to a stream

is defined by the second term in equation 2.6 (i.e., the age of groundwater at $z = 0$ at the aquifer outlet equals the travel time through the confined portion of the aquifer). Equation 2.6 describes the age versus depth relationship in an aquifer, but the ratio of z to L is equivalent to the "fraction younger" value plotted for any given τ in Figure 2.7 because groundwater discharge from the aquifer is assumed to be constant and uniform over the aquifer thickness L.

The two-parameter gamma travel time distribution

$$h(\tau) = \frac{\tau^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-\tau/\beta} \tag{2.7}$$

is defined by a shape factor (α) and a scaling factor (β), where $MTT = \alpha\beta$ [*Kirchner et al.*, 2010; their equation 6]. $\Gamma(\alpha)$ is the gamma function [e.g., *Andrews and Phillips*, 2003]. Gamma distributions with small α (e.g., $\alpha < 1$) have a large fraction of young water and have been used to describe residence time distributions in catchments [e.g., *Kirchner et al.*, 2000, 2010], while the gamma distribution with $\alpha = 1$ is equivalent to the EM distribution (Fig. 2.7) commonly applied to unconfined groundwater systems. The *TTDs* observed in this study are well fit by gamma distributions with large α (e.g., $\alpha \geq 5$, Fig. 2.7), though with some under-prediction at small transit times < 20 yr. To our knowledge, a relationship between the gamma distribution and spatial variation in recharge has not been established, but a preliminary 2-D groundwater model with no dispersion suggests the *TTD* could fit the shape of a gamma distribution with $\alpha \gg 1$ if the aquifer receives lower recharge near the stream and higher recharge further from the stream.

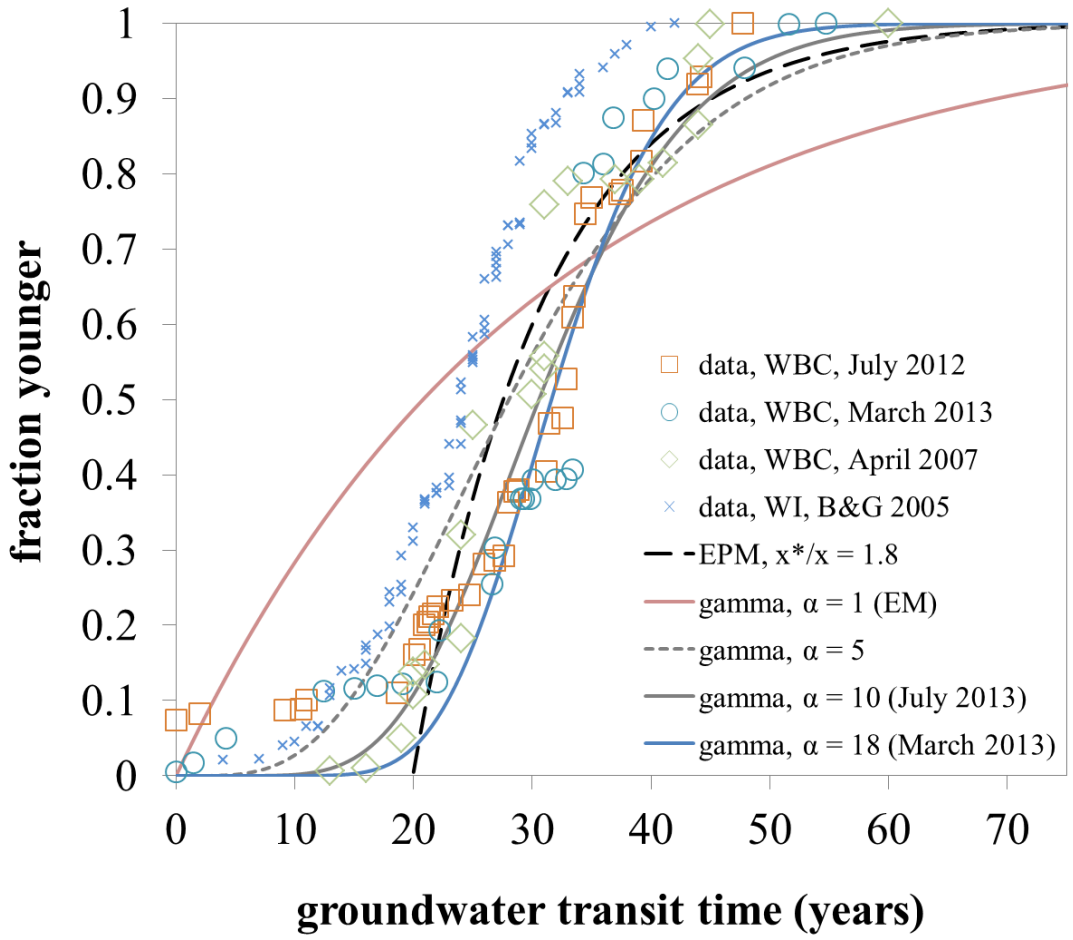

Figure 2.7. The cumulative transit time distribution of groundwater discharging through the West Bear Creek (WBC) streambed in April 2007 [*Kennedy et al.,* 2009a], July 2012 and March 2013. Data from streambed sampling in Wisconsin ("B&G 2005"; *Browne and Guldan* [2005]) are also shown. Apparent groundwater age is from streambed point sampling for CFCs for *Browne and Galdun* [2005] and April 2007, and $^{3}H/^{3}He$ for July 2012 and March 2013 (or SF_6 where $^{3}H/^{3}He$ was unavailable, 1 sample in July and 3 in March). Curves show *TTD* predicted by an exponential-piston flow model (EPM), or gamma distribution ("gamma", with α ranging from 1 – 18, where α = 1 is equivalent to an exponential model (EM)). See text for details on the calculation of curves.

Curves shown in Figure 2.7 were calculated in Excel® using the function GAMMA.DIST(τ, α, β, TRUE), where "TRUE" gives the cumulative distribution. Model parameter values for the EPM (e.g., R_{EPM}, x^*/x) and gamma distributions (α and β) were determined by using Solver® in Excel® to minimize the misfit between observed and modeled "fraction younger" (i.e., minimizing $SSE = \Sigma(\text{observed} - \text{modeled})^2$). Using best estimates for aquifer thickness and porosity (L = 16 m, θ = 0.35 [*Gilmore et al.* 2015a]) and constraining *MTT* to a range of 29 – 31 years (Table 2.3) the EPM was a reasonable fit to the rising limb and tail of the distribution, if the minimum age was set equal to the average apparent age of the left bank and right bank point samples (20 years for July and March field campaigns, combined; Fig. 2.7). The resulting x^*/x and R_{EPM} suggest that the length scale of the confined portion of the aquifer is about 1.8 times that of the unconfined portion, and that the unconfined portion receives recharge at a rate of 50 cm/yr. The *SSE* was large (58), mainly due to poor fit for groundwater ages < 20 years. With only L and θ constrained (all other variables in equation 2.6 were unconstrained), the EPM was visually a poor fit to the observed data and suggested *MTT* = 40 years, which is higher than the *MTT* estimates calculated as flow-weighted mean ages (Table 2.3). Narrow constraints on the gamma distribution fitting parameters (α and β) were not necessary to achieve a good fit to the data. Compared to the EPM, the gamma distribution better fit the observed transit time distributions and had low *SSE* (0.2 for both July 2012 and March 2013 data sets). *MTT* calculated from the gamma parameters was 32 years for both campaigns, very close to the *MTT* values of 29 – 31 years calculated as flow-weighted mean ages (Table 2.3).

Viewed with the *TTD* from *Browne and Guldan* [2005] and *Kennedy et al.* [2009a], the results of this study suggest that gamma distribution with large α may be a better fit than the commonly assumed exponential distribution for the *TTD* in unconfined aquifers. For West Bear Creek, we hypothesize that the groundwater *TTD* may be influenced by spatial variability in recharge. Age distributions at well nests near West Bear Creek (Fig. 2.1) suggest that recharge could vary across the contributing area [*Gilmore et al*., 2015a]. It is possible that shallow low permeability layers (e.g., observed in the borehole next to West

Bear Creek [*Kennedy et al.*, 2009a]) and/or the roughly 150-200 m wide floodplain composed of poorly drained soils on each side of West Bear Creek also limit recharge along some portions of the stream. Conceptually, this would be similar to the distribution of age observed near a stream in Minnesota [*Böhlke et al.*, 2002], where the groundwater system was semi-confined near the stream, and mostly older water was discharging toward the stream (although limited streambed sampling occurred in that study). Variability in recharge could also be magnified by agricultural ditches, which may drain shallow groundwater or capture runoff (from the gentle slopes along the floodplain) that might otherwise recharge the groundwater system. Perhaps minor zones of recharge near the stream could be responsible for the small amount of 0-20 yr old groundwater discharge that was observed (Fig. 2.7).

While the highly idealized EPM distribution can be reasonably fit to the observed *TTD*s, the gamma distribution seems to better fit the general shape of the observed *TTD*s. A related paper [*Gilmore et al.*, 2015a] further explores the issue of spatial variability in groundwater recharge and its relationship to groundwater age distributions measured in the streambed and in wells.

2.6.6. Appropriateness of apparent age estimates

In interpreting groundwater age from age-dating tracer data, a critical question is whether the piston-flow model is appropriate when interpreting age-dating tracer data from streambed point and streambed blanket sampling. The data set from July 2012 and March 2013 includes groundwater concentrations suitable for determining up to five independent age estimates ($^3H/^3He$, SF_6, CFC-11, CFC-12, CFC-113) from blanket samples and a subset of point samples collected adjacent to the blankets. This allows testing for extensive groundwater mixing (i.e., mixing of a wide range of ages, such that the apparent age is no longer a reasonable estimate of the mean age of the sample) with three tracer plots (Fig. 2.8-2.10) that have previously been used for this purpose.

Plots comparing initial tritium ($[^3H_{meas}] + [^3He_{trit}] = [^3H_{initial}]$) to ^3H input from precipitation (Fig. 2.8) are commonly used to evaluate the appropriateness of $^3H/^3He$ apparent ages [e.g., *Friedrich et al.*, 2013; *Visser et al.*, 2013, 2007; *Happell et al.*, 2006; *Koh et al.*, 2006; *Price et al.*, 2003; *Shapiro et al.*, 1999; *Aeschbach-Hertig et al.*, 1998; *Stute et al.*, 1997; *Ekwurzel et al.*, 1994; *Dunkle et al.*, 1993]. No local long-term record of ^3H was available, so we compared results to ^3H data from Cape Hatteras, NC and Washington, DC precipitation. Cape Hatteras is closer to the study site, but is a coastal observation site and may therefore slightly underestimate the ^3H input at West Bear Creek (probably by <10%, e.g., *Ingraham* [1998], their Fig. 3.16). Washington, DC data may slightly overestimate ^3H input at West Bear Creek (perhaps by ~20%, *Ferronsky and Polyakov* [2012], their Fig. 13.15), given that it is a similar distance from the coast but about 400 km north of West Bear Creek. The choice of ^3H input curves is also relevant to the $[^3He_{trit}]$ curves in Fig. 2.9.

Deviations of $[^3H_{initial}]$ in groundwater from the ^3H input curve are usually attributed to (1) dispersion in the groundwater system, which broadens and flattens the bomb peak, (2) mixing of pre- and post-bomb peak water, which usually causes $[^3H_{initial}]$ to plot below the ^3H precipitation curve, or (3) loss of ^3He. ^3He loss could be due to degassing in the ground, artifacts of sampling or analysis, or diffusion. Diffusive losses of ^3He across the water table may occur when recharge rates are low (e.g., <30 mm/yr, [*Solomon and Cook*, 2000]), and/or during recharge years when the bomb-peak ^3H was entering the groundwater system [*Solomon et al.*, 1993] and $^3He_{trit}$ concentration gradients were large.

While the initial tritium plot is useful for evaluating results for evidence of processes as described above, mixtures representing a range of groundwater age (e.g., 10-20 year range) are not distinguishable from unmixed groundwater samples over much of the $[^3H_{initial}]$ input curve (Fig. 2.8, and *Aeschbach-Hertig et al.* [1998]). Deviations of our $[^3H_{initial}]$ estimates from the ^3H precipitation curve (Fig. 2.8) were mostly similar in magnitude to deviations observed in previously published groundwater data [e.g., *Friedrich et al.*, 2013; *Visser et al.*, 2013, 2007; *Koh et al.*, 2006; *Price et al.*, 2003; *Shapiro et al.*, 1999; *Aeschbach-Hertig et al.*, 1998; *Stute et al.*, 1997; *Ekwurzel et al.*, 1994; *Dunkle et al.*, 1993], and could be

explained by moderate dispersion in the aquifer. In March 2013, three groundwater point samples from the downstream end of the 2.5 km reach were apparently recharged during 1955-1965 and stand out because [$^3H_{initial}$] is very low (Fig. 2.8). The low [$^3H_{initial}$] suggests that the samples were composed mostly of older pre-1953 groundwater, and is consistent with the presence of elevated $^4He_{terr}$ in the samples. The small amounts of $^3He_{trit}$ and ^3H in the samples from the 1700C, 1910R, and 1910RB locations may be from mixing or diffusion of ^3He and ^3H from younger groundwater. Removing these three samples from the March 2013 *MTT* calculation reduced *MTT* by only 1.3 years (to 29.7 years), but the detection of ^3H-free groundwater could be conceptually important, given recent studies suggesting the potential importance of old groundwater discharge to streamflow in some hydrologic systems [e.g., *Genereux et al.*, 2009; *Gardner et al.*, 2011; *Frisbee et al.*, 2013]. One corrected blanket sample (481R) stood out in the July 2012 data set (Fig. 2.8) because the apparent age suggested recharge prior to the bomb peak, but [$^3H_{initial}$] was very high. This discrepancy may be an artifact of the blanket correction calculation, as corrections to [SF_6] and all CFC concentrations resulted in negative age-dating tracer concentrations for water from the blanket at 481R. Corrected data from the blanket sampled at 516L also gave high [$^3H_{initial}$], but the value was similar to the ^3H input curve.

If the Cape Hatteras, NC [$^3He_{trit}$] shown in Fig. 2.9 is the most appropriate for West Bear Creek, then corrected concentrations from two blanket samples could be interpreted using the exponential mixing model (EMM). For the 481RB sample the PFM age was 36.1 years compared to an EMM age of about 42 years; for the 481C sample the apparent age could be interpreted as 50 years (PFM) or 22 years (EMM). The corrected data from the blanket at 516L also stood out in the [$^3He_{trit}$] versus x_{SF6} plot (Fig. 2.9), because the results showed high SF_6 concentration (corresponding to young water) compared to the large amount of [$^3He_{trit}$] (corresponding to older water). The cause of this apparent discrepancy is unclear; [$^3H_{initial}$] plotted near the ^3H input curve and ^3H was low (suggesting older water), but CFC-12 data suggested that an apparent age as young as 2 years was possible for this corrected blanket data. With the exception of July 2012 corrected blanket data (biased somewhat by the 516L

blanket data), the flow-weighted mean concentrations (FWM, Fig. 2.9) from points and corrected blanket data plotted near the PFM.

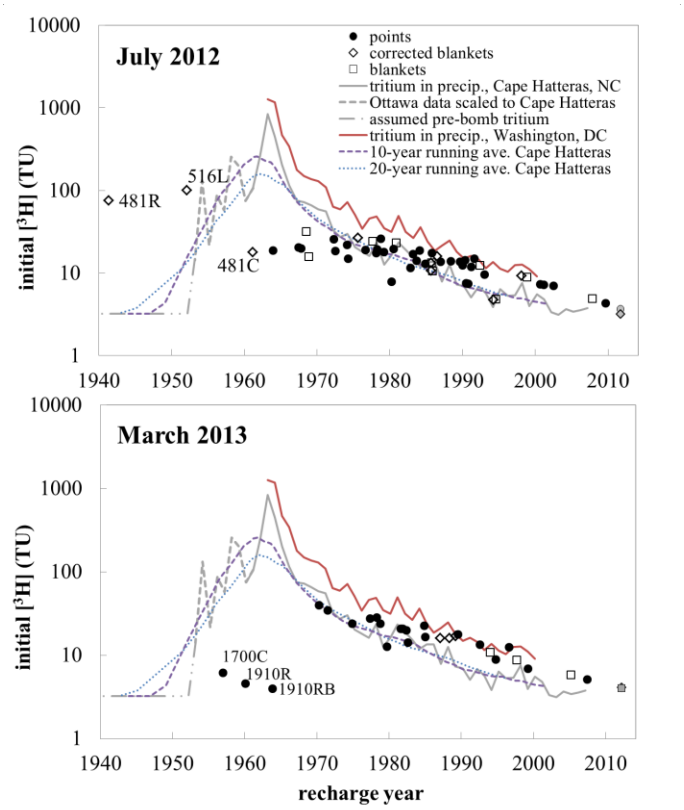

Figure 2.8. Reconstructed initial (recharge) tritium in July 2012 and March 2013 samples, from modeled [$^3He_{trit}$] and measured [^3H], plotted with [^3H] measured in precipitation at Cape Hatteras, NC and Washington, DC. [^3H] data from Cape Hatteras, NC precipitation was not available for recharge years prior to 1960. [^3H] in Cape Hatteras precipitation for recharge years prior to 1960 was estimated using the published correlations with Vienna, Austria ^3H data for Ottawa, and Cape Hatteras [IAEA, 1992]. For recharge years prior to 1954, background [^3H] was assumed to be 5 TU in Vienna [*Kaufman and Libby*, 1954], which translated to 3.2 TU at Cape Hatteras, NC. The "10-year running ave." curve is the 10-year running average of initial [^3H] from Cape Hatteras, NC plotted against the running average apparent age from $^3H/^3He$, after *Aeschbach-Hertig et al.* [1998]. The "20-year running ave." curve was plotted in the same manner.

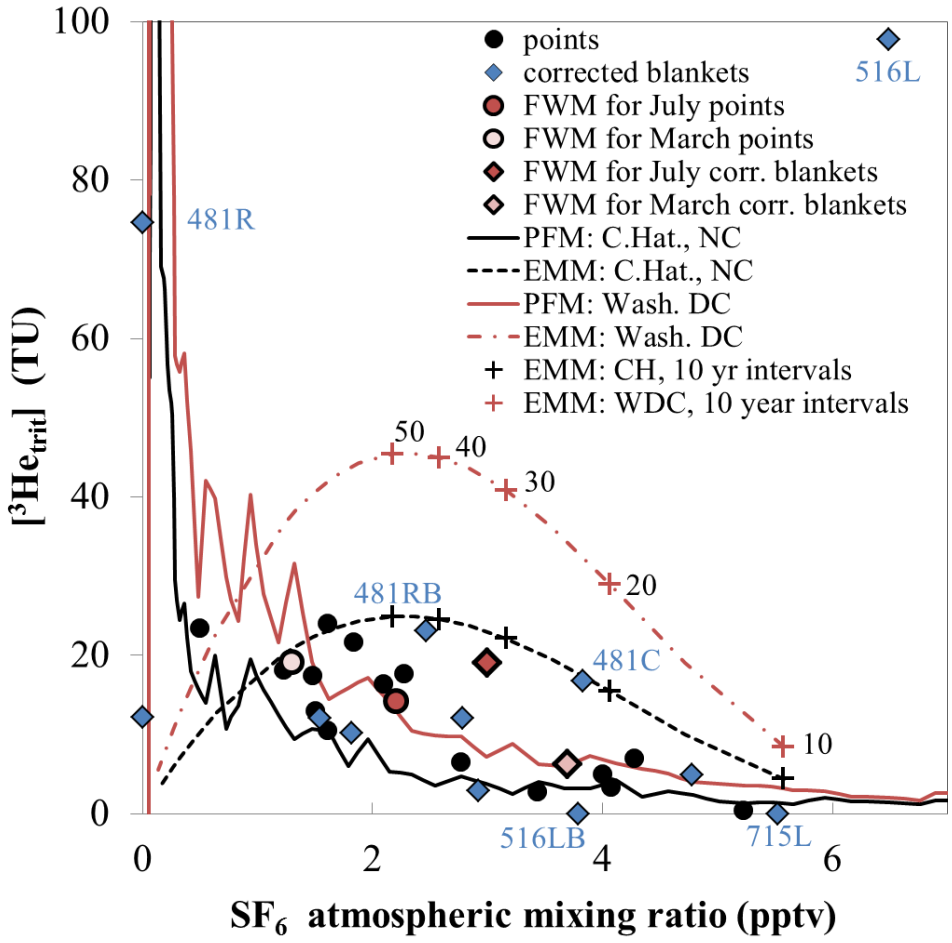

Figure 2.9. Atmospheric mixing ratios modeled from streambed point and streambed blanket dissolved gas data plotted with the measured atmospheric mixing ratio curve (piston-flow model, PFM) and calculated mixing ratio curve from the exponential mixing model (EMM) equation. PFM and EMM models were used to predict [$^3He_{trit}$] for sampling in 2013, based on [^3H] from Cape Hatteras, NC (C.Hat., NC) and Washington, DC (Wash. DC). Crosses on the EMM lines indicate 10-year intervals, beginning at 10 years on the right side of the figure. Blue data labels correspond to corrected blanket data. For 481R, correcting blanket data resulted in a negative concentration for SF_6 (the concentration is shown here as zero). Corrected data from 516LB and 715L blanket locations showed slightly negative [$^3He_{trit}$] (-0.4 and -1.5 TU, respectively) and bothwere interpreted as [$^3He_{trit}$] = 0 for the purpose of estimating age from $^3H/^3He$. Flow-weighted mean (FWM) concentrations for points and corrected blankets are also shown.

CFC-11 and CFC-113 mixing ratios generally fell below both the EMM and PFM lines, likely due to sorption or degradation in the surficial aquifer (Fig. 2.10). Several CFC-12 mixing ratios from July 2012 were clearly affected by contamination (5 samples are not shown in Fig. 2.10, as they were roughly 2x – 10x the scale shown for CFCs), as were two CFC-11 and three CFC-113 mixing ratios in March 2013. CFC-12 can also be affected by microbial degradation in the surficial aquifer, though to a lesser extent than CFC-11 or CFC-113 [*Darling et al.,* 2012]. Degradation of CFC-12 likely explains the lower CFC-12 atmospheric mixing ratios relative to SF_6 (i.e., some mixing may have occurred, but CFC degradation likely explains most of the apparent agreement of some streambed point and blanket samples with the EMM line). Overall, plots of SF_6 and CFC mixing ratios are likely not reliable for identifying a specific model for interpreting point and blanket data (because of likely contamination and/or degradation of CFCs), and overall show little distinction between point and corrected blanket data.

Although streambed sampling occurs in the discharge zone where there is potential for mixing between converging flowpaths [e.g., *Pint et al.,* 2003, *McCallum et al.,* 2014a], the [$^3H_{initial}$] plot (Fig. 2.8) and x_{SF6} (Fig. 2.9) seem to suggest that the piston flow model is reasonable for interpreting most of the data from streambed sampling. Based on the plot of CFCs versus x_{SF6} (Fig. 2.10), the majority of CFC concentrations have been affected by contamination and/or degradation. It is clear that some mixing of groundwater occurred, particularly in blankets, but apparent age estimates from corrected blanket data and from point measurements appear to be representative of the mean age of the mixtures in most cases.

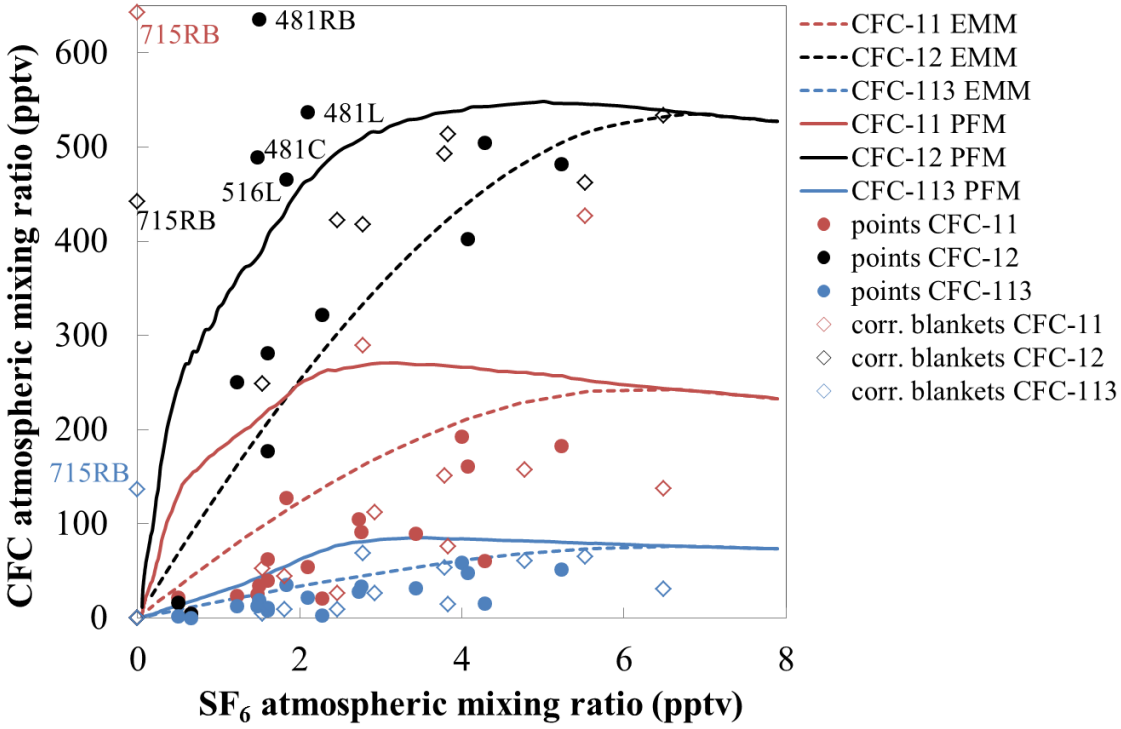

Figure 2.10. Atmospheric mixing ratios modeled from July 2012 and March 2013 point and corrected blanket ("corr. blankets") dissolved gas data. Theoretical lines for piston-flow (PFM) and exponential (EMM) models are also shown. Locations for CFC-12 samples with possible contamination are labeled by the location in the stream (e.g., 481RB). Three point samples and two blanket samples are not shown due to CFC-12 contamination (CFC-12 mixing ratios ranged from 1131 – 6319 pptv). Correcting 715RB blanket data resulted in negative SF_6 concentration (plotted here as zero).

2.6.7. Spatial variability in apparent groundwater age

On average $^3H/^3He$ apparent groundwater ages from point sampling showed a symmetric lateral pattern of higher age in the center of the streambed and younger apparent age toward the stream banks (Fig. 2.11), although this pattern was not present in every point transect (Fig. 2.12 and 2.13). The overall result of greater apparent ages toward the center of the

stream from point samples was consistent with the conceptual model of *Modica et al*. [1998] and with the streambed sampling results of *Kennedy et al.* [2009a]. Based on $^3H/^3He$ apparent ages, uncorrected blanket results followed the trend of greater apparent groundwater age toward the center of the stream in July 2012, but the corrected blanket ages at the "right" and "left" sampling locations were slightly higher than at the center (Fig. 2.11), mostly because of the old (possibly over-corrected) ages at the 481R and 516L blankets.

There were significant lateral gradients in apparent age in the groundwater in/below the streambed in both July 2012 and March 2013, similar to results of Kennedy et al. (2009a). The mean age gradient (difference in two groundwater ages divided by the distance between the two points) was about 10.6 years/m, with a maximum of about 26 years/m in each campaign (based primarily on ages from the $^3H/^3He$, results from SF_6 were substituted for 4 points where the $^3H/^3He$ was not available). The large age gradients indicate that groundwater age signals along flowlines are not lost by transverse dispersion as groundwater converges on the 6-7 m wide streambed of West Bear Creek.

The age gradients across the streambed may also have implications for interpreting groundwater age from streambed blanket sampling, given that each blanket integrates along the length of about 1 m in the lateral direction across the stream. For apparent age estimates from $[SF_6]$ and/or $^3H/^3He$, the apparent age of a mixture of groundwater that spans about 10 years will result in a reasonably close estimate of the true mean age of the mixture, at least for groundwater recharged in the last 30 – 40 years. The likely agreement between the true age and apparent age of a groundwater mixture is because atmospheric $[SF_6]$ has increased in a roughly linear trend for most of the last 30 years, while $[^3H]$ in precipitation has dropped smoothly over the last 40 years (e.g., the difference between apparent and "true" age was <1 year along the 10-year mixing window in Fig. 2.8). If age gradients are large (e.g., >20-30 years/m), then the blanket would integrate a wider groundwater age distribution, with greater potential for bias in apparent age [e.g., *Bethke and Johnson*, 2008; *McCallum et al.*, 2014b] and a greater likelihood of deviating from the tracer curves in tracer-tracer plots (Figs. 2.8-2.10). Two of the four blanket samples that stand out in the $[^3He_{trit}]$ versus x_{SF6} plot (Fig. 2.9)

were located near apparently large groundwater age gradients in the streambed. The largest age gradient calculated from July 2012 data was between the LB and L points at the 516 m transect (26 years/m), the streambed area associated with the ambiguous blanket results at 516L (e.g., Fig. 2.9 and related discussion). Similarly, the streambed blanket at 481C was located near a gradient of 22 years/m, between the points sampled at 481C and 481L.

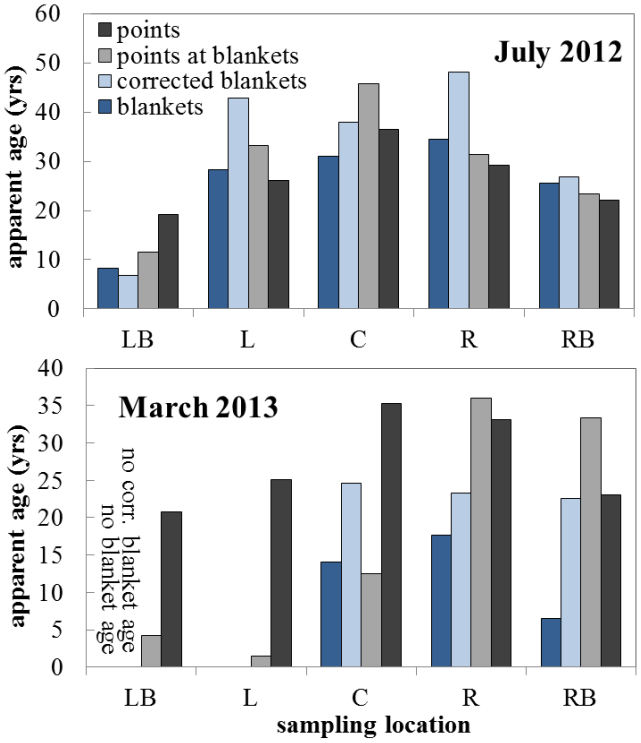

Figure 2.11. Lateral patterns in apparent groundwater age for point and blanket transects, from left bank to right bank. Age estimates are from $^{3}H/^{3}He$. No samples were collected from the left bank (LB) blanket in March 2013. Both corrected and uncorrected blanket data from the left (L) sampling location gave negative apparent age (-3.6 and -8.4 years, respectively) from $^{3}H/^{3}He$ in March 2013, and those samples were interpreted as modern groundwater (apparent age = 0 years).

In the upper half of the July 2012 sampling reach (Fig. 2.12, 466 – 491 m transects) older water was detected along the banks compared to the downstream half of the reach. Profiles of streambed depth (Fig. 2.12) suggest that West Bear Creek is, on average, slightly deeper in the upper half of the reach (25 cm depth) compared to the lower half of the reach (18 cm), which could promote younger groundwater discharging laterally through steep (unsampled) faces of the stream bank in the upper portion of the reach. Given that these steep faces were mostly fine-grained or clayey material (low permeability), groundwater discharge would have likely been low and difficult to sample. In future studies, it would be interesting to assess the importance of groundwater seepage from such steep faces, perhaps using a modified (L-shaped) tube seepage meter [*Solder,* 2014] pushed horizontally into the stream bank.

The left side of the stream in the lower half of the July 2012 reach tended to have deposits of fine-grained sediment on the streambed within ~1 meter of the water line due to clayey materials sloughing off the stream banks. Low permeability of these sediments may have prevented young groundwater from discharging through the streambed close to the waterline. In March 2013, the 715 m transect (where both points and blankets were sampled) is likely an example of this same phenomena, because there was a wide deposit of fine-grained sediment toward the left bank. Apparent groundwater age from both point and blanket samples along the left side of the stream show young water at the 715 m transect (< 5 years, Fig. 2.11; no LB blanket sample was collected because groundwater discharge was too low to adequately flush the blanket), similar to results from the left bank in the downstream half of the July 2012 reach (<10 years from LB, n = 4, Fig. 2.12).

March 2013 point data suggested older groundwater in the lower half of the 2.5 km reach (Fig. 2.13). Some groundwater sampled at the 1700 m and 1910 m transects likely contained some very old groundwater (likely recharged before 1950, prior to significant [^{3}H] in precipitation; Fig. 2.8), and if so, the age for those samples is under-estimated.

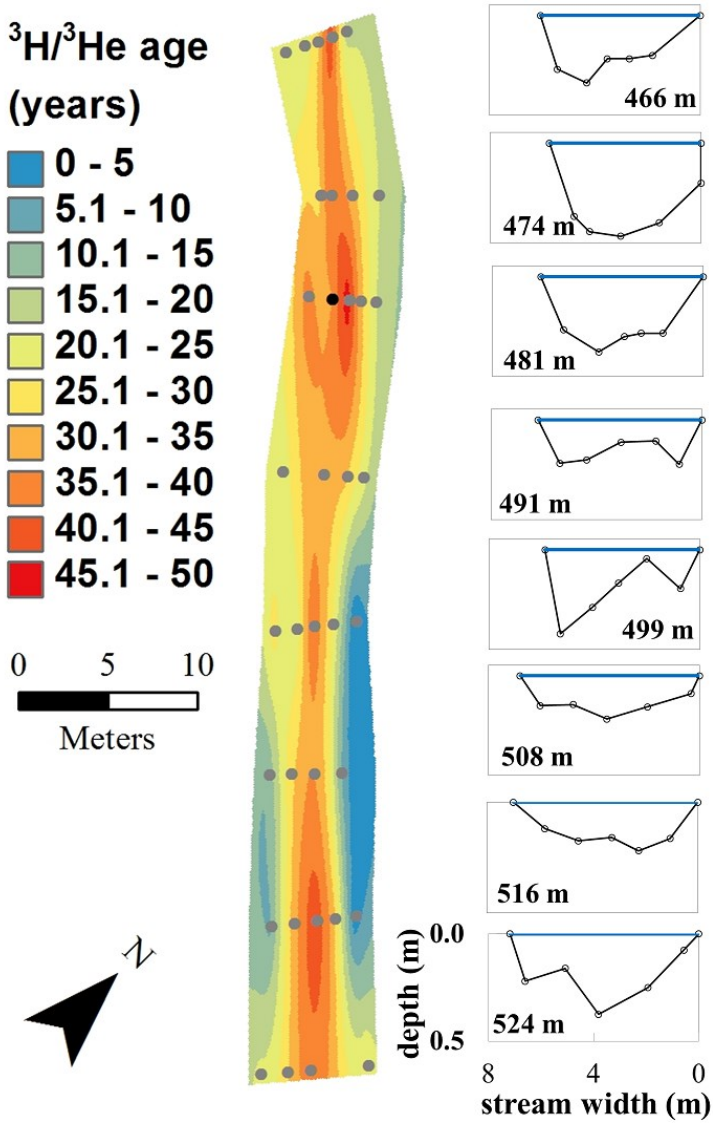

Figure 2.12. Streambed map of apparent groundwater age from the $^3H/^3He$ method, with stream depth profiles for each transect, from July 2012 data. Ages are from the $^3H/^3He$ method (grey circles on streambed map) with the exception of one age estimated from SF_6 (black circle). The direction of stream flow is from the top to the bottom of the map, thus the right bank (RB) is on the left side of the streambed map. Depth profiles were plotted using the axes labeled on the bottom-most plot, with zero stream width starting at the left bank (right side of the figure) and zero water depth along the blue line. Depth profiles are shown in the same order (upstream to downstream) as transects on the streambed map. Blankets were deployed along the transects at 481 m and 516 m. The streambed map was created in ESRI™ ArcMap using the multiquadric radial basis function with anisotropy ratio of 8 and smoothing parameter set = 0.

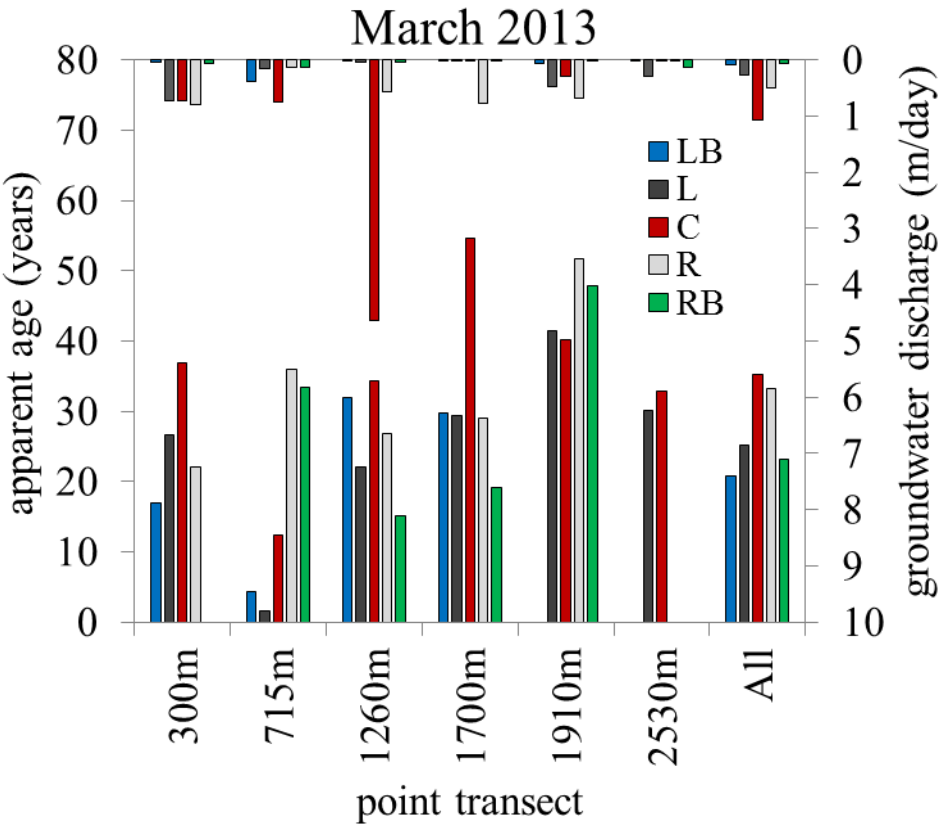

Figure 2.13. Apparent groundwater age from streambed point sampling (n = 26) in March 2013. Groundwater discharge (used to weight ages when calculating MTT) is shown on the upper axis. Ages are from $^{3}H/^{3}He$, and from SF_6 for 3 points at which $^{3}H/^{3}He$ was unavailable. Groundwater age could not be calculated for samples from 1910LB, 2530LB, 2530R, and 2530RB because the samples were lost during analysis (n = 2) or because noble gas concentrations were anomalously high or low (n = 2). Age of zero was estimated for the point sample collected at 300RB. Blankets were deployed along the 715 m transect.

Overall, results from streambed point sampling in July 2012 and March 2013 show lateral patterns in groundwater apparent age that are consistent with previous work [*Kennedy et al.*, 2009a] and conceptual models [e.g., *Modica et al.*, 1998], even though the sampling scales (58 m versus 2.5 km), distances between transects (8.3 m versus 800-900 m), and streamflow

conditions (57 L/s vs 500 L/s) differ by at least an order of magnitude between the two sampling campaigns. Comparisons between apparent ages for groundwater sampled by points and blankets at the same streambed locations were varied, but streambed blanket sampling seemed to capture a similar overall picture of groundwater age across the streambed.

2.7. Summary and conclusions

We used a streambed point approach and seepage meters (flexible streambed seepage "blankets") to sample groundwater exiting the surficial aquifer and entering West Bear Creek in the agricultural coastal plain of North Carolina. Apparent groundwater age was determined from groundwater samples and weighted by groundwater discharge through the streambed to determine flow-weighted mean apparent groundwater age (i.e., groundwater mean transit time, or *MTT*) and flow-weighted distribution of age in aquifer discharge (i.e., groundwater transit time distribution, or *TTD*).

Results from two field campaigns show good agreement in *MTT* based on $^{3}H/^{3}He$ ages (Table 2.3) from dense point measurements in a 58 m reach during low flow (July 2012, *MTT* = 29 yr, n = 35) and more widely dispersed point measurements in a 2230 m reach during high flow (March 2013, *MTT* = 31 yr, n = 23). These *MTT* values were essentially the same as that determined in previous streambed sampling (30 yr, n =21, *Kennedy et al.*[2009a]), even though we sampled closer to the stream bank and detected younger (<10 years) groundwater entering the stream. We found that for these campaigns with n > 20, *MTT* from streambed sampling was not sensitive to sampling arrangement (e.g., point sampling density or transect locations), stream flow conditions, temperature, or the age-dating tracer used. *MTT* derived from groundwater well sampling near West Bear Creek was slightly younger (27 yr) than that from the streambed sampling approach (Table 2.3), as was an estimate from a reach mass-balance study of SF_6 in stream water (26 yr, *Solomon et al.*, [2015]). The reproducibility of *MTT* from various streambed point sampling arrangements in West Bear Creek, and reasonable agreement with the more traditional well sampling approach, suggests

that where streambed point sampling is feasible it is a robust alternative to more traditional well sampling for estimation of *MTT*.

The *TTDs* from streambed point sampling were well fit by a gamma distribution with shape parameter α equal to 10 for July 2012 data and 18 for March 2013 data (Fig. 2.7). The α values for the unconfined aquifer at West Bear Creek and for another sandy unconfined aquifer in Wisconsin are large relative to the values less than 1 that have been observed for transit time distributions on small watersheds [e.g., *Kirchner et al.*, 2000, 2010]. *MTT* estimated by fitting a gamma distribution to the age data was within 1 – 3 years of the *MTT* values mentioned above and computed as flow-weighted mean ages. The exponential-piston flow model (EPM) based on confined groundwater flow without recharge downgradient of unconfined flow with recharge did not fit the measured *TTD* as well as the gamma model, but both models were much closer to the data than the exponential model (EM) often-assumed for unconfined aquifers and based in part on uniform recharge. The superior fit of the gamma model has important implications for transport of non-point-source pollutants through surficial unconfined aquifers and into surface water, e.g., compared to the EM, the observed gamma *TTD* suggests that nitrate output from the groundwater system to streams would initially respond much more slowly to a reduction in N loading, but then the nitrate output would drop steadily for about 20 years (a relatively narrow *TTD*, ~76% of groundwater discharge to West Bear Creek was 20-40 years old), and not show the problematic decades-long tailing behavior characteristic of an exponential distribution.

The better fit of EPM compared to EM suggests that spatial variation in recharge may be important to the observed *TTD* of surficial aquifers, and perhaps that the good fit of the gamma model suggests that the values of its parameters α and β (Eqn. 2.7) for a given aquifer may be related to the spatial distribution of recharge to the aquifer. Other work has drawn a connection between the value of α and watershed hydrological characteristics. For example, in a study of catchment transit times in Scotland, *Hrachowitz et al.* [2010] related the α parameter to drainage density, the presence of hydrologically responsive soils, and catchment water storage. Similar connections may exist for aquifers. We hypothesize that the relatively

small contribution of young water through the streambed of West Bear Creek (~12% was <20 years old) may be linked to low recharge near the stream. The poorly drained soils and shallow low permeability layer in the floodplain likely cause the groundwater system to operate there as a semi-confined aquifer, where groundwater flowing through the system increases in age with little addition of modern water before discharging to the stream. Discharge of younger groundwater likely occurs at locations other than the mainly-horizontal streambed of the main channel, such as agricultural ditches, tributaries, or steep near-vertical faces just above or below the waterline on the stream banks.

Apparent groundwater age was generally older for point samples compared to blanket samples, as illustrated by the slope of regressions (Fig. 2.3) forced through the origin (0.77 for uncorrected blanket values). The slope moved closer to 1 (0.92) when blanket data were corrected for stream water intrusion under the blankets by hyporheic flow, using tracer Br^- released into the stream water at a steady rate. In general, *MTT* results from blanket sampling were similar to *MTT* results from streambed point sampling (Table 2.4) when blanket data were corrected for the presence of stream water in the blanket samples. Overall, our approach of correcting dissolved gas concentrations based on the presence of a stream water tracer (Br^-) seems to result in real improvement in estimates of groundwater age from blanket sampling (they move closer to the age estimates from pure groundwater samples collected below the hyporheic zone by point sampling) (Fig. 2.3).

Apparent groundwater age from most point and corrected blanket data suggested the use of a piston flow model to estimate apparent ages was a reasonable approximation (Fig. 2.8-2.9). In other words, despite the fact that samples were collected from groundwater converging beneath the streambed, mixing across groundwater ages was not strong and samples could be considered to represent groundwater of a single age. Corrected age dating tracer concentrations from only two of the blanket samples (out of 12 blankets where age from SF_6 and the $^3H/^3He$ could be determined) seemed to better fit an exponential model consistent with mixing of groundwater flowlines (Fig. 2.9) but streambed seepage devices

require more complex analyses (e.g., correcting for stream water in samples, or the greater probability for mixed groundwater samples).

Degassing (loss of gas from solution due to bubble formation) was more prevalent in July 2012, when groundwater sampling temperatures were higher compared to March 2013, and may have contributed to differences between apparent ages from $^{3}H/^{3}He$ and SF_6. Production of biogenic gases was a likely driver of degassing, but lower hydrostatic pressure as groundwater approached the streambed (e.g., sampling roughly 0.5 – 1.0 m below stream surface) likely contributed to the formation of bubbles in over-saturated water and subsequent degassing. Degassing can be accounted for by using noble gas modeling, but sampling in colder conditions when gas solubility is higher may minimize degassing and could be an important consideration in choosing the time of sample collection, particularly in areas where biogenic gas production is prevalent (e.g., agricultural areas). We believe $^{3}H/^{3}He$ gave the most robust age estimates for degassed samples, as possible fractionation in some July 2012 samples may have caused SF_6 concentrations to be over-corrected.

As a whole, our results from two field campaigns suggest that groundwater collected in the streambed may provide reasonable estimates of apparent groundwater age, and that *MTT* can be determined from different age-dating tracers and from sampling with different groundwater collection devices. Coupled streambed point measurements of groundwater age and groundwater seepage rate represent a novel and effective approach to estimating aquifer *TTD* as well as *MTT*. *TTD* for groundwater discharging to the stream was very consistent between different point sampling arrangements and suggest that the *TTD* for some surficial groundwater systems may differ from the commonly-assumed exponential distribution for unconfined aquifers.

REFERENCES

Aeschbach-Hertig, W., and D. K. Solomon (2013), Noble gas thermometry in groundwater hydrology, in *The Noble Gases as Geochemical Tracers*, edited by P. Burnard, pp. 81–122, Springer, Heidelberg.

Aeschbach-Hertig, W., P. Schlosser, M. Stute, H. J. Simpson, A. Ludin, and J. F. Clark (1998), A $^{3}H/^{3}He$ study of ground water flow in a fractured bedrock aquifer, *Ground Water*, *36*(4), 661–670, doi:10.1111/j.1745-6584.1998.tb02841.x.

Aeschbach-Hertig, W., H. El-Gamal, M. Wieser, and L. Palcsu (2008), Modeling excess air and degassing in groundwater by equilibrium partitioning with a gas phase, *Water Resour. Res.*, *44*(8), W08449, doi:10.1029/2007WR006454.

Amin, I. E., and M. E. Campana (1996), A general lumped parameter model for the interpretation of tracer data and transit time calculation in hydrologic systems, *J. of Hydrol.*, *179*(1–4), 1–21, doi:10.1016/0022-1694(95)02880-3.

Andrews, L. C., and R. L. Phillips (2003), *Mathematical techniques for engineers and scientists*, SPIE - The International Society for Optical Engineering, Bellingham, WA.

Bethke, C. M., & Johnson, T. M. (2008), Groundwater age and groundwater age dating, *Ann. Rev. Earth Pl. Sc.*, *36*(1), 121–152, doi:10.1146/annurev.earth.36.031207.124210

Böhlke, J. K., & Denver, J. M. (1995), Combined use of groundwater dating, chemical and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, atlantic coastal plain, Maryland, *Water Resour. Res., 31*(9), 2319, doi:10.1029/95WR01584

Böhlke, J. K. (2002), Groundwater recharge and agricultural contamination, *Hydrogeol. J.*, *10*, 153–179, doi:10.1007/s10040-001-0183-3.

Böhlke, J. K., R. Wanty, M. Tuttle, G. Delin, and M. Landon (2002), Denitrification in the recharge area and discharge area of a transient agricultural nitrate plume in a glacial outwash sand aquifer, Minnesota, *Water Resour. Res.*, *38*(7), 10–1, doi:10.1029/2001WR000663.

Browne, B. A., and N. M. Guldan (2005), Understanding long-term baseflow water quality trends using a synoptic survey of the ground water-surface water interface, Central Wisconsin, *J. Environ. Qual.*, *34*(3), 825–35.

Bullister, J. L., Wisegarver, D. P., & Menzia, F. A. (2002), The solubility of sulfur hexafluoride in water and seawater, *Deep Sea Research Part I: Oceanographic Research Papers*, *49*(1), 175–187, doi:10.1016/S0967-0637(01)00051-6

Busenberg, E., & Plummer, L. N. (2000), Dating young ground water with sulfur hexafluoride: Natural and anthropogenic sources of sulfur hexafluoride, *Water Resour. Res., 36*(10), 3011–3030.

Cook, P. G., & Solomon, D. K. (1995), Transport of atmospheric trace gases to the water table: Implications for groundwater dating with chlorofluorocarbons and krypton 85, *Water Resour. Res., 31*(2), 263–270, doi:10.1029/94WR02232

Cook, P. G., & Solomon, D. K. (1997), Recent advances in dating young groundwater: chlorofluorocarbons, ^{3}H-^{3}He and ^{85}Kr. *J. Hydrol.*, *191*(1–4), 245–265, doi:10.1016/S0022-1694(96)03051-X

Darling, W. G., Gooddy, D. C., MacDonald, A. M., & Morris, B. L. (2012), The practicalities of using CFCs and SF_6 for groundwater dating and tracing. *Appl. Geochem.*, *27*(9), 1688–1697, doi:10.1016/j.apgeochem.2012.02.005

Dunkle, S. A., L. N. Plummer, E. Busenberg, P. J. Phillips, J. M. Denver, P. A. Hamilton, R. L. Michel, and T. B. Coplen (1993), Chlorofluorocarbons (CCl_3F and CCl_2F_2) as dating tools and hydrologic tracers in shallow groundwater of the Delmarva Peninsula, Atlantic Coastal Plain, United States, *Water Resour. Res.*, *29*(12), 3837–3860, doi:10.1029/93WR02073.

Ekwurzel, B., P. Schlosser, W. M. Smethie, L. N. Plummer, E. Busenberg, R. L. Michel, R. Weppernig, and M. Stute (1994), Dating of shallow groundwater: Comparison of the transient tracers $^{3}H/^{3}He$, chlorofluorocarbons, and ^{85}Kr, *Water Resour. Res., 30*(6), 1693–1708, doi:10.1029/94WR00156.

Ferronsky, V. I., & Polyakov, V. A. (c2012.), *Isotopes of the earth's hydrosphere [electronic resource]*. Dordrecht: Springer Science Business Media, [available at http://catalog.lib.ncsu.edu/record/NCSU2693065.]

Friedrich, R., Vero, G., von Rohden, C., Lessmann, B., Kipfer, R., & Aeschbach-Hertig, W. (2013), Factors controlling terrigenic SF_6 in young groundwater of the Odenwald region (Germany), *Appl. Geochem.*, *33*, 318–329, doi:10.1016/j.apgeochem.2013.03.002.

Frisbee, M. D., J. L. Wilson, J. D. Gomez-Velez, F. M. Phillips, and A. R. Campbell (2013), Are we missing the tail (and the tale) of residence time distributions in watersheds?, *Geophys. Res. Lett.*, *40*(17), 4633–4637, doi:10.1002/grl.50895.

Gardner, W. P., G. A. Harrington, D. K. Solomon, and P. G. Cook (2011), Using terrigenic ^{4}He to identify and quantify regional groundwater discharge to streams, *Water Resour. Res.*, *47*(6), doi:http://dx.doi.org.prox.lib.ncsu.edu/10.1029/2010WR010276.

Genereux, D. P., Leahy, S., Mitasova, H., Kennedy, C. D., & Corbett, D. R. (2008), Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA. *J. of Hydrol.*, *358*(3–4), 332–353, doi:10.1016/j.jhydrol.2008.06.017.

Genereux, D. P., M. Webb, and D. K. Solomon (2009), Chemical and isotopic signature of old groundwater and magmatic solutes in a Costa Rican rain forest: Evidence from carbon, helium, and chlorine, *Water Resour. Res.*, *45*(8), W08413, doi:10.1029/2008WR007630.

Gilmore, T. E., D. P. Genereux, D. K. Solomon, K. M. Farrell, and H. Mitasova (2015a), Aquifer nitrate legacy and dynamics observed from groundwater sampling in streambeds and well nests, *in prep*.

Gilmore, T. E., Genereux, D. P., Solomon, D. K., Solder, J. E., Kimball, B.A., Mitasova, H., & Birgand, F. (2015b), Quantifying the fate of agricultural nitrogen in an unconfined aquifer: stream-based observations at three measurement scales, *in prep*.

Hamilton, S. K. (2012), Biogeochemical time lags may delay responses of streams to ecological restoration." *Freshwater Biol.* 57: 43–57, doi:10.1111/j.1365-2427.2011.02685.x.

Happell, J. D., Opsahl, S., Top, Z., & Chanton, J. P. (2006), Apparent CFC and ^{3}H/^{3}He age differences in water from Floridan Aquifer springs, *J. of Hydrol.*, *319*(1–4), 410–426, doi:10.1016/j.jhydrol.2005.07.033.

Hrachowitz, M., C. Soulsby, D. Tetzlaff, I. A. Malcolm, and G. Schoups (2010), Gamma distribution models for transit time estimation in catchments: Physical interpretation of parameters and implications for time-variant transit time assessment, *Water Resour. Res.*, *46*(10), doi:http://dx.doi.org.prox.lib.ncsu.edu/10.1029/2010WR009148.

International Atomic Energy Agency (IAEA), (1992), Statistical treatment of isotope data in precipitation, *Technical Reports Series No. 206*, 255 pp., Vienna, Austria.

Ingraham, N. L. (2006), Isotopic variations in precipitation, in *Isotope tracers in catchment hydrology*, edited by C. Kendall and J. J. McDonnell, pp. 87–118, Elsevier, Boston, MA.

Kaufman, S., & Libby, W. F. (1954), The natural distribution of tritium, *Phy. Rev.*, *93*(6), 1337–1344, doi:10.1103/PhysRev.93.1337.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2007), Design of a light-oil piezomanometer for measurement of hydraulic head differences and collection of groundwater samples, *Water Resour. Res., 43*, W09501, doi:10.1029/2007WR005904.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2009a), Relationships among groundwater age, denitrification, and the coupled groundwater and nitrogen fluxes through a streambed, *Water Resour. Res., 45*, W09402, doi:10.1029/2008WR007400.

Kennedy, C. D., Genereux, D. P., Corbett, D. R., & Mitasova, H. (2009b), Spatial and temporal dynamics of coupled groundwater and nitrogen fluxes through a streambed in an agricultural watershed, *Water Resour. Res., 45*, W090401, doi:10.1029/2008WR007397.

Kennedy, C. D., Genereux, D. P., Mitasova, H., Corbett, D. R., & Leahy, S. (2008), Effect of sampling density and design on estimation of streambed attributes, *J. Hydrol.*, *355*, 164–180, doi:10.1016/j.jhydrol.2008.03.018.

King, D. B., & Saltzman, E. S. (1995), Measurement of the diffusion coefficient of sulfur hexafluoride in water, *J. of Geophy. Res.: Oceans*, *100*(C4), 7083–7088, doi:10.1029/94JC03313.

Kirchner, J. W., X. Feng, and C. Neal (2000), Fractal stream chemistry and its implications for contaminant transport in catchments, *Nature*, *403*(6769), 524–527, doi:10.1038/35000537.

Kirchner, J. W., D. Tetzlaff, and C. Soulsby (2010), Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments, *Hydrol. Process.*, *24*(12), 1631–1645, doi:10.1002/hyp.7676.

Koh, D. C., Niel Plummer, L., Kip Solomon, D., Busenberg, E., Kim, Y. J., & Chang, H. W. (2006), Application of environmental tracers to mixing, evolution, and nitrate contamination of ground water in Jeju Island, Korea, *J. Hydrol.*, *327*(1–2), 258–275, doi:10.1016/j.jhydrol.2005.11.021.

Lindsey, B. D., S. W. Phillips, C. A. Donnelly, G. K. Speiran, L. N. Plummer, J. K. Böhlke, M. J. Focazio, W. C. Burton, and E. Busenberg (2003), Residence times and nitrate transport in ground water discharging to streams in the Chesapeake Bay watershed, *Water-Resources Investigations Report 03-4035*, U.S. Geol. Surv., Reston, VA.

McCallum, J. L., Cook, P. G., & Simmons, C. T. (2014a), Limitations of the use of environmental tracers to infer groundwater age, *Groundwater*, n/a–n/a, doi:10.1111/gwat.12237.

McCallum, J. L., Cook, P. G., Simmons, C. T., & Werner, A. D. (2014b), Bias of apparent tracer ages in heterogeneous environments. *Groundwater*, *52*(2), 239–250, doi:10.1111/gwat.12052.

Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010), Lag time in water quality response to best management practices: A review, *J. Environ. Qual.*, *39*(1), 85, doi:10.2134/jeq2009.0108.

Modica, E., Buxton, H. T., & Plummer, L. N. (1998), Evaluating the source and residence times of groundwater seepage to streams, New Jersey Coastal Plain, *Water Resour. Res., 34*, 2797. doi:10.1029/98WR02472.

Molénat, J., Gascuel-Odoux, C., Aquilina, L., & Ruiz, L. (2013), Use of gaseous tracers (CFCs and SF_6) and transit-time distribution spectrum to validate a shallow groundwater transport model, *J. Hydrol.*, *480*, 1–9. doi:10.1016/j.jhydrol.2012.11.043.

Newman, B. D., Osenbrück, K., Aeschbach-Hertig, W., Kip Solomon, D., Cook, P., Różański, K., & Kipfer, R. (2010), Dating of "young" groundwaters using environmental tracers: advantages, applications, and research needs, *Isotopes in Environmental and Health Studies*, *46*(3), 259–278, doi:10.1080/10256016.2010.514339.

Pint, C. D., Hunt, R. J., & Anderson, M. P. (2003), Flowpath delineation and ground water age, Allequash Basin, Wisconsin, *Groundwater*, *41*(7), 895–902, doi:10.1111/j.1745-6584.2003.tb02432.x.

Plummer, L. N., & Busenberg, E. (2000), Chloroflurocarbons, in P. G. Cook & A. Herczeg (Eds.), *Environmental tracers in subsurface hydrology*, pp. 441–478, Kluwer Academic Publishers, Boston, MA.

Plummer, L. N., Busenberg, E., Böhlke, J. K., Nelms, D. L., Michel, R. L., & Schlosser, P. (2001), Groundwater residence times in Shenandoah National Park, Blue Ridge Mountains, Virginia, USA: a multi-tracer approach, *Chemical Geology*, *179*(1–4), 93–111, doi:10.1016/S0009-2541(01)00317-5.

Plummer, L.N., Busenberg, E., & Cook, P. G. (2006), Principles of chlorofluorocarbon dating. In E. Busenberg (Ed.), *Use of Chlorofluorocarbons in Hydrology: A Guidebook*, pp. 17–30, International Atomic Energy Agency (IAEA), Vienna, Austria.

Poreda, R. J., Cerling, T. E., & Salomon, D. K. (1988), Tritium and helium isotopes as hydrologic tracers in a shallow unconfined aquifer, *J. Hydrol.*, *103*(1–2), 1–9, doi:10.1016/0022-1694(88)90002-9.

Portniaguine, O., & Solomon, D. K. (1998), Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts. *Water Resour. Res., 34*(4), 637–645, doi:10.1029/97WR03361.

Price, R. M., Top, Z., Happell, J. D., & Swart, P. K. (2003), Use of tritium and helium to define groundwater flow conditions in Everglades National Park, *Water Resour. Res., 39*(9), 1267, doi:10.1029/2002WR001929.

Puckett, L. J. (2004), Hydrogeologic controls on the transport and fate of nitrate in ground water beneath riparian buffer zones: results from thirteen studies across the United States, *Water Science and Technology*, *49*(3), 47.

Puckett, L. J., & Cowdery, T. K. (2002), Transport and fate of nitrate in a glacial outwash aquifer in relation to ground water age, land use practices, and redox processes, *J. Environ. Qual.*, *31*(3), 782.

Puckett, L. J., Tesoriero, A. J., & Dubrovsky, N. M. (2011), Nitrogen contamination of surficial aquifers - A growing legacy, *Environ. Sci. Technol.*, *45*(3), 839–844, doi:10.1021/es1038358.

Reilly, T. E., L. N. Plummer, P. J. Phillips, and E. Busenberg (1994), The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer, *Water Resour. Res.*, *30*(2), 421–433, doi:10.1029/93WR02655.

Sanford, W. (2011), Calibration of models using groundwater age, *Hydrogeol. J.*, *19*(1), 13–16, doi:10.1007/s10040-010-0637-6.

Sanford, W. E., & Pope, J. P. (2013). Quantifying groundwater's role in delaying improvements to Chesapeake Bay water quality, *Environ. Sci. Technol.*, *47*(23), 13330–13338, doi:10.1021/es401334k.

Schlosser, P., Stute, M., Dörr, H., Sonntag, C., & Münnich, K. O. (1988). Tritium/^{3}He dating of shallow groundwater, *Earth Planet Sc. Lett.*, *89*(3–4), 353–362, doi:10.1016/0012-821X(88)90122-7.

Shapiro, S. D., LeBlanc, D., Schlosser, P., & Ludin, A. (1999). Characterizing a sewage plume using the ^{3}H-^{3}He dating technique. *Ground Water*, *37*(6), 861–878, doi:10.1111/j.1745-6584.1999.tb01185.x.

Solder, J. E. E. (2014), Quantifying groundwater-surface water exchange: development and testing of shelby tubes and seepage blankets as discharge measurement and sample collections devices, M.S. thesis, Univ. of Utah, Salt Lake City, UT. [available at: http://content.lib.utah.edu/cdm/singleitem/collection/etd3/id/3170/rec/1.]

Solomon, D. K., Gilmore, T. E., Solder, J. E., Kimball, B. A., & Genereux, D. P. (2015), Evaluating a groundwater flow system using SF_6 in streamflow, *in prep*.

Solomon, D. K., Cook, P. G., & Plummer, L. N. (2006), Models of groundwater ages and residence times. In E. Busenberg (Ed.), in *Use of Chlorofluorocarbons in Hydrology: A Guidebook*, pp. 73–78, International Atomic Energy Agency (IAEA), Vienna, Austria.

Solomon, D. K., Schiff, S. L., Poreda, R. J., & Clarke, W. B. (1993), A validation of the $^3H/^3He$ method for determining groundwater recharge, *Water Resour. Res., 29*(9), 2951–2962, doi:10.1029/93WR00968.

Solomon, D. K., & Sudicky, E. A. (1991), Tritium and Helium 3 isotope ratios for direct estimation of spatial variations in groundwater recharge. *Water Resour. Res., 27*(9), 2309–2319, doi:10.1029/91WR01446.

Stolp, B. J., D. K. Solomon, A. Suckow, T. Vitvar, D. Rank, P. K. Aggarwal, and L. F. Han (2010), Age dating base flow at springs and gaining streams using helium-3 and tritium: Fischa-Dagnitz system, southern Vienna Basin, Austria, *Water Resour. Res.*, *46*, W07503, doi:201010.1029/2009WR008006.

Stute, M., Deák, J., Révész, K., Böhlke, J. K., Deseö, E., Weppernig, R., & Schlosser, P. (1997), Tritium/^{3}He dating of river infiltration: an example from the Danube in the Szigetköz area, Hungary, *Ground Water*, *35*(5), 905–911, doi:10.1111/j.1745-6584.1997.tb00160.x.

Tesoriero, A. J. (2005), Nitrogen transport and transformations in a coastal plain watershed: Influence of geomorphology on flow paths and residence times, *Water Resour. Res., 41*(2), doi:10.1029/2003WR002953.

Visser, A., Broers, H. P., & Bierkens, M. F. P. (2007), Dating degassed groundwater with $^3H/^3He$, *Water Resour. Res., 43*(10), W10434, doi:10.1029/2006WR005847.

Visser, A., Broers, H. P., Purtschert, R., Sültenfuß, J., & de Jonge, M. (2013). Groundwater age distributions at a public drinking water supply well field derived from multiple age tracers (^{85}Kr, $^3H/^3He$, and ^{39}Ar), *Water Resour. Res., 49*(11), 7778–7796, doi:10.1002/2013WR014012.

Visser, A., Fourré, E., Barbecot, F., Aquilina, L., Labasque, T., Vergnaud, V., & Esser, B. K. (2014), Intercomparison of tritium and noble gases analyses, $^3H/^3He$ ages and derived parameters excess air and recharge temperature, *Appl. Geochem.*, *50*, 130–141, doi:10.1016/j.apgeochem.2014.03.005.

Vogel, J. C. (1967), Investigation of groundwater flow with radiocarbon, in *Isotopes in Hydrology*, pp. 355–369, International Atomic Energy Agency, Vienna, Austria.

Weeks, E. P., D. W. Ericson, and C. L. R. Holt (1965), Hydrology of the Little Plover River basin, Portage County, Wisconsin, and the effects of water resources development, *Water Supply Paper 1811*, U.S. Geol. Surv., Reston, VA.

Weiss, R. F. (1970), The solubility of nitrogen, oxygen and argon in water and seawater, *Deep-Sea Res. and Oceanographic Abstracts*, *17*(4), 721–735, doi:10.1016/0011-7471(70)90037-9.

Weiss, R. F. (1971), Solubility of helium and neon in water and seawater, *J. Chem. Eng. Data*, *16*(2), 235–241, doi:10.1021/je60049a019.

Welch, H. L., Green, C. T., & Coupe, R. H. (2011), The fate and transport of nitrate in shallow groundwater in northwestern Mississippi, USA, *Hydrogeol. J.*, *19*(6), 1239–1252, doi:10.1007/s10040-011-0748-8.

Winner, M. D., and R. W. Coble (1996), Hydrogeologic framework of the North Carolina coastal plain, *Professional Paper 1404-I*, U.S. Geol. Surv., Reston, VA.

Wise, D. L., & Houghton, G. (1966), The diffusion coefficients of ten slightly soluble gases in water at 10–60°C, *Chemical Engineering Science*, *21*(11), 999–1010, doi:10.1016/0009-2509(66)85096-0.

CHAPTER 3

AQUIFER NITRATE LEGACY AND DYNAMICS OBSERVED FROM GROUNDWATER SAMPLING IN STREAMBEDS AND WELL NESTS

3.1. Abstract

Fluxes of nitrate from groundwater systems may contribute to lag times in water quality response to nutrient management initiatives in agricultural watersheds. We investigated the magnitudes of groundwater contamination, denitrification extent, and groundwater transit times in a surficial aquifer in the coastal plain of North Carolina, USA, by sampling groundwater in a streambed (8 transects in 58 m reach and 6 transects in 2.5 km reach) and in well nests (two 3-well nests). All approaches indicated elevated groundwater nitrate concentrations (210 – 808 μM) despite the loss by denitrification of 51–78% of all nitrate entering the aquifer. Streambed sampling showed patterns of groundwater [NO_3^-] and extent of denitrification relating to recharge areas along the two sides of the stream, which could be useful for targeting high-priority areas for nutrient management or remediation. Age-dating tracer data from both streambed and well nest sampling suggested some areas of little or no recharge the watershed, based on age distributions with relatively little young groundwater compared to the commonly-applied exponential age distribution model. This finding has potentially important implications for predicting future nitrate discharge flux (f_{NO3}) from the aquifer. Well data were consistent with an exponential-piston flow model and suggested a lag of about a decade before f_{NO3} would respond to lower N use in the watershed. Using N concentration data, and the observed age distribution from streambed sampling (which was consistent with a gamma distribution with large shape parameter, $\alpha = 18$), to predict f_{NO3} from the aquifer suggested that even with no N use in the watershed f_{NO3} may increase in the next 1-2 decades before declining. The different predictions from well data and streambed data highlight the potential importance of sampling at the aquifer outlet (e.g., the streambed)

to determine the actual age distribution of groundwater discharge. The time required for f_{NO3} to stabilize is about 1-2 times the length of groundwater mean transit times estimated from streambed (29-31 years) and well nest sampling (27 years). All sampling approaches suggested that N use may have been lower in the last decade, but based on groundwater NO_3^- mass balance calculations the storage of NO_3^- in the aquifer was increasing modestly at the time of sampling, at a rate of about 28 mmol m^{-2} yr^{-1} (about one-fifth of annual NO_3^- input to the aquifer).

3.2. Introduction

Significant amounts of agricultural nitrogen (N) have been temporarily stored in surficial aquifers since the mid-20th century, due to increasing use of N-based fertilizers and the lag between N recharge to groundwater systems and eventual discharge of that N to streams [e.g., *Puckett et al.*, 2011]. Groundwater discharge generally represents a large proportion of annual streamflow [e.g., *Lindsey et al.*, 2003] and the flux of N from groundwater to streams poses a potential threat to surface water quality. Quantifying the transfer of N from groundwater to surface water is critical for understanding the impact of groundwater discharge on stream water quality.

The agricultural N legacy (present-day contamination from N release in past years) has been observed using well nests in recharge areas [e.g., *Böhlke and Denver*, 1995; *Böhlke*, 2002; *Böhlke et al.,* 2002, 2007; *Green et al.*, 2008]. In some cases, wells are installed along transects oriented along groundwater flow paths from recharge to discharge. Such transect sampling has led to insights regarding links between age and N concentrations in groundwater and changes in redox status along flowlines in N contaminated groundwater. In other cases, a single well nest has been used to estimate groundwater recharge rates and the vertical profile of N contamination in the surficial aquifer [e.g., *Welch et al.*, 2011].

Several studies have used groundwater sampling in both the recharge area and in streams to show that N use in agricultural areas has affected groundwater on a decadal time scale

[e.g., *Böhlke and Denver*, 1995; *Böhlke et al., 2007*; *Spruill et al.*, 2004; *Tesoriero et al.*, 2005]. *Modica et al*. [1998] and *Lindsey et al.* [2003] sampled groundwater along transects across streambeds, and results suggested higher nitrate concentrations in younger groundwater that entered the stream closer to the banks. *Modica et al.* [1998] also delineated source areas for sampled transects using a particle tracking model and showed longer flow paths to transects with older groundwater. In two headwater streams *Böhlke and Denver* [1995] found examples of greater groundwater age toward the centers of the streams and lower nitrate concentrations in older groundwater, which was consistent with increasing groundwater age with depth in the surficial aquifer. Differences in nitrate concentrations in the two streams were attributed to a reducing layer of sediments that was shallower beneath the stream with lower nitrate and thus influenced a larger fraction of the groundwater flow toward that stream. These studies showed direct linkages between groundwater age and nitrate in surficial aquifers and beneath streams, but did not directly evaluate fluxes of N into the stream or define the distribution of groundwater age in aquifer discharge to the stream.

Few studies have focused on extensive sampling of groundwater in streambeds to determine flow-weighted distributions of both groundwater nitrate and apparent age [*Browne and Guldan*, 2005; *Kennedy et al.*, 2009a]. *Browne and Guldan* [2005] sampled groundwater discharge in the thalweg of a meandering sand plains stream in Wisconsin and used the groundwater age and nitrate distributions to reconstruct historic baseflow concentrations. Groundwater accounted for about 90% of annual discharge to the stream and in-stream denitrification was apparently negligible, which simplified the analysis. In a study focused solely on streambed sampling in a channelized stream located in the Neuse River basin in North Carolina, *Kennedy et al.* [2009a] developed a streambed map of groundwater apparent age, showing older water discharging toward the center of the stream. Groundwater nitrate concentrations (both corrected and uncorrected for denitrification in the surficial aquifer) were also lower toward the center of the stream, consistent with lower N use in the watershed when this older groundwater was recharged. The majority of water entering the stream was 20-40 years old, which was inconsistent with the exponential age distribution that has been

used to approximate the age-depth relationships in some unconfined aquifers (e.g., *Böhlke* [2002], their Fig. 3.5). It was hypothesized that sampling closer to the stream banks might account for a greater amount of younger water discharging into the stream. The flow-weighted mean apparent age of groundwater seeping through the streambed (30 years) gave an indication of the time scale for flushing of the connected surficial aquifer, which may be relevant to understanding an observed lag in water quality response to the nutrient management initiative in the Neuse River basin [e.g., *Lebo et al.*, 2010; *Meals et al*., 2012].

In this study we link groundwater nitrate and apparent age results from well nests with results from detailed streambed sampling to characterize the agricultural nitrogen legacy in groundwater. Specifically, the goal of the study was to:

- quantify the magnitude of nitrate contamination and denitrification in the surficial aquifer
- investigate potential controls on nitrate fluxes to and from the surficial aquifer (e.g., groundwater mean transit time and spatial patterns in recharge)
- estimate the balance between nitrate input to and loss from the surficial aquifer (i.e., whether nitrate pollution in the aquifer is increasing, decreasing, or holding steady)
- predict future nitrate flux from the surficial aquifer.

We measured nitrate concentrations in groundwater from two well nests and from two streambed sampling campaigns. The amount of denitrification that had occurred in the surficial aquifer was determined by modeling excess N_2 in groundwater samples. Groundwater apparent age was also modeled for each groundwater sample. The streambed sampling campaigns and results for individual streambed point samples (apparent age, N concentrations) are described in detail in two additional papers on the same study site in North Carolina, United States [*Gilmore et al.*, 2015a,b].

3.3. Study Site and Hydrologic Conditions

The Bear Creek watershed lies in the middle coastal plain in North Carolina, in the same hydrogeologic subregion (subregion 4, Ator et al., 2005) where other work on N contamination has occurred [e.g., *Böhlke and Denver*, 1995; *Spruill et al.*, 2004; *Tesoriero et al.*, 2005]. About 50% of the land in the watershed is used for row crop production and intensive animal production (poultry, swine). The landscape in Bear Creek watershed consists of floodplain with poorly-drained soils transitioning to nearly flat uplands which contain mostly moderately well-drained to excessively well-drained soils, but with pockets of poorly-drained to somewhat poorly-drained soils, particularly near the watershed divide. This landscape is typical for the middle coastal plain [*Daniels et al.*, 1999], which makes up roughly a third of the mid-Atlantic coastal plain.

Nitrate is the predominant form of N in groundwater in this region [*Tesoriero et al.*, 2004; *Kennedy et al.*, 2009a; *Gilmore et al.*, 2015b]. Soils contain up to 5% organic C and dissolved organic C is commonly < 333 μM (< 4 mg/L) [*Tesoriero et al.*, 2004]. Significant levels of denitrification along groundwater flow paths (>50% removal of NO_3^-) have been observed in the middle coastal plain ([e.g., *Böhlke and Denver*, 1995; *Spruill et al.*, 2004; *Tesoriero et al.*, 2005; *Kennedy et al.*, 2009a; *Gilmore et al.*, 2015b]).

The streambed sampling site was in West Bear Creek (Fig. 3.1), a channelized and entrenched stream that drains into Bear Creek, which then drains to the Neuse River. The stream is about 7 m wide, with a sandy streambed containing 2-3% organic matter [*Elkins*, 2007]. Agricultural drainage has been implemented through much of the Bear Creek watershed, mainly surface drainage in the floodplain adjacent to the West Bear Creek study reach (Fig. 3.1) but some subsurface drains that discharge directly to Bear Creek in the floodplain down-gradient of the BC2 well nest (field observations; J. Gray, local farmer, personal communication, January 2015). The fraction of annual stream discharge attributed to baseflow (i.e., base flow index) in this region is about 0.4 [*Wolock*, 2003] and N application rates are high [*Ruddy et al.*, 2006] which suggests that streams are vulnerable to the effects of legacy groundwater contamination [*Tesoriero et al.*, 2013]. Large export of N from Bear Creek watershed (810 kg N km^{-2}) was reported previously by *Usry* [2006].

Well nests were installed at two locations in Bear Creek watershed (Fig. 3.1), outside of the topographically defined contributing area to West Bear Creek but in areas that are likely representative of recharge near West Bear Creek (e.g., predominantly agricultural land use in the upland area).

A geologic core ("Geoprobe core", Fig. 3.1) extracted along the left bank of West Bear Creek revealed a low permeability mud layer between about 1.6 m to 2.8 m depth below ground surface [*Elkins*, 2007]. Below the mud layer was a sand unit about 7.5 m thick, which was underlain by a mud unit that was interpreted as the lower confining layer for the surficial aquifer [*Kennedy et al.*, 2009a], possibly the Black Creek confining unit [*Winner and Coble*, 1996]. Cores extracted during installation of one well nest (BC1, Fig. 3.1) also showed a shallow low permeability layer from about 1.2 m to 1.9 m (Fig. 3.2), while at the second well nest (BC2) site, interbedded sand and mud was observed between 2.4 to 3.9 m. The water table observed during groundwater sampling and/or drilling was roughly at the base of the shallow low permeability layer (BC1, Geoprobe core) or within it (BC2). The cretaceous Black Creek confining unit was observed at 17.7 m below land surface (about 16.0 m below the water table) at the BC1 well nest site. Depth to the confining unit was assumed to be about 17.7 m at the BC2 site. Topographical relief between the BC1 well site and the streambed in West Bear Creek is roughly 15 meters.

West Bear Creek discharge differed by about an order of magnitude during the two streambed sampling campaigns (57 L/s in July 2012 and 500 L/s in March 2013, based on discharge calculated at the 200 m station in Fig. 3.1 from a steady bromide injection to the stream [*Gilmore et al.*, 2015b]). Water table elevation in a nearby NC Division of Water Resources monitoring well was near 30-year monthly median values during both well nest sampling (June 2013) and streambed sampling (Table 3.1) [*Gilmore et al.*, 2015b].

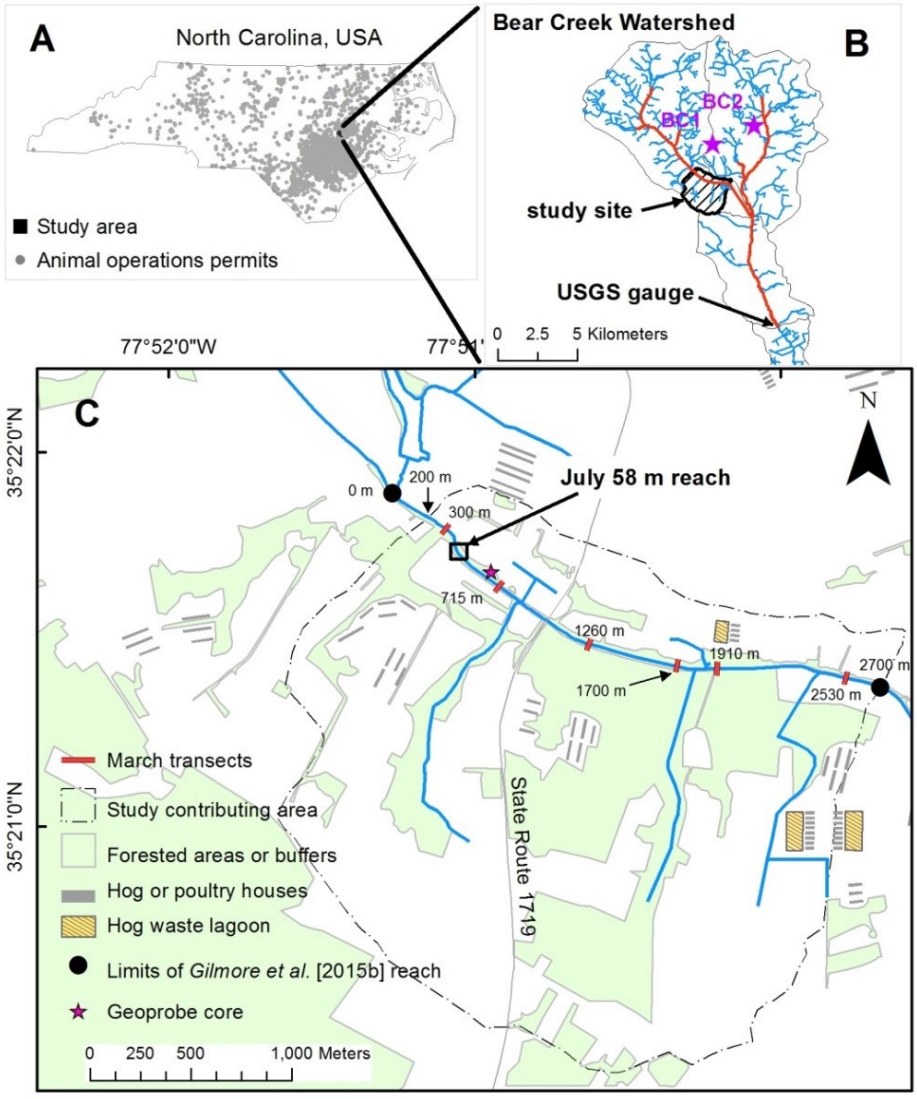

Figure 3.1. Study area map showing streambed point sampling and recharge area well nest locations. A: study area location in eastern North Carolina. B: Well nests (denoted by stars) located in the Bear Creek watershed, just to the east of the West Bear Creek watershed (outlined by gray line), the topographically-defined contributing area for the 2.5 km study reach (cross-hatched area), and location of USGS stream gauge on Bear Creek. The main channel of Bear Creek and West Bear Creek are highlighted in red (see discussion in Section 5.1). C: All streambed sampling occurred between the 200 m and 2700 m stations (roughly 2.5 km reach). Eight streambed point transects were sampled in the "July 58 m reach". In March 2013, six transects were widely spaced, as shown by red lines hash marks along the stream, labeled by distance downstream from the 0 m station. All GIS data were accessed via the NC OneMap Geospatial Portal (data.nconemap.com), with the exception of the detailed hydrography data in inset B [CGIA, 2004]. Forested areas, agricultural facilities, and

tributaries were defined using digital orthophotos (2010 North Carolina Statewide Digital Orthoimagery) and field observations. Digital elevation data from the North Carolina Division of Transportation were used to defined the 2.5 km contributing area. The Bear Creek watershed is from the USDA NC NRCS 12-Digit Hydrologic Units data set. North Carolina Division of Environment and Natural resources data sets were utilized to show the main channel of West Bear Creek (1:24,000 scale hydrography) and the locations for animal operations permits. The Geoprobe core location was from *Elkins* [2007].

Table 3.1. Water table (WT) elevations at NC Division of Water Resources monitoring well, and Bear Creek stream discharge (Q) at USGS gauge.

sampling date	30-year median WT elevation[e]	WT elevation during sampling[f]	Q at USGS gauge[g]
	m	*m*	$m^3 s^{-1}$
June 2013[a]	27.4	27.3	3.6
July 2012[b]	27.5	27.4	0.5
March 2013[c]	27.9	27.7	1.8
April 2007[d]	27.8	27.6	0.8[h]

[a]wells
[b]58 m reach
[c]2.5 km reach
[d]*Kennedy et al.* [2009a], 75 m reach
[e]historic median groundwater level (depth below land surface) at Saulston, NC for the month when sampling occurred; well O30J4, Saulston well site, 12 km northwest of the study site, http://www.ncwater.org/
[f]groundwater level in the Saulston, NC well during roughly the middle of the sampling period
[g]Bear Creek stream discharge at USGS stream gauge at Mays Store, NC during the middle of the sampling period (http://waterdata.usgs.gov/nwis/uv?site_no=0208925200)
[h]*Kennedy et al.* [2009a]

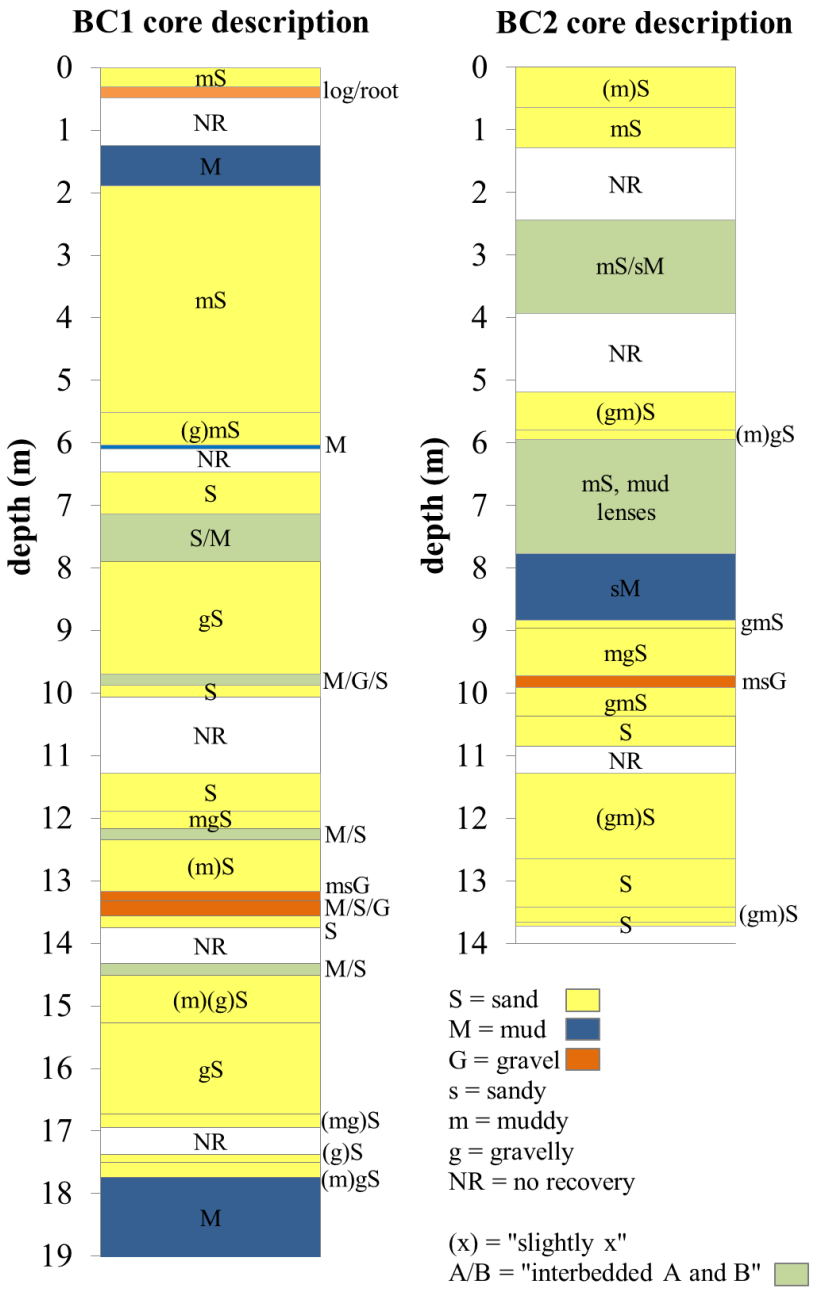

Figure 3.2. BC1 and BC2 core descriptions, with texture class codes from *Farrell et al.* [2012]. Textural class codes list classes in order of abundance, e.g., "(m)gS" indicates slightly muddy, gravelly sand where "slightly" indicates <5% by volume.

3.4. Methods

Two well nests (BC1 and BC2, each with 3 wells) were installed on April 9-10, 2013 (Fig. 3.1). Depth from land surface to the middle of the 45 cm well screens was 6.1 m, 10.2 m, and 16.5 m at BC1 and 5.4 m, 7.7 m, and 14.2 m at BC2. Water table depth was about 1.8 m and 2.7 m below ground surface at BC1 and BC2, respectively, when groundwater samples were collected on June 19-21, 2013.

Groundwater samples from wells and from the streambed were analyzed for ^3H, NO_3^-, sulphur hexafluoride (SF_6), and other dissolved gases (N_2, He, Ne, Ar, Kr, Xe).

The wells were purged before sampling by pumping at least 3x the water volume in the well (the PVC screen and casing and the sand-packed annulus around the screen). Groundwater samples for dissolved gases and ^3H were collected using an inertial pump (Waterra® check-valve installed on the bottom of the sampling line). Samples for analysis of NO_3^- were collected using a peristaltic pump.

Streambed sampling is described in detail by *Gilmore et al.* [2015a,b]. Piezomanometers [*Kennedy et al.*, 2007] with 5 cm screens were inserted to 31 cm depth (top of screen) in the streambed and after purging and measurement of vertical head gradient (J) groundwater samples were extracted by syringe (^3H and NO_3^-) or peristaltic pump (SF_6, N_2, and Ar; USGS method, http://water.usgs.gov/lab). A second piezometer was inserted into the streambed within about 10 cm of the piezomanometer and an inertial pump was used to sample groundwater for noble gas analysis (He, Ne, Ar, Kr, Xe). Vertical hydraulic conductivity (K) was also measured at each streambed sampling location [*Genereux et al.*, 2008], and a vertical groundwater flux (seepage rate, v) was calculated as $v = KJ$. With v known, fluxes of NO_3^- from the surficial aquifer to the stream were calculated as $f_{NO3} = v[NO_3^-]$ (where the brackets denote aqueous concentration).

^3H samples were collected in 500 mL HDPE bottles and noble gas samples were stored in copper tubes sealed with steel pinch clamps [*Aeschbach-Hertig and Solomon,* 2013]. ^3H and

noble gas samples (and SF_6 samples from the well nests only) were analyzed at the Dissolved and Noble Gas Laboratory at the University of Utah in Salt Lake City, UT (analytical methods for noble gases are described in [*Solomon et al.*, 2015]; ^3H was determined by helium ingrowth (http://www.noblegaslab.utah.edu/tritium.html)). SF_6, N_2, and Ar samples were collected in glass bottles and analyzed at the USGS CFC lab in Reston, VA (http://water.usgs.gov/lab/). For the SF_6 samples collected from the wells, the average SF_6 concentrations from lab analyses at the two different labs were used to determine apparent groundwater age.

3.5. Modeling

3.5.1. Apparent age and denitrification

The closed equilibrium (CE) model [*Aeschbach-Hertig et al.,* 2008] was fit to noble gas data (Ar and Ne for groundwater from streambed sampling, Ar, Ne, Kr, and Xe for groundwater from wells) to model tritiogenic He (He_{trit}), SF_6, and N_2 concentrations [*Gilmore et al.,* 2015a,b]. Only Ar and Ne were used for streambed groundwater samples because Kr and Xe were injected into the stream as part of a reach mass-balance experiment that was concurrent with much of the streambed sampling period. The model formulation used was:

$$C_{i-\mathrm{mod}} = \frac{C_i^{eq}(1 + AH_{i-rech})}{(1 + BH_{i-sam})} \tag{3.1}$$

where $C_{i\text{-}mod}$ is the modeled concentration of gas i (Ne, Ar), C_i^{eq} is the solubility equilibrium concentration of gas i at recharge conditions (recharge temperature, atmospheric pressure, and recharge salinity), $H_{i\text{-}rech}$ and $H_{i\text{-}sam}$ are the Henry's Law constants for gas i at recharge conditions and discharge conditions, respectively, and A and B are gas-to-water volume ratios in pore space at recharge and discharge, respectively [*Aeschbach-Hertig et al.,* 2008]. Equation 3.1 was applied according to *Gilmore et al.*[2015a,b], after classifying each

groundwater sample as either "degassed" [e.g., *Visser et al.*, 2007; *Aeschbach-Hertig et al.*, 2008] or having "excess air" [e.g., *Heaton and Vogel*, 1981]. Degassed samples have noble gas concentrations below C_i^{eq} (i.e., gases have been removed from the groundwater by bubble formation), while samples with excess air have noble gas concentrations greater than C_i^{eq}. The recharge temperature (12.8° C, 1σ = 1.7° C) used in equation 3.1 was modeled from noble gas concentrations in groundwater samples from the wells (i.e., noble gas thermometry [*Aeschbach-Hertig and Solomon*, 2013]); details are found in *Gilmore et al.* [2015a,b].

Groundwater apparent age (τ), the travel time from recharge to sampling assuming piston-flow transport, was determined by converting groundwater [SF_6] values to atmospheric mixing ratios (x_i, pptv) (after *Friedrich et al.* [2013], *Solomon et al.* [2010]):

$$x_{SF_6} = \frac{[SF_6]_{sam} K_{rech}(1+BH_{sam})}{\left(P_a - p_{H_2O}\right)(1+AH_{rech})} \tag{3.2}$$

where K_{rech} is the Henry's Law constant at the recharge temperature and salinity in units of kg-atm/mol. P_a is the atmospheric pressure at the recharge elevation (m) and temperature (°C), and p_{H20} is the water vapor pressure at recharge temperature and salinity. The SF_6 apparent age was determined by matching x_{SF6} from equation 3.2 to historic atmospheric mixing ratios, based on the assumption that groundwater was at equilibrium with the atmosphere at recharge, and subtracting the recharge year from the sampling year (taken as 2012.5 for July 2012, 2013.0 for March 2013, and 2013.5 for June 2013, the months of our three sampling campaigns).

The $^3H/^3He$ method of age dating was also used for groundwater samples [e.g., *Poreda et al.*, 1988]. Based on previous work [*Schlosser et al.*, 1988; *Solomon et al.*, 1993; *Aeschbach-Hertig et al.*, 2008], [$^3He_{trit}$] was calculated as:

$$\left[{}^3He_{trit}\right] = \left(\left[{}^4He_{meas}\right]\left(R_{meas} - R_{terr}\right) - \left[{}^4He_{mod}\right]\left(\alpha R_{atm} + R_{terr}\right)\right)\left(1+BH_{sam}\right) \tag{3}$$

where the subscripts "*meas*", "*mod*", and "*terr*" represent measured, modeled, and terrigenic concentrations of ^{3}He. R is the [^{4}He]/[^{3}He] ratio, α is the isotope fractionation factor (0.983) for ^{3}He and ^{4}He (R_{gas}/R_{water}). An R_{terr} of 2.0 x 10^{-8} was assumed [*Schlosser et al.*, 1988; *Solomon et al.*, 1993]. The factor ($1 + BH_{sam}$) corrects for degassing, under the assumption that degassing occurred near the stream (i.e., degassing occurred *after* decay of ^{3}H to ^{3}He$_{trit}$ in the groundwater [*Aeschbach-Hertig et al.*, 2008]).

The parameters A and B from the CE model were also used to calculate the concentration of N_2 in groundwater samples that was produced by denitrification, [N_2-den]:

$$[N_2 - den] = \left([N_2]_{meas} - [N_2]_{\mathrm{mod}}\right)(1 + BH_N) \tag{4}$$

Initial nitrate concentrations (what observed [NO_3^-] would have been in the absence of denitrification) were calculated on a molar basis as $[NO_3^-]^0 = [NO_3^-] + 2[N_2\text{-den}]$.

3.5.2. Age distributions and groundwater mean transit time

Groundwater transit time is the travel time through the groundwater system from recharge to discharge, and mean transit time (*MTT*) represents the average across all flowlines in the groundwater system. *MTT* was calculated as the flow-weighted apparent age: $MTT = \Sigma\tau v/\Sigma v$ [*Kennedy et al.*, 2009a]. The distribution of groundwater age from streambed sampling was calculated by categorizing the samples into 10-year age bins (1-10 yr, 10.1-20 yr, etc.) and summing the v measured at the streambed sampling locations associated with samples in each bin (Σv_{bin}). Then the fraction of total groundwater discharge within each bin was calculated as $\Sigma v_{bin}/\Sigma v$. By definition, the age distributions from streambed sampling are transit time distributions, *TTD*s (transit time = age at the aquifer discharge face, e.g., the streambed). Age distributions from well nests are not, however, *TTD*s (wells are not at the aquifer outlet), so for the sake of simplicity the term "age distribution" is used throughout this paper for both streambed sampling and well nest sampling results.

The relationship between apparent age and depth in uniform unconfined aquifers with uniform recharge may approximate an exponential age distribution (EM) [*Vogel*, 1967; *Solomon et al.*, 2006]:

$$\tau = \frac{L\theta}{R}\ln\left(\frac{L}{L-z}\right) \tag{3.5}$$

where R is the rate of recharge, L is the aquifer thickness, z is the depth below the water table, and θ is the porosity. Equation 3.5 was used to calculate the R value that gave the best fit to data on apparent age versus depth from well sampling. With R known, MTT was calculated as $MTT = L\theta/R$ [*Solomon et al.*, 2006].

Another idealized model is the exponential-piston flow model (EPM), where water is recharged into an unconfined aquifer and then flows into a confined section of the same aquifer. The age distribution in the unconfined portion of the aquifer is equivalent to that of the exponential model (described above). As groundwater flows through the confined portion, no new recharge (i.e., no modern water) is added to the system. As a result, all of the groundwater increases in age in proportion to the distance the groundwater flows in the confined section of the aquifer. The age distribution in the confined section is calculated as:

$$\tau = \frac{L\theta}{R_{EPM}}\ln\left(\frac{L}{L-z}\right) + \frac{L\theta}{R_{EPM}}\left(\frac{x^*}{x}\right) \tag{3.6}$$

where x is the length of the unconfined portion of the aquifer, x^* is distance downgradient (in the confined portion of the aquifer) from the edge of the confining layer, and R_{EPM} is the recharge rate in the unconfined portion of the aquifer. Equation 3.6 was used to find the R_{EPM} and x^*/x ratio that gave the best fit to the age data observed from sampling the BC1 and BC2 well nests, and MTT was calculated as $MTT = (L\theta/R_{EPM})(1+x^*/x)$ [*Solomon et al.*, 2006]. To estimate average recharge rate over the entire contributing area, R was calculated as $R = R_{EPM}(1+ x^*/x)^{-1}$.

The exponential-piston flow model and exponential models are both idealized, but the exponential model has been shown to give relatively (± 25%) robust estimates of R (and by

extension, *MTT*) even in moderately heterogeneous environments that violate the underlying assumption of a homogenous surficial aquifer [*Kozuskanich et al.*, 2014]. For the assumed uniform porosities in equations 3.5 and 3.6 we used 0.35, similar to θ observed at 7 sites throughout the North Carolina coastal plain (11 core samples, $\theta = 0.34$, $1\sigma = 0.03$; *Coes et al.* [2007]).

3.6. Results

3.6.1. Groundwater nitrate concentration and denitrification

Widely spaced streambed point sampling (2.5 km reach) and wells gave a similar picture of mean groundwater [NO_3^-] (210 µM and 296 µM, respectively; Table 3.2), while the results from closely-spaced streambed point transects (58 m reach in Fig. 3.1) showed higher groundwater [NO_3^-] (808 µM). All sampling approaches showed [NO_3^-] in groundwater was significantly lower than [NO_3^-]^0, illustrating that denitrification has a major impact on the [NO_3^-] in groundwater discharging from the surficial aquifer (Table 3.2). [NO_3^-]^0 in groundwater samples from the closely-spaced transects in the 58 m reach was, on average, reduced by about half (E_{den}, the percent of [NO_3^-]^0 denitrified in the aquifer, was 51%), less than the 78% and 62% reductions observed for samples from the 2.5 km reach and the wells, respectively. However, the removal of NO_3^- on a molar basis was higher in the 58 m reach, where NO_3^- loss by denitrification, 2[N_2-den], was 656 µM, compared to 394 µM and 292 µM from the 2.5 km reach and wells, respectively.

Table 3.2. Mean values and uncertainty estimates (in parentheses) for groundwater variables from streambed and well nest sampling.

sampling type	$[NO_3^-]$	$[NO_3^-]^0$	E_{den}	*MTT*	*R*	*v*
	μM	*μM*	*%*	*yrs*	*cm/yr*	*m/day*
BC wells[a]	296 (343)	588 (418)	62 (31)	27 (4)[f]	20 (5)[f]	n/a
58 m reach[b]	808 (529)	1453 (502)	51 (30)	29 (6[g])	19 (6)[i]	0.35 (0.37)
2.5 km reach[c]	210 (423)	629 (564)	78 (27)	31 (6[g])	18 (6)[i]	0.40 (0.85)
75 m reach[d]	421 (314)	655 (302)	50	30 (8[h])	18 (6)[i]	0.50 (0.60)
62-263 m reach[e]	437 (469)	n.a.	n.a.	n.a.	n.a.	0.51 (0.53)

[a]sampled June 2013, n = 6 wells, screen length = 45 cm
[b]sampled July 2012, n = 39 for $[NO_3]$, *R* and *v*; n = 36 for $[NO_3^-]^0$ and E_{den}, and n = 35 for *MTT*
[c]sampled March 2013, n = 30 for $[NO_3]$, *R* and *v*; n = 26 for $[NO_3^-]^0$ and E_{den}, and n = 23 for *MTT*
[d]*Kennedy et al.* [2009a], sampled April 2007, n = 21
[e]*Kennedy et al.* [2009b], repeated sampling between December 2005 – December 2006, total n = 422
[f]*MTT* and *R* were calculated for each well nest (based on ages from SF_6 and $^3H/^3He$ at each nest) and the mean value for the exponential-piston flow age distribution model is shown here; uncertainty estimates are based on sensitivity analyses as discussed in the text
[g]value shown is 20% of *MTT*, based on 15-20% uncertainty estimate from *Gilmore et al.* [2015a] supplemental information
[h]value shown is 25% of *MTT*, based on estimated uncertainty of 20-25% in *MTT* [*Kennedy et al.*, 2009a]
[i]calculations for *R* and uncertainty ranges are detailed in Section 5.3

Some of the results in Table 3.2 were used to compare the magnitudes of aquifer and watershed discharge, for both water and NO_3^-, to at least approximately gauge the potential significance of groundwater-derived fluxes to watershed export. A starting point is to compare historic discharge from the watershed with the long-term annual recharge derived from groundwater age-dating results. Stream discharge from the USGS stream gauge (Bear Creek at Mays Store, USGS site 0208925200) suggests annual watershed export of 42 cm/yr (s.d. = 16 cm/yr), based on mean annual discharge of 2.0 m^3s^{-1} (1988 – 2014 data) divided by the upstream watershed area of 152 km^2; watershed area was defined by the USDA NC NRCS 12-Digit Hydrologic Unit boundary upstream of the USGS stream gauge on Bear Creek (Fig. 3.1B). Based on groundwater age dating results, long-term annual recharge in the Bear Creek watershed is roughly 18 – 20 cm/year (Table 3.2) which would account for about

44 – 48% of the 42 cm/yr outflow water budget, similar to the base flow index estimate for the Bear Creek area (0.4, [*Wolock*, 2003]).

Estimating the NO_3^- discharge from the surficial aquifer to the watershed outlet is complicated due to NO_3^- retention in the stream, spatial and temporal variability in groundwater [NO_3^-] and the input of N to the groundwater system, and other factors. Two estimates of aquifer NO_3^- discharge were calculated and compared to the stream export of NO_3^- at the USGS stream gauging station on Bear Creek. The flow-weighted mean [NO_3^-] from streambed sampling in the 2.5 km reach (142 μM, each point groundwater concentration weighted by its associated point groundwater seepage rate) was multiplied by the long-term recharge rate (18 cm/yr, Table 3.2) and the watershed area upstream of the USGS stream gauge on Bear Creek (152 km^2) to calculate an annual groundwater discharge from the surficial aquifer of 5.5 x 10^4 kg N/yr. A second estimate was calculated by scaling up the nitrate flux observed from groundwater sampling in the 2.5 km reach (57 mmol m^{-2} d^{-1} [*Gilmore et al.*, 2015b]) to the entire stream network upstream of the USGS stream gauge (Fig. 3.1B). Assuming groundwater discharges steadily to the main channel (estimated 6.3 m wide x 32 km long) and one third of tributaries (assumed width of 0.5 m x 79 km) throughout the year, the total nitrate load exiting the aquifer would be about 7.0 x 10^4 kg N/yr (5.8 x 10^4 kg N/yr to the main channel alone). The main channel width was estimated from field observations at the upstream ends of the main channel (roughly 2 m), width measured along the West Bear Creek study site (6.5 m), and historic width at the Bear Creek USGS gauge (12 ± 3m, n = 43, Jaime Marlowe, USGS Technician, personal communication). This second estimate of nitrate flux might be considered an upper limit, given that sampling in the 2.5 km reach occurred during a period of high discharge.

For comparison, an estimate of nitrate export from the watershed is available from *Usry* [2006]: 1.0 x 10^5 kg N/yr for the period May 2002 to May 2003. Mean discharge at the USGS gauge for May 2002 to May 2003 was 1.4 m^3s^{-1} compared to 1.6 m^3s^{-1} for 2012-2013, suggesting fairly similar hydrologic conditions between the period of *Usry*'s work and the period of this study. Thus, the magnitude of annual NO_3^- discharge from the aquifer seems to

be roughly 55-70% the magnitude of annual nitrate export from the watershed. Some portion of NO_3^- in aquifer discharge will be retained in the stream channel, but groundwater discharge to the stream seems potentially significant to both water and nitrate watershed export.

3.6.2. Groundwater apparent age and age distributions

The range of groundwater apparent ages (<1 year to ~50 years) was similar for groundwater sampled from wells and from the 58 m and 2.5 km reaches (Figs. 3.3 and 3.4; results from individual point samples are reported in *Gilmore et al.* [2015a]). Age profiles at the wells are reasonably consistent with an exponential increase in age with depth, but apparent ages in groundwater from intermediate and shallow wells were older than predicted by an exponential fit with age = 0 at the water table (Fig. 3.3). An exponential-piston flow model gave a better fit to the data (Fig. 3.3), particularly for the BC1 well nest, where a shallow mud unit (Fig. 3.2) may have restricted recharge near the well site. At the BC2 site, shallow muddy sands may have also contributed to lower recharge in the vicinity of the well nest, but to a lesser extent than at BC1. Given the better fit to the well data, values derived from the exponential-piston flow model are presented as best estimates for *MTT* and *R* in this paper (Table 3.2).

The physical interpretation of the x^*/x ratio derived from the BC1 groundwater age data (0.9) is that recharge occurred over only about half of the distance ($x^* = 0.9x$) between the groundwater divide and the BC1 well site. While only a first approximation of the groundwater divide, the topographical divide (based on the USDA NC NRCS 12-Digit Hydrologic Unit boundary; Fig. 3.1B) is roughly 100 m from the BC1 well nest site, which suggests that a relatively small-scale restrictive horizon (length scale of decameters, e.g., [*Delin et al.*, 2000]) could cause a significant "offset" in the age of groundwater at the shallow wells at the BC1 site. The distance between the BC2 site and the topographic divide is on the order of 800 m and the EPM gives a lower x^*/x ratio (0.4), which suggests a shallow

confining unit on the scale of 230 m in length, a length scale comparable to dimensions of agricultural fields in the watershed.

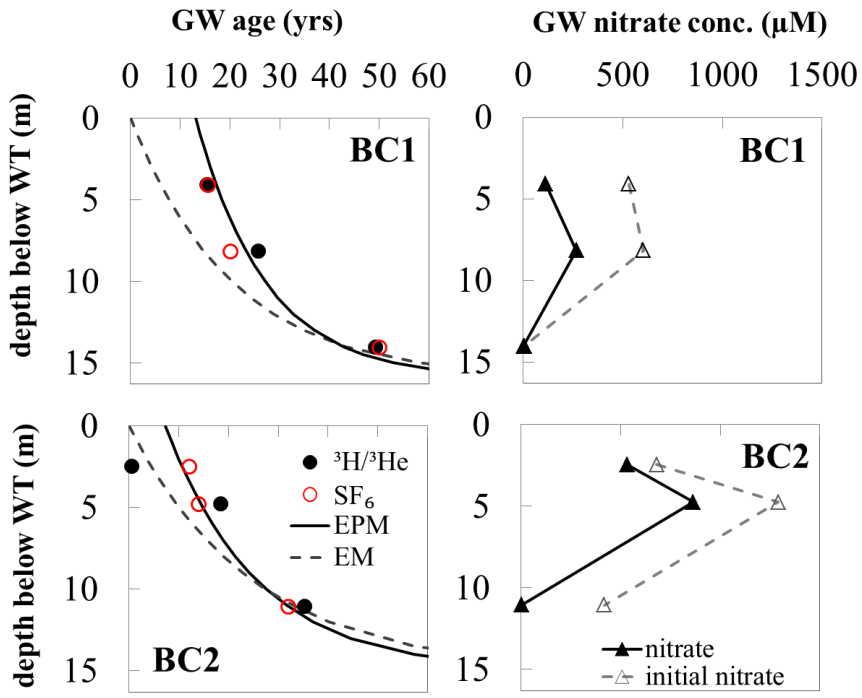

Figure 3.3. Groundwater nitrate concentration and apparent age versus depth below the water table (WT) from well nests BC1 and BC2. The curves fit to the age versus depth data are the exponential (EM) and exponential-piston flow (EPM) models (equations 3.5 and 3.6, respectively).

Groundwater age distributions from streambed sampling (Fig. 3.4) were very consistent between the two field campaigns, despite the different sampling arrangements (closely spaced transects within 58 m versus widely spaced transects over 2.5 km) and different hydrologic conditions (stream discharge of 57 L/s versus 500 L/sec) in July 2012 and March 2013, respectively [*Gilmore et al.*, 2015a]. In both cases, relatively little (≤ 12%) of the groundwater entering the stream was <10 years old, and discharge from the aquifer to the stream was predominantly 20-40 year-old groundwater (76% for both the 58 m and 2.5 km reaches; Fig. 3.4). *Gilmore et al.* [2015a] showed that a gamma distribution [e.g., *Amin and Campana*, 1996; *Kirchner et al.*, 2010] with a large shape factor (α on the order of 10-18) was a good fit to the age distributions from streambed sampling and hypothesized that the absence of very young water may be related to a lack of recharge near the stream. This observation is consistent with the well nest data. It seems possible the surficial aquifer operates as a confined or semi-confined aquifer in the immediate vicinity of West Bear Creek, due to: (1) poorly-drained soils in the stream valley (similar to the situation described by *Böhlke et al.* [2002] for a watershed in Minnesota), and/or (2) the shallow low permeability layer observed in a core near the stream [*Kennedy et al.*, 2009a]. If so, there would be little recharge near the stream and thus even shallow groundwater discharging at the banks would have age significantly >0 years, as was generally observed (Fig. 3.5 and 3.6). Also, this would suggest that groundwater age data from both streambed and well sampling show the effects of spatial variability in groundwater recharge, possibly from confined or semi-confined zones within the unconfined surficial aquifer.

Groundwater discharging at the streambanks of West Bear Creek had an average age of about 20 yr (Figs. 3.5, 3.6), greater than would be expected if there were uniform recharge in the watershed but consistent with an EPM-style distribution of recharge: a confined area without recharge near the stream and unconfined area with recharge further from the stream. We tested whether the idealized EPM could explain the presence of older water along the banks of West Bear Creek and considered whether other factors could have contributed to the 20-year “offset” in groundwater age at the streambanks. Based on the floodplain width (about

150-200 m) and the width of the topographically defined contributing area on each side of the stream (about 600 – 1900 m), the x^*/x ratio with a confined floodplain and unconfined upland (from the outer edge of the floodplain to the topographic divide) would be about 0.1-0.5. Using $x^*/x = 0.5$, $L = 16$ m, $\theta = 0.35$, and an R estimated from streambed sampling (18 cm/year; this value is equivalent to $R_{EPM} = R(1 + x^*/x) = 27$ cm/year), the predicted age "offset" is 10 years (i.e., minimum groundwater age at the streambanks would be about 10 years, on average; this value would closer to 3 years if $x^*/x = 0.1$ were used), lower than the 20-year age offset observed from streambed sampling. If the topographically defined contributing area is a good approximation to the actual groundwater contributing area, this suggests that a confined, no-recharge conditions would need to extend at least about twice the width of the floodplain adjacent to West Bear Creek to produce the average 20-year age "offset" observed from streambed sampling. But other factors may have contributed to the age offset. For instance, the topographical relief between the upland and floodplain may violate the assumption of uniform aquifer thickness for the EPM and may promote (1) discharge of young groundwater to agricultural ditches (effectively lowering discharge of young groundwater at the main channel) and (2) runoff to agricultural ditches (effectively reducing recharge near the stream). Some young groundwater may have discharged through steep unsampled streambed areas just below the water line of West Bear Creek, although given that average stream width (6.5 m) was much greater than average depth along the banks (~15 – 30 cm), the "vertical" portion of the streambed could account for less than 10% of total streambed area and likely a small proportion of groundwater discharge to the stream. For comparison, the EPM described above (10-year minimum groundwater age) predicts that about 40% of aquifer discharge would have 10-20 year age, which is unlikely to be accounted for by bank seepage. Groundwater may have also discharged through seepage faces above the water line (more likely perhaps at higher flow, as in March 2013).

Field-scale variation in recharge (i.e., variation on length scales of 10s to 100s of meters) does seem to be indicated by both streambed and well sampling, two very different sampling approaches. Variability in recharge may have important implications for the transport of

nitrate into the surficial aquifer and surface water. Depending on the location of shallow low permeability layers in the contributing area, shallow groundwater (above the low permeability layer) may rapidly transport NO_3^- to small tributary streams/ditches, or in the absence of a nearby tributary, flow laterally to zones of preferential recharge. Preferential recharge in areas lacking a shallow low permeability layer may induce locally greater flux of NO_3^- to the surficial aquifer [e.g., *Delin et al.*, 2000] similar to irrigated aquifers [*Böhlke*, 2002] where NO_3^- fluxes may outpace the capacity of the groundwater system to remove NO_3^-. In contrast, *Nolan et al.* [2003] showed lower [NO_3^-] in groundwater below high-recharge areas in southern New Jersey, due to greater dilution. In either case, the implication is that field-scale differences in N use, in combination with spatial variation in recharge, could be a potential control on the overall level of NO_3^- contamination in the surficial aquifer (i.e., a "non-point source" contaminant approaching point-like behavior relative to the watershed scale [*Böhlke*, 2002]).

Overall, apparent ages from well sampling and the streambed were similar, mostly ranging from modern to about 50 years (Fig. 3.4). The age distributions from both sampling methods were likely affected by spatial variation in recharge, which suggests that differences in land and N use at the length scale of agricultural fields in the watershed could have a strong influence on groundwater quality in the watershed. The shapes of the age distributions from well sampling were consistent with the exponential-piston flow model, while the distributions from streambed sampling showed a more unimodal pattern (Fig. 3.4) and were better fit by a gamma distribution [*Gilmore et al.*, 2015a].

3.6.3. Groundwater mean transit time and recharge estimates

Groundwater mean transit times (*MTT*) were very consistent between the two streambed sampling campaigns (29 and 31 years, Table 3.2) and slightly greater than the *MTT* derived from sampling groundwater in the well nests (27 years). The apparent ages from SF_6 and $^3H/^3He$ showed strong agreement at 5 of the 6 wells (Fig. 3.3), strengthening confidence in apparent age estimates and the model parameters derived from the data. Overall,

uncertainties in individual apparent ages are likely about 2 – 6 years for $^3H/^3He$ [*Gilmore et al.*, 2015a] and 1 – 4 years for SF_6 (based on differences in $[SF_6]$ from the two different labs where samples were analyzed). The difference in apparent ages derived from SF_6 and $^3H/^3He$ at the shallow BC2 well could be explained by a lag in SF_6 equilibration in the unsaturated zone (i.e., a lag in the when the groundwater "clock" was set [*Solomon et al.*, 2015]), the fact that the $^3H/^3He$ clock is set at the seasonal low water table [*Cook and Solomon*, 1997], higher uncertainty in very young $^3H/^3He$ apparent ages [*Visser et al.*, 2014; *Gilmore et al.*, 2015a], or a combination of those factors.

The choice of age distribution model was important for estimating *MTT* at the BC1 well site, because the exponential model gave a lower *MTT* (20.7 years) than the exponential-piston flow model (27.3 years). In contrast, the *MTT* from the two age distribution models differed by only about 0.8 years for the data from the BC2 site (25.0 versus 25.8 years, respectively), assuming that the surficial aquifer had the same thickness at BC1 and BC2 (the top of the underlying confining unit was observed in drilling at BC1 but not at BC2). Varying the estimated aquifer thickness at BC2 by ±2 m resulted in roughly ±5 years in *MTT* from the BC2 data. For the samples collected at the BC1 site, we also assessed the possible effect of lagged SF_6 transport in the unsaturated zone on the *MTT* estimate by assuming that the SF_6 groundwater age "clock" was set at the land surface rather than the water table (in this case apparent age > 0 years at the water table would be expected). Using the land surface clock assumption, *MTT* derived from the exponential model was about 20 years, 7 years less than the 27 year mean from both well nests (Table 3.2). If either the BC1 or BC2 *MTT* estimate was biased by a maximum of 7 years, *MTT* would change by 3.5 years. This is roughly equivalent to two extreme test cases in which all $^3H/^3He$ and SF_6 apparent ages were assumed to be biased either high or low by 4 years in which case the uncertainty in *MTT* was ± 4 years. Uncertainty in individual apparent age estimates are likely on the order of 1 – 6 years, so a 4-year bias in age seemed a reasonable test of uncertainty. Based on these sensitivity calculations, a reasonable upper limit for uncertainty in *MTT* from well data is probably about 4 years (Table 3.2).

Recharge rate (R) is calculated in the process of determining *MTT* from well data. Recharge estimated from fitting the EPM to well data (R_{EPM}) was 34 cm/yr. When R_{EPM} was averaged over the total aquifer length (confined plus unconfined, see discussion of equation 3.6), R was 20 ± 5 cm/yr (Table 3.2). Uncertainty was estimated using the same sensitivity analysis as described for *MTT*; maximum potential bias in R from either well was estimated at 10 cm/yr, which would translate to a 5 cm/yr change in the mean value. Ultimately, the recharge estimate from well data is similar to recharge estimates from groundwater age dating in the North Atlantic coastal plain (6 – 40 cm/yr, *McMahon et al.* [2011] and references therein) and the coastal plain in North Carolina (4 – 30 cm/yr; [*Coes et al.*, 2007]).

The simplest model ($R = L\theta/MTT$) was used to estimate recharge from the streambed *MTT* values (Table 3.2). Based on $\theta = 0.35$ and the aquifer thickness observed at the BC1 well nest (16.0 m) R was estimated as 19 ± 6 cm/yr and 18 ± 6 cm/yr for the 58 m and 2.5 km reaches, respectively (Table 3.2). The upper and lower bounds of uncertainty (+5 to -7 cm/yr) were estimated by varying aquifer thickness and porosity values. A minimum recharge estimate (12 cm/yr and 11 cm/yr for the 58 m and 2.5 km reaches, respectively) was calculated using a low porosity (0.30) and the mean of the saturated aquifer thicknesses determined at the BC1 site and next to West Bear Creek (11.5 m). Maximum recharge estimates (24 and 23 cm/yr, respectively) were calculated using the total thickness (17.7 m; land surface to confining unit at the BC1 site) and higher porosity (0.40). Using instead the exponential-piston flow model parameters that gave a best fit to the observed streambed age distributions 58 m and 2.5 km data sets ($x^*/x = 1.8$, $MTT = 31$ years, $R_{EPM} = 51$ cm/yr [*Gilmore et al.*, 2015a]), the spatially averaged R was 18 cm/yr, essentially the same as the estimate from the exponential model.

Mean transit times from well sampling and streambed sampling were similar (27 – 31 years). Recharge rates from the two sampling approaches (20 cm/day and 18-19 cm/day, respectively) were also similar.

3.6.4. Historic and spatial trends in groundwater nitrate concentrations

Streambed maps from the closely-spaced transects (58 m reach) show that groundwater age was generally greatest and $[NO_3^-]$ generally lowest beneath the center of the stream (Fig. 3.5), consistent with results from previous sampling in West Bear Creek [*Kennedy et al.*, 2009a]. Higher nitrate was observed, on average, in 15 – 20 year-old groundwater recharged in the 1990s (Fig. 3.6). Groundwater $[NO_3^-]$ was lower on the right (south) bank in both the 58 m reach and 2.5 km reach, in association with a higher proportion of forest land cover south of the stream (Fig. 3.1); lower $[NO_3^-]$ seems due to a combination of lower N input and greater denitrification along groundwater flowpaths contributing to the right side of the 58 m reach (Fig. 3.6). These cross-channel patterns are an advantage of the streambed sampling approach, compared to sampling in well nests where information from only one up-gradient direction is obtained.

The highest $[NO_3^-]$ observed in samples from the well nests corresponded to roughly 18-26 year-old water (602 μM and 1292 μM for BC1 and BC2 nests, respectively). The $[NO_3^-]^0$ in the youngest groundwater (<10 years, based on apparent ages from $^3H/^3He$) was lower than in the 10-20 year-old water, which was consistent with observations from streambed sampling (Fig. 3.4). Sample sizes were small ($n \leq 3$) and differences in mean $[NO3]^0$ between the 0-10 year-old groundwater and 10-20 year-old groundwater were not significant at the 95% level ($p>0.05$ for two-tailed t-test assuming unequal variances (no test could be conducted for the well samples because of the low number of samples in each bin, n = 1-2), but all three sampling campaigns (Fig. 3.4) were consistent in showing lower mean $[NO_3^-]^0$ for 0-10 year-old groundwater than for 10-20 year-old groundwater. With the exception of the youngest groundwater (0 – 10 years old), both streambed and well sampling suggest decreasing $[NO_3^-]^0$ with increasing groundwater age (Fig. 3.4), which is consistent with the pattern of county-level N fertilizer sales estimates (Wayne County, NC) in the Bear Creek

watershed area [*Ruddy et al.*, 2006] and with previous work in West Bear Creek [*Kennedy et al.*, 2009a].

Slightly lower $[NO_3^-]^0$ in the youngest groundwater (Fig. 3.4) may indicate progress toward reduction of agricultural N use in the watershed in response to the nutrient management efforts in the Neuse River basin. This would be consistent with reported reductions in N use by the agricultural community [NCDENR, 2012], even though not all growers in the watershed have adopted specific "nutrient management plans" [*Osmond et al.*, 2015]. In general, lower $[NO_3^-]^0$ for young groundwater could result from several factors, including a shift in dominant crop type in the watershed [e.g., NCDENR, 2012], changes in crop rotation patterns [e.g., *Jaynes et al.*, 2001], or if the groundwater was recharged in wooded areas near the stream or well nests. Detailed crop data was not available for the areas of the watershed where groundwater was collected, so the potential impact of changes in dominant crop or cropping systems could not be assessed. Both the BC1 and BC2 well nests were installed in the edges of wooded areas but the upgradient areas were dominated by agricultural fields, and the age distributions suggest that vertical recharge in the immediate vicinity of the wells was very low, so the lower $[NO_3^-]$ from the shallowest wells may be related to lower N use in up-gradient fields. Similarly, for streambed sampling in both the 58 m and 2.5 km reaches, the data suggest that lower $[NO_3^-]$ in the youngest groundwater could also be related to lower recent N use in agricultural fields. For instance, four of the six sampling locations with young low-$[NO_3^-]$ groundwater (two out of three for both the 58 m and 2.5 km reaches) were along the left bank where $[NO_3^-]$ in groundwater older than 10 years was generally higher (Fig. 3.6). In other words, $[NO_3^-]$ in young groundwater was unlikely to be low as a result of being recharged only in wooded areas; $[NO_3^-]$ in groundwater recharged beneath some agricultural fields seems to be lower in recent years.

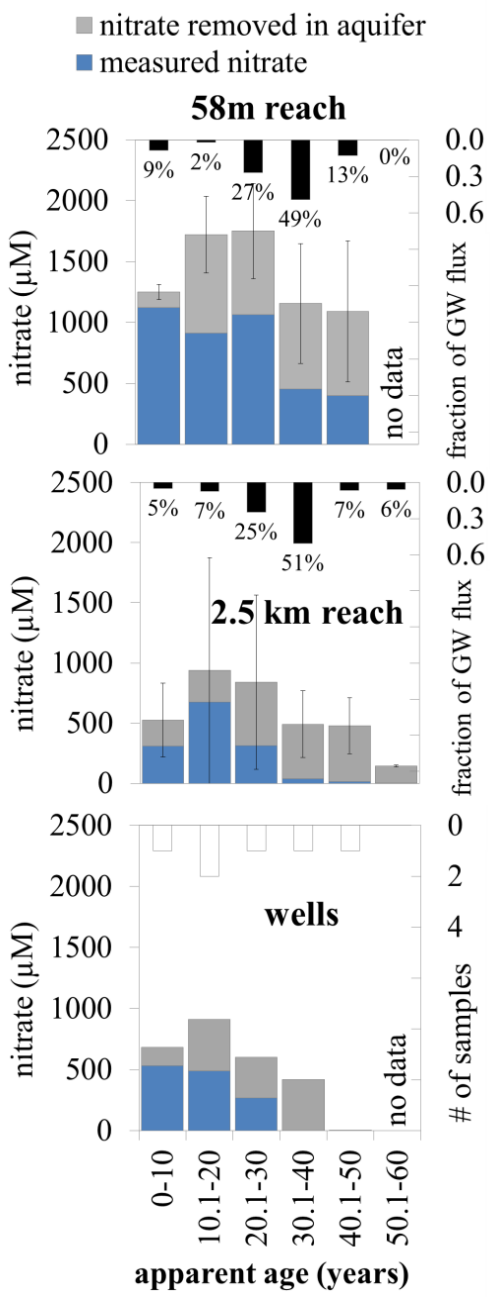

Figure 3.4. Patterns of nitrate concentration from the last 5-6 decades, from well and streambed groundwater sampling. The age distributions of groundwater discharging through the streambed (black bars in top two plots) were calculated after [*Kennedy et al.*, 2009a]. Total height of a stacked bar (blue + gray) indicates $[NO_3]^0$. Whiskers show standard deviations representing spatial variability in $[NO_3]^0$. No whiskers are shown for the well results, because each bar represented only 1-2 samples.

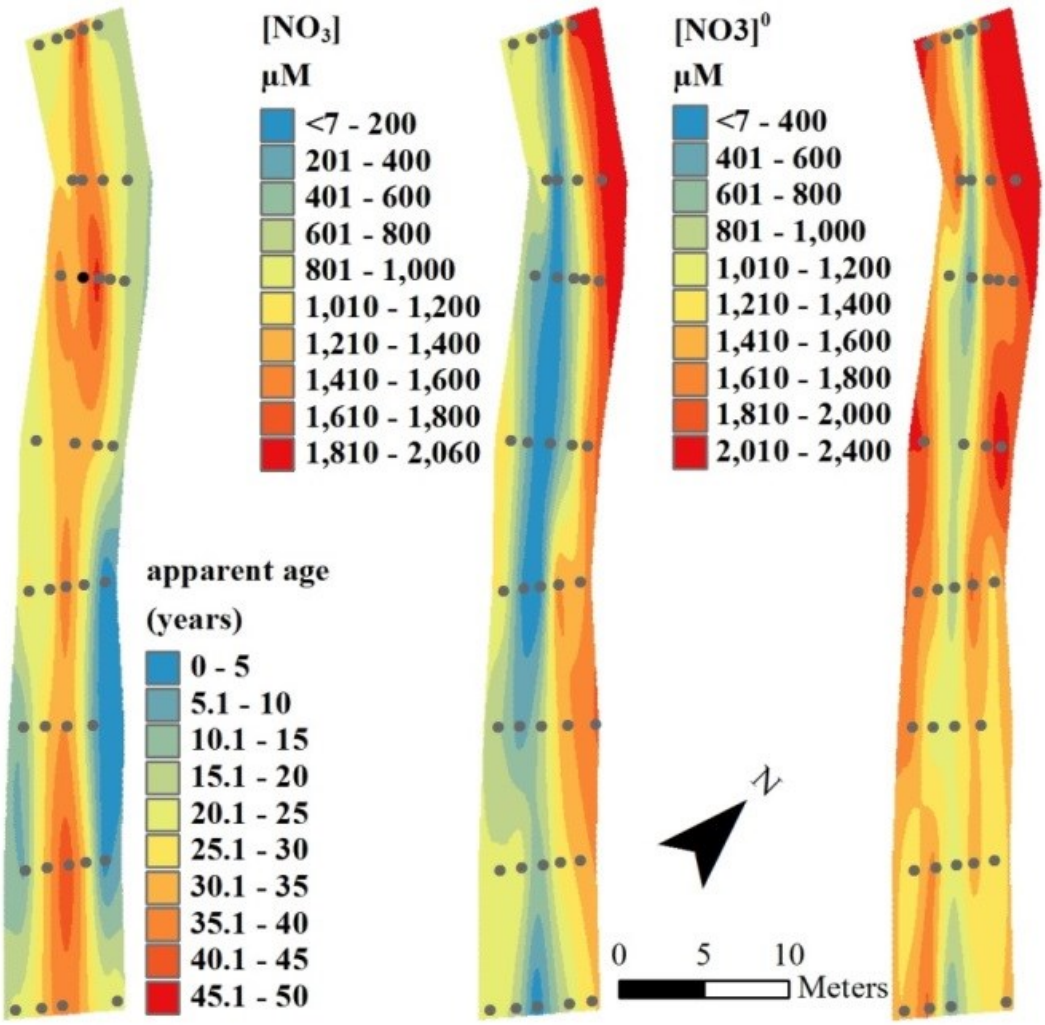

Figure 3.5. Streambed maps of groundwater nitrate concentration and age for the 58 m reach, July 2012. One apparent age from SF_6 (black dot) was substituted where $^3H/^3He$ data was not available. Maps were created in ESRI™ ArcMap using the multiquadric radial basis function with anisotropy ratio of 8 and smoothing parameter set = 0.

Results from the well nests and the 2.5 km stream reach (Table 3.2, Fig. 3.4) suggest that nitrate contamination in the surficial aquifer is typically lower than observed in the 58 m reach. The 58 m reach, and a roughly 300-400 m reach that contains it [*Kennedy et al.*, 2009a,b; *Gilmore et al.*, 2015b], represents a "hotspot" of significant nitrate flux into West Bear Creek.

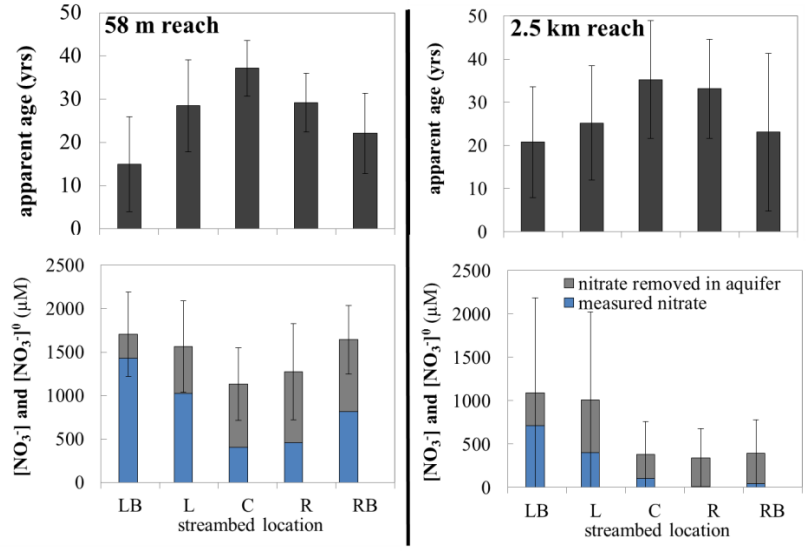

Figure 3.6. Cross-channel variation in groundwater apparent age, nitrate removed by denitrification in the surficial aquifer, and observed nitrate concentrations in groundwater sampled in the 58 m and 2.5 km reaches. Whiskers show standard deviations representing spatial variability. Total height of a stacked bar (blue+gray) indicates $[NO_3]^0$. Each bar represents the mean of 6-8 samples in the 58 m reach and 4-6 samples in the 2.5 km reach.

3.6.5. Groundwater NO_3^- budget

Knowing the rate of change of NO_3^- storage in the surficial aquifer (S_{NO3}, mass of NO_3^- per unit area of aquifer) could be useful for deciding whether more aggressive nutrient management initiatives are necessary (if S_{NO3} is increasing), or if a more passive approach may be suitable (if S_{NO3} is decreasing). Although variability in NO_3^- was significant (e.g., Figs. 3.4, 3.6), similar historic patterns of NO_3^- contamination observed from groundwater sampling in the 2.5 km reach and the well nests (Fig. 3.4) offered an opportunity to compare strengths and weaknesses of the two sampling approaches in calculating a groundwater NO_3^- budget and also to explore potential uncertainties. We used equation 3.7 below to estimate whether NO_3^- storage in the surficial aquifer (S_{NO3}) was increasing, decreasing, or staying roughly the same at the time of sampling (Table 3.3):

$$\frac{dS_{NO3}}{dt} = f_{NR} - f_{NO3} - 2f_{N2-den} \tag{3.7}$$

where dS_{NO3}/dt is the rate of change of NO_3^- storage in the surficial aquifer (positive if storage is increasing, negative if storage is decreasing), f_{NR} is the current NO_3^- flux into the surficial aquifer by groundwater recharge, f_{NO3} is the current NO_3^- flux out of the surficial aquifer to West Bear Creek, and $2f_{N2\text{-}den}$ represents the current denitrification sink in the aquifer ("current" refers to the time of sampling, 2013).

For a surficial aquifer at steady state with regard to groundwater storage, and without significant net groundwater exchange with an underlying regional aquifer or groundwater loss by ET, the groundwater flux into the aquifer (R) is equal to the groundwater flux out of the aquifer to surface water. Assuming this, and using the streambed sampling data from the 2.5 km reach, dS_{NO3}/dt was calculated for the surficial aquifer surrounding West Bear Creek by using the following values in equation 3.7:

- f_{NR}: calculated as $R[NO_3^-]^0_{young}$, where R is from streambed sampling data (18 cm/yr) and $[NO_3^-]^0_{young}$ is the flow-weighted mean of the three streambed samples from the 2.5 km reach with age < 10 years (637 μM, Table 3.3)

- f_{NO3}: calculated as $R[NO_3^-]_{FWM}$, where $[NO_3^-]_{FWM}$ is the flow-weighted NO_3^- concentration in groundwater discharging to West Bear Creek (n = 30, 142 μM, Table 3.3)

- $f_{N2\text{-}den}$: calculated as $R[N_2\text{-den}]_{FWM}$, where $[N_2\text{-den}]_{FWM}$ is the flow-weighted concentration of N_2-den in groundwater discharging to West Bear Creek, modeled from dissolved gases in groundwater (n = 26, 375 μM, Table 3.3)

All terms are mass of NO_3^- per unit contributing area (recharge area, not streambed area) per year.

In the case of streambed sampling, each NO_3^- or N_2-den concentration measured at each streambed point was weighted by the rate of aquifer discharge (groundwater flux measured at each sampling point in the streambed) to calculate the flow-weighted mean concentrations used in equation 3.7.

In using well data to solve for dS_{NO3}/dt, the overall approach was the same except that the rate of aquifer discharge associated with each NO_3^- or N_2-den concentration was not known, so a weighting function was required to calculated flow-weighted concentrations in aquifer discharge. Lumped-parameter models (LPM) such as the EM or EPM are commonly used as weighting functions, because the models define a fraction of aquifer discharge (q_i) for different ranges of groundwater age associated with different depth intervals in the aquifer. For instance, for an aquifer with thickness of 16 m, $\theta = 0.35$, and recharge = 20 cm/year, the EM predicts q_i = 30% for the 0-10 year age interval, 21% for the 10-20 year age interval is 21%, and so on (see Appendix E for q_i values from the EM, EPM, and gamma distributions used in this analysis). With q_i known for each age bin in Figure 3.4, a flow-weighted concentration in groundwater was calculated ($C_{LPM} = \Sigma C_i q_i$, where C is either $[NO_3^-]$ or $[N_2$-

den] in age bin *i*). The flow-weighted groundwater concentrations and *R* used to calculate dS_{NO3}/dt are shown in Table 3.3 and details of each LPM used to calculate the concentrations are in Appendix E.

Table 3.3. NO_3^- and N_2-den concentrations and recharge rates used for aquifer NO_3^- mass balance.

	$[NO_3^-]_{FWM}$	$2[N_2\text{-den}]_{FWM}$	$[NO_3^-]^0_{young}$	*R*
	μM	*μM*	*μM*	*cm/yr*
2.5 km reach	142	375	637[b]	18
wells EM[a]	302	226	682	20
wells EPM[a]	289	332	682	20
wells gamma[a]	123	332	682	20

[a]FWM concentrations from well data are $[NO_3^-]_{LPM}$ and $2[N_2\text{-den}]_{LPM}$ values, see Appendix E for details of LPMs

[b]$[NO_3^-]^0_{young}$ is the flow-weighted concentration for the 3 groundwater samples with apparent age < 10 years, see Table E5 in Appendix E for concentrations and *v*

For the groundwater system connected to the 2.5 km reach, dS_{NO3}/dt was positive (21 mmol m^{-2} d^{-1}, Table 3.3), although it's likely the value was not significantly different than zero based on its relatively large uncertainty (170%). Estimates from well nest data also suggested that storage was increasing (dS_{NO3}/dt = 12-31 mmol m^{-2} yr^{-1} for EPM and EM, respectively) and in the case of the gamma model the rate of change may be significantly different than zero (47 ± 43 mmol m^{-2} yr^{-1} for the gamma distribution, the distribution best fitting the observed age data [*Gilmore et al.*, 2015a]). Overall, the dS_{NO3}/dt estimates suggest that nitrate storage in the aquifer is increasing at an annual rate of about one-fifth of the year's input (average of dS_{NO3}/dt divided by f_{NR}, Table 3.4).

The f_{NO3} and $2f_{N2\text{-}den}$ values were very similar for the gamma model and the 2.5 km reach, because the historic patterns in NO_3^- were similar (Fig. 3.4) and the gamma model used to calculate the flow-weighted concentrations from well data was a good fit to the groundwater age distribution observed in the 2.5 km reach [*Gilmore et al.*, 2015a]. The total rate of nitrate loss from the aquifer ($f_{OUT} = f_{NO3} + 2f_{N2\text{-}den}$) for the gamma model (90 mmol m^{-2} yr^{-1}) was low relative to the value from the EM or EPM because the gamma distribution had a q_i of 0% for groundwater < 10 years old and only 4% for groundwater in the 10-20 year age range with high nitrate ($[NO_3^-]^0$ = 911 µM, Fig. 3.4). The q_i values for the EM and EPM for those same age intervals were 30% and 21%, and 0% and 46%, respectively, giving higher f_{OUT} values for those models. Additionally, the relative contribution of the denitrification sink ($2f_{N2\text{-}den}/f_{OUT}$) from the different LPMs gives a slightly different picture of the importance of denitrification in the NO_3^- budget. While the EM and EPM results suggest denitrification accounts for 42% and 53% of f_{OUT}, respectively, the gamma model suggests 72% of f_{OUT}, which is similar to the $2f_{N2\text{-}den}/f_{OUT}$ from the 2.5 km reach data (73%), and slightly higher than mean E_{den} from well data (62%, Table 3.2).

Table 3.4. Aquifer NO_3^- mass balance, with estimated uncertainties in parentheses.

approach	f_{NR}	f_{NO3}	$2f_{N2\text{-}den}$	dS_{NO3}/dt
	mmol m^{-2} $year^{-1}$, where m^{-2} is contrib. area			
2.5 km reach	115 (41%)	26 (45%)	68 (42%)	21 (170%)
wells EM[a]	136 (36%)	61 (48%)	45 (33%)	31 (143%)
wells EPM[a]	136 (36%)	58 (48%)	66 (33%)	12 (360%)
wells gamma[a]	136 (35%)	25 (48%)	65 (32%)	47 (86%)

[a]see Appendix E for details of each lumped-parameter model used to calculate $[NO_3^-]_{LPM}$ and $[N_2\text{-den}]_{LPM}$

A Monte Carlo analysis was used to gauge uncertainty for dS_{NO3}/dt. The concentrations and recharge rates in Table 3.3 were randomly varied and the coefficient of variation (the standard deviation divided by the mean for the 5000 iterations) was taken as an estimate of percent uncertainty (Table 3.4, details in Appendix E). There was a high degree of spatial variability in $[NO_3^-]$ in the 2.5 km reach (Fig. 3.6), but the similarity in overall mean values and in historic NO_3^- trends between the 2.5 km reach and the well nests (Fig. 3.4) increased confidence that the $[NO_3^-]$ from each sampling approach may reasonably represent groundwater nitrate contamination in the aquifer. However, the estimated uncertainties in the concentrations and recharge rates (mostly 20-30%, Appendix E) propagated into substantial uncertainty in f_{NR} and f_{OUT} (35-42%) and dS_{NO3}/dt (≥86%). Among the four cases in Table 3.4, those in the first and last rows ("2.5 km reach" and "wells gamma") had the best agreement in f_{OUT} and the relative contribution of the denitrification sink (72-73%) to f_{OUT}.

3.6.6. Future NO_3^- flux from the aquifer

Some studies have explored how hypothetical changes in recharge NO_3^- fluxes may affect $[NO_3^-]$ in groundwater samples collected from wells (*Böhlke* [2002]; their Fig. 14), the fraction of baseflow contaminated by NO_3^- (*Modica et al.* [1998]; their Fig. 15), the magnitude of baseflow $[NO_3^-]$ (*Browne and Guldan* [2005]; their Fig. 8), and the depth of NO_3^- contamination in the surficial aquifer (*Welch et al.* [2011]; their Fig. 8). Results from this study offer an opportunity to calculate future groundwater nitrate fluxes from the surficial aquifer to the stream (f_{NO3}, normalized by streambed area), based on estimates of $[NO_3^-]^0$ and the known age distribution of groundwater in aquifer discharge (Fig. 3.4). As with the groundwater NO_3^- balance, the similar patterns of historic $[NO_3^-]$ from the well nest and streambed sampling approaches (Fig. 3.4) and consistent age distributions from streambed sampling (from this study, *Kennedy et al.* [2009a], and *Browne and Guldan* [2005]) give some confidence that the predicted long-term patterns of f_{NO3} may be realistic.

The flux of NO_3^- from the surficial aquifer in future decades was predicted (Fig. 3.7) assuming steady hydrologic conditions (constant groundwater seepage rate, v, and age

distribution, Table 3.5) as well as steady denitrification rate (ε) in each 10-year age bin (Fig. 3.4). $[NO_3^-]_{FWM}$ was calculated and multiplied by v to determine flux at each 10-year time step in Fig. 7. The v measured in the 58 m reach was believed to be a plausible representation of long-term discharge (Appendix E) and was used to calculate fluxes in the 58 m plot (Fig. 3.7). For the 2.5 km reach and well nest plots, a long-term annual recharge rate (20 cm/yr, from well nest data, Table 3.2) was converted to a v through the streambed of 0.20 m/day (Appendix E, $v = RA_R/A_S$, where A is the topographically-defined groundwater contributing area for the 2.5 km reach (subscript R), or the streambed area (subscript S)) and multiplied by $[NO_3^-]_{FWM}$. The v from the 2.5 km reach (0.40 m/day) was not used because it was measured on the falling limb of a hydrograph during wet spring-time conditions, and may have been higher than the long-term average v.

Two cases were explored for $[NO_3^-]^0$. In the first case, groundwater $[NO_3^-]^0_{young}$ from the 58 m and 2.5 km reaches was assumed to represent current conditions across the contributing area for the respective reach, and was held constant over time. In the second case, $[NO_3^-]^0 = 0$ was used to simulate a sudden elimination of agricultural N use in the watershed. The second case may seem unrealistic for a predominantly agricultural watershed, but could give insight into system response to remediation projects (e.g., conversion of agricultural land to forest or other land use).

Two different denitrification assumptions were also explored. In the first case, constant removal of NO_3^- on a molar basis was assumed for each 10-year age bin (i.e., the quantity $[NO_3^-]^0 - [NO_3^-]$, represented by each gray bar in Fig. 3.4, was held constant). In the second case, percent removal (i.e., E_{den}, equal to the ratio of the gray bar height to the total height of the blue plus gray stacked bar in Fig. 3.4) was assumed constant for each age bin. Given that average τ for each age bin is constant and that ε is assumed to be constant, the two cases are functionally equivalent to assuming zero-order and first-order denitrification rates, respectively. The zero-order expression may be written as $\varepsilon\tau = [NO_3^-]^0 - [NO_3^-]$ [*Böhlke* 2002], while the first-order expression may be written $e^{-\varepsilon\tau} = [NO_3^-]/[NO_3^-]^0$ (where percent

removal, E_{den} = 1 - $[NO_3^-]/[NO_3^-]^0$). Given constant τ and ε, these equations translate to constant NO_3^- removal in each age bin on molar and percentage bases, respectively.

Table 3.5. Inputs used to predict future NO_3^- flux from the surficial aquifer.

variable	58 m reach	2.5 km reach	well nests
$[NO_3^-]^0$ (μM)	1251 or 0	525 or 0	682 or 0
v (m/day)	0.35	0.20	0.20
age distribution	empirical (Fig. 3.4)	empirical (Fig. 3.4)	EPM[a] or gamma[a] (e.g., Appendix E)
denitrification rate, ε	zero or first-order	zero or first-order	zero-order[b]
age range where complete NO_3^- removal was assumed[c]	>50 years	>50 years	>40 years

[a]EPM and gamma distributions were fit to well nest and streambed data, respectively. See Appendix E for details.
[b]Zero- and first-order curves were very similar for the EPM. For the gamma distribution, the first-order curve paralleled the EPM curves, but was shifted downward by roughly 65 mmol m^{-2} d^{-1}.
[b]Older age bins with no (or very low) $[NO_3^-]^0$ were assumed to have complete removal of NO_3^-

First-order denitrification rates have been calculated for some aquifer systems [*Böhlke* 2002], but zero-order rates are commonly considered appropriate for aquifer systems [e.g., *Green et al.*, 2008 and references therein; *Liao et al.*, 2012]. The assumption of zero-order or first-order denitrification is important for predicting f_{NO3} (Fig. 3.7), particularly for the 2.5 km reach. Results from sampling the 58 m reach suggest $[NO_3^-]^0$, on average, exceeds the denitrification capacity of the aquifer in all age ranges. This may give some insight into the capacity of the aquifer system to remove NO_3^-, and therefore which rate assumption may be

more appropriate for the groundwater system. For example, even for the groundwater with the longest residence time (30 to 50 years, Fig. 3.4) denitrification was not complete (2[N_2-den] was about 700 μM, on average). This could suggest that at least some flowpaths in the aquifer have a finite capacity for NO_3^- removal on a molar basis, and that the zero-order curves may be realistic for this groundwater system.

For age bins with little or no $[NO_3^-]^0$ (e.g., > 40 years for well data), complete removal of NO_3^- was assumed (Table 3.5). This assumption had relatively little impact on f_{NO3} predictions from streambed sampling, because the fraction of groundwater discharge was small (0-6%) for age ranges with little or no observed $[NO_3]^0$. For the well data, the assumption of complete denitrification for older groundwater did tend to lower f_{NO3} predictions, particularly for the zero-order denitrification rate case, because the EPM and gamma age distributions had groundwater discharge fractions of 11% and 15%, respectively, for groundwater > 40 years old. In other words, as the NO_3^- "peak" in Fig. 3.4 was shifted to ages > 40 years during the course of predicting future f_{NO3}, it was removed from the system via denitrification and no longer contributed to f_{NO3}. Assuming constant removal of NO_3^- on a molar basis (e.g., using a rough maximum value of 700 μM for ages > 40 years) had essentially the same effect. The f_{NO3} predicted from well data may be minimum estimates for the zero-order ε case.

After 4-5 decades, future predicted f_{NO3} approaches zero when $[NO_3^-]^0 = 0$, but remains elevated near current f_{NO3} if $[NO_3^-]^0$ remains at current values (Fig. 3.7). The time scale required to approach steady nitrate discharge from the aquifer is 1-2x the *MTT* derived from the streambed and well nest sampling approaches (about 30 years). Less intuitive are the exact pathways, during the first 1-3 decades, to these eventual steady conditions. Predictions from the 2.5 km reach data show two different paths associated with the different denitrification rate assumptions. Under the assumption of zero-order denitrification, f_{NO3} increases rapidly for a decade before returning to current levels after about 30 years, regardless of NO_3^- concentration in groundwater recharge (zero, or current $[NO_3^-]^0$). The peak in the f_{NO3} curve is a result of the high $[NO_3^-]^0$ from the late 1990s to early 2000s (10-20

year age bin, Fig. 3.4) being shifted over time toward age intervals associated with a large fraction of aquifer discharge (25% for 20-30 years; 51% for 30-40 years). There is no peak in f_{NO3} for the first-order ε curve, because the 30-40 year age bin (51% of discharge) has a denitrification proportion of almost 100%. As a result, groundwater with high nitrate passes through the system without increasing f_{NO3}. For the 58 m reach, all f_{NO3} curves have a peak, except the curve associated with no $[NO_3^-]^0$ and first-order ε, because $[NO_3^-]$ of groundwater currently in the aquifer overwhelms the denitrification capacity of the aquifer (there is currently significant $[NO_3^-]$ in groundwater of all ages, top panel of Fig. 3.4).

Future f_{NO3} predicted from streambed sampling raises the possibility that fluxes of nitrate from the surficial aquifer could increase in the next 10-20 years, despite apparently lower N use in recent years, and even if N use in the watershed were eliminated (zero-order curves, Fig. 3.7). Applying a gamma age distribution to predict f_{NO3} from well data also suggested that f_{NO3} could increase in the next 20 years, regardless of $[NO_3^-]^0$. These predictions could only be made by having the actual age distribution of aquifer discharge to the stream, or by weighting groundwater N data from wells by an age distribution model (gamma distribution with $\alpha = 18$) that was a good fit to age data collected in the streambed. In contrast, predictions from applying an EPM age distribution to N concentrations observed at well nests suggested that f_{NO3} would remain unchanged for a decade before responding to the elimination of N use in the watershed. The response to the no N-use scenario after 10 years is rapid, typical of an age distribution with an exponential component (i.e., the predicted f_{NO3} response is similar to that of an EM, but lagged by a decade). The tailing (long-term persistence of low levels of contaminant) usually associated with an EPM prediction is dampened in the wells graph in Fig. 3.7 because of the assumption of complete denitrification for groundwater with age > 40 years. Interestingly, the initial f_{NO3} (in 2013, Fig. 3.7) predicted by the EPM was roughly double the f_{NO3} calculated for the 2.5 km reach (Fig. 3.7), even though the f_{NO3} was calculated using the same v (Table 3.5). This suggests that the predicted f_{NO3} from the EPM may not be a good representation of actual f_{NO3} to the stream. The primary reason for this difference is that the EPM is weighted heavily with high

$[NO_3^-]^0$ (911 μM) observed at the wells in the 10-20 year interval which represents 46% of groundwater discharge. Only 7% of discharge to the 2.5 km reach was actually in the 10-20 year interval. The initial f_{NO3} predicted using the gamma distribution was more consistent with the initial f_{NO3} calculated for the 2.5 km reach, because in the gamma model only 4% of groundwater discharge was associated with the high-nitrate groundwater.

Thus, a relevant question for prediction of future f_{NO3} is whether the age distribution in groundwater discharge to the main channel of West Bear Creek is representative of the age distribution in discharge to the full stream network in the watershed. Our analysis of the work of *Browne and Guldan* [2005] in a meandering sand plains stream suggested that the age distribution observed in the main channel of West Bear Creek may hold for other unconfined siliciclastic aquifers [*Gilmore et al.*, 2015a]. However, age distributions with large fractions of young groundwater have been shown to be reasonably consistent with age dating tracers measured in groundwater samples from well nests (e.g., exponential model [e.g., *Böhlke*, 2002]), or from hydrologic tracers such as Cl^- in the stream water of some small watersheds (gamma distribution with $\alpha < 1$ [e.g., *Kirchner et al.*, 2010]). This suggests that further investigation is warranted, perhaps by sampling groundwater discharge to the full stream network in the watershed, along both the lateral and longitudinal axes of the channels, to determine whether the age distribution observed in West Bear Creek is representative of aquifer discharge throughout the watershed. Finding that the age distribution in groundwater discharge to West Bear Creek is representative of groundwater discharge watershed wide would support the predicted increases in f_{NO3} based on streambed sampling results (zero-order ε assumption) and from the gamma distribution curve (Fig. 3.7). In other words, it would seem likely that nitrate discharge from the surficial aquifer would increase, and thus surface water quality may decrease, for 1-3 decades into the future, regardless of future nitrate loading to the aquifer by groundwater recharge.

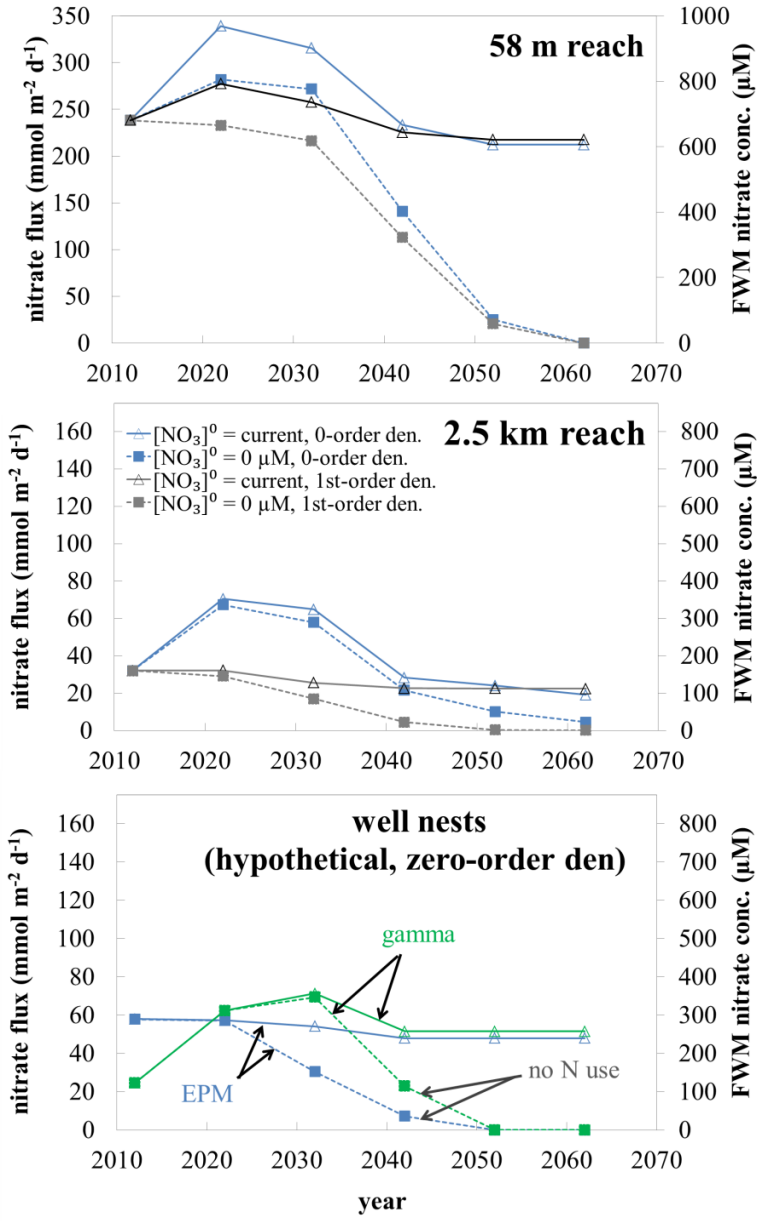

Figure 3.7. Predicted nitrate flux and $[NO_3^-]_{FWM}$ from the surficial aquifer to West Bear Creek. For the stream reaches, predicted fluxes are based on observed age distributions (Fig. 3.4). Using the well data, predicted fluxes were based on age distributions; either an exponential-piston flow model (EPM) with minimum groundwater age of 10 years, or a gamma age distribution ($\alpha = 18$) (Appendix E). For simplicity, only curves based on zero-order denitrification are shown on the well nest plot. See Table 3.4 and text for additional detail. Note that the vertical scale of the 58 m reach plot differs from the scale of the 2.5 km and well nest plots.

3.6.7. Summary and Conclusions

Groundwater mean transit time (*MTT*) and the magnitudes of NO_3^- contamination and denitrification were determined for an agricultural coastal plain watershed based on groundwater sampling in a streambed and in two 3-well nests. Streambed sampling was conducted in a 58 m reach using 8 closely-spaced point transects and in a 2.5 km reach using 6 widely-spaced point transects (Fig. 3.1). Results from the streambed and well nest sampling were also used to calculate a NO_3^- budget for the groundwater system, and to predict how NO_3^- discharge from the aquifer to streams may change in coming decades.

Groundwater [NO_3^-] was variable among the sampling approaches, with higher mean from the 58 m reach (808 μM) than from the 2.5 km reach (210 μM) or well nests (296 μM). The 2.5 km reach and the well nests may capture a more representative picture of average levels of NO_3^- contamination in the surficial aquifer (Table 3.2), while the 58 m reach seems to represent a "hotspot" for NO_3^- flux from the aquifer. Both streambed and well nest sampling indicated that denitrification in the surficial aquifer significantly reduced groundwater [NO_3^-] (51 – 78% on average, Table 3.2). The streambed sampling approach has the advantage of showing streambed patterns of groundwater [NO_3^-] and extent of denitrification (Figs. 3.5 and 3.6) with the streambed patterns relating to general recharge locations in the watershed (left vs right side of the stream, upstream vs downstream along the study reach, and near to vs farther from the stream for streambank vs channel-center samples), all without wells. Such observations could be useful for guiding remediation or nutrient management; for example, in the West Bear Creek watershed, targeting the left side of the stream where denitrification is lower could have the most beneficial impact on water quality. Overall, where streambed sampling is feasible (e.g., sandy streambed) the approach is ideal for determining fluxes of groundwater NO_3^- to a stream reach. Aquifer thickness and porosity are required to determine groundwater recharge rates from either sampling approach, and to derive *MTT* from well sampling (but not from streambed sampling).

A major finding is that SF_6 and $^3H/^3He$ age results for groundwater samples from both wells and streambed points were consistent with a lack of groundwater recharge (i.e., confined or semi-confined conditions) in some areas of the watershed, both near the stream and further upgradient in the watershed. The primary evidence supporting this conclusion was: (1) age profiles at the well nests were better fit by an exponential-piston flow model than an exponential model (Fig. 3.3), suggesting little to no recharge in zones near the well nests and (2) a general lack of discharge of very young groundwater (0-10 yr old) to West Bear Creek (Figs. 3.4-3.6) suggests a lack of recharge near the stream channel. Near-channel areas of little to no recharge were reported in an agricultural watershed in Minnesota [*Böhlke et al.*, 2002]. The floodplain areas adjacent to West Bear Creek (~200 m wide) may restrict recharge near the stream, due to poorly-drained soils and/or shallow low permeability layers. In an environment with such spatial variation in recharge, the exponential model will not be the most realistic framework for extracting information such as recharge rate or *MTT* from age-dating tracer data. Streambed sampling to obtain coupled point estimates of groundwater age and seepage rate may offer a better approach, independent of and simple lumped model, for estimation of *MTT* (as the flow-weighted mean groundwater age from the sampling points). But then, it is important to fully sample the groundwater discharge in an unbiased manner. Sampling across the channel, the primary expected axis of variation in the age of groundwater discharge (Figs. 3.4-3.6), and along the channel, a potential secondary axis of variation, is important. Sampling streambed groundwater in small shallow tributaries to the main channel, which was not done at West Bear Creek, should be a goal of future work, in order to assess whether such sampling gives a different picture of the groundwater transit time distribution compared to that from the main channel.

The exponential model has been widely used in interpreting groundwater age dating tracer data but is brought into question for the West Bear Creek area by age data from wells and streambed sampling (Fig. 3.4). The exponential model would predict that N flux from the surficial aquifer (f_{NO3}) would decline rapidly in response to lower N use in the watershed but also show long tailing behavior in later years. Using age distributions and N concentration

data from streambed sampling to predict future f_{NO3} from the surficial aquifer suggested that rather than decreasing rapidly, f_{NO3} may hold steady or increase in the next 1-2 decades before declining (Fig. 3.7). Predictions of f_{NO3} from the distribution of age (exponential-piston flow model) and N concentrations observed at the well nest suggested that f_{NO3} could decrease rapidly after about a decade, assuming no N use in the watershed in the future, or f_{NO3} would remain steady if current N use persisted. However, applying an age distribution more consistent with streambed sampling (gamma distribution, $\alpha = 18$) to the N concentrations from wells gave a view of future f_{NO3} similar to streambed sampling where f_{NO3} is likely to increase before decreasing; this result highlights the uncertainty associated with predicting N fluxes from well data, because with well sampling alone the actual age distribution of groundwater discharge to streams is unknown. It also highlights the potential significance of the observed gamma distribution in the age of discharging groundwater, in unconfined aquifers both at West Bear Creek [*Gilmore et al.*, 2015a] and Wisconsin [*Browne and Guldan*, 2005, with additional analysis by *Gilmore et al.* 2015a]. The distribution of age in the surficial aquifer at West Bear Creek, including the lack of young water observed at two well nests and from streambed sampling, could contribute to slower response to N management initiatives than would be predicted by age distribution models with a large fraction of young groundwater (e.g., exponential model); on the bright side, the observed gamma distribution does not have the long slow tailing behavior of the exponential distribution, suggesting pollutant flushing from the aquifer may start slower but finish sooner.

All sampling approaches suggested that N use in the watershed has declined in the last decade (based on lower $[NO_3^-]$ in < 10 year-old groundwater, Fig. 3.4), which raised the question of whether NO_3^- inputs were currently higher than outputs at the time of sampling (2012–2013). For the larger-scale streambed sampling (2.5 km reach) and well nest sampling we compared the current rate of input of NO_3^- to the groundwater system, f_{NR}, (recharge $[NO_3^-]$ in young groundwater times recharge rate) to the total NO_3^- loss rate from the groundwater system (f_{OUT}) by discharge to surface water (f_{NO3}) and denitrification in the

aquifer ($f_{N2\text{-}den}$). Results from both streambed and well nest sampling approaches suggested that f_{OUT} was less than f_{NR}, i.e., storage of nitrate in the aquifer was increasing slowly at a rate of about 28 mmol m^{-2} yr^{-1} (average of four estimates in Table 3.4, about one-fifth the rate of NO_3^- input to the aquifer by recharge); percentage uncertainty is high in the estimates of the rate of change of NO_3^- storage, but all four estimates are positive (Table 3.4) Agreement in f_{NO3} and $f_{N2\text{-}den}$ was strongest between streambed sampling and well sampling approaches when the latter was weighted by a lumped-parameter model that gave the best fit to the age distribution from streambed sampling (gamma model). This highlights the potential importance of knowing the distribution of age in aquifer discharge, which is an advantage of streambed sampling.

It is ideal to combine both streambed sampling and well nest sampling, when site logistics and budget allow, because corroborating evidence of important hydrologic processes may be observed (e.g., in this work, evidence of spatial variability in recharge). However, for a study focused on legacy groundwater contamination, and the contaminant discharge from aquifer to the streams is of concern, the streambed sampling approach has the advantage of providing direct measure of actual fluxes from the groundwater system to a stream reach, robust estimates of *MTT* for the aquifer, and potential insights into the distribution of groundwater age and nitrogen concentrations in aquifer discharge.

REFERENCES

Aeschbach-Hertig, W., and D. K. Solomon (2013), Noble gas thermometry in groundwater hydrology, in *The Noble Gases as Geochemical Tracers*, edited by P. Burnard, pp. 81–122, Springer, Heidelberg.

Aeschbach-Hertig, W., H. El-Gamal, M. Wieser, and L. Palcsu (2008), Modeling excess air and degassing in groundwater by equilibrium partitioning with a gas phase, *Water Resour. Res.*, *44*(8), W08449, doi:10.1029/2007WR006454.

Amin, I. E., and M. E. Campana (1996), A general lumped parameter model for the interpretation of tracer data and transit time calculation in hydrologic systems, *J. of Hydrol.*, *179*(1–4), 1–21, doi:10.1016/0022-1694(95)02880-3.

Böhlke, J. K. (2002), Groundwater recharge and agricultural contamination, *Hydrogeol. J.*, *10*, 153–179, doi:10.1007/s10040-001-0183-3.

Böhlke, J. K., and J. M. Denver (1995), Combined use of groundwater dating, chemical and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, atlantic coastal plain, Maryland, *Water Resour. Res.*, *31*(9), 2319, doi:10.1029/95WR01584.

Böhlke, J. K., R. Wanty, M. Tuttle, G. Delin, and M. Landon (2002), Denitrification in the recharge area and discharge area of a transient agricultural nitrate plume in a glacial outwash sand aquifer, Minnesota, *Water Resour. Res.*, *38*(7), 10–1, doi:10.1029/2001WR000663.

Böhlke, J. K., M. E. O'Connell, and K. L. Prestegaard (2007), Ground water stratification and delivery of nitrate to an incised stream under varying flow conditions, *J. Environ. Qual.*, *36*(3), 664, doi:10.2134/jeq2006.0084.

Browne, B. A., and N. M. Guldan (2005), Understanding long-term baseflow water quality trends using a synoptic survey of the ground water-surface water interface, Central Wisconsin, *J. Environ. Qual.*, *34*(3), 825–35.

CGIA: Hydrology-Lines by River Basin (2004). [available at http://www.lib.ncsu.edu/gis/search/datainfo.php?datasetid=2296_bp8]

Coes, A., T. Spruill, and M. Thomasson (2007), Multiple-method estimation of recharge rates at diverse locations in the North Carolina Coastal Plain, USA, *Hydrogeol. J.*, *15*(4), 773–788, doi:10.1007/s10040-006-0123-3.

Cook, P. G., and D. K. Solomon (1997), Recent advances in dating young groundwater: chlorofluorocarbons, $^{3}H^{3}He$ and ^{85}Kr, *J. of Hydrol.*, *191*(1–4), 245–265, doi:10.1016/S0022-1694(96)03051-X.

Daniels, R. B., S. W. Buol, H. J. Kleiss, and C. A. Ditzler (1999), *Soil systems in North Carolina*, *Technical Bulletin 314*, North Carolina State University Soil Science Department, Raleigh, NC.

Delin, G. N., R. W. Healy, M. K. Landon, and J. K. Böhlke (2000), Effects of topography and soil properties on recharge at two sites in an agricultural field, *J. Am. Water Resour. Assoc.*, *36*(6), 1401–1416, doi:10.1111/j.1752-1688.2000.tb05735.x.

Dunkle, S. A., L. N. Plummer, E. Busenberg, P. J. Phillips, J. M. Denver, P. A. Hamilton, R. L. Michel, and T. B. Coplen (1993), Chlorofluorocarbons (CCl_3F and CCl_2F_2) as dating tools and hydrologic tracers in shallow groundwater of the Delmarva Peninsula, Atlantic Coastal Plain, United States, *Water Resour. Res.*, *29*(12), 3837–3860, doi:10.1029/93WR02073.

Elkins, B. (2007), Nutrient transport between the riparian and hyporheic zones in a small agricultural watershed in the North Carolina Coastal Plain, M.S. thesis, Dept. of Geol. Sci., East Carolina Univ., Greenville, N. C.

Farrell, K. M., W. B. Harris, D. J. Mallinson, S. J. Culver, S. R. Riggs, J. Pierson, J. M. Self-Trail, and J. C. Lautier (2012), Standardizing texture and facies codes for a process-based classification of clastic sediment and rock, *J. Sediment. Res*, *82*(6), 364–378, doi:10.2110/jsr.2012.30.

Friedrich, R., G. Vero, C. von Rohden, B. Lessmann, R. Kipfer, and W. Aeschbach-Hertig (2013), Factors controlling terrigenic SF_6 in young groundwater of the Odenwald region (Germany), *Appl. Geochem.*, *33*, 318–329, doi:10.1016/j.apgeochem.2013.03.002.

Genereux, D. P., S. Leahy, H. Mitasova, C. D. Kennedy, and D. R. Corbett (2008), Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA, *J. Hydrol.*, *358*(3–4), 332–353, doi:10.1016/j.jhydrol.2008.06.017.

Gilmore, T. E., D. P. Genereux, D. K. Solomon, and J. E. Solder (2015a), Groundwater age and mean transit time from streambed sampling in an agricultural coastal plain watershed, North Carolina, USA, *in prep*.

Gilmore, T. E., D. P. Genereux, D. K. Solomon, J. E. Solder, Kimball, B.A., H. Mitasova, and F. Birgand (2015b), Quantifying the fate of agricultural nitrogen in an unconfined aquifer: stream-based observations at three measurement scales, *in prep*.

Green, C. T., L. J. Puckett, J. K. Böhlke, B. A. Bekins, S. P. Phillips, L. J. Kauffman, J. M. Denver, and H. M. Johnson (2008), Limited occurrence of denitrification in four shallow aquifers in agricultural areas of the United States, *J. Environ. Qual.*, *37*(3), 994, doi:10.2134/jeq2006.0419.

Heaton, T. H. E., and J. C. Vogel (1981), "Excess air" in groundwater, *J. Hydrol.*, *50*, 201–216, doi:10.1016/0022-1694(81)90070-6.

Jaynes, D. B., T. S. Colvin, D. L. Karlen, C. A. Cambardella, and D. W. Meek (2001), Nitrate loss in subsurface drainage as affected by nitrogen fertilizer rate, *J. Environ. Qual.*, *30*(4), 1305–1314.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2007), Design of a light-oil piezomanometer for measurement of hydraulic head differences and collection of groundwater samples, *Water Resour. Res.*, *43*, doi:10.1029/2007WR005904.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2009a), Relationships among groundwater age, denitrification, and the coupled groundwater and nitrogen fluxes through a streambed, *Water Resour. Res.*, *45*, doi:10.1029/2008WR007400.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2009b), Spatial and temporal dynamics of coupled groundwater and nitrogen fluxes through a streambed in an agricultural watershed, *Water Resour. Res.*, *45*, doi:10.1029/2008WR007397.

Kirchner, J. W., D. Tetzlaff, and C. Soulsby (2010), Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments, *Hydrol. Process.*, *24*(12), 1631–1645, doi:10.1002/hyp.7676.

Kozuskanich, J., C. T. Simmons, and P. G. Cook (2014), Estimating recharge rate from groundwater age using a simplified analytical approach: Applicability and error estimation in heterogeneous porous media, *J. Hydrol.*, *511*, 290–294, doi:10.1016/j.jhydrol.2014.01.058.

Lebo, M., H. Paerl, and B. Peierls (2012), Evaluation of progress in achieving TMDL mandated nitrogen reductions in the Neuse River Basin, North Carolina, *Environ. Manage.*, *49*(1), 253–266, doi:10.1007/s00267-011-9774-5.

Liao, L., C. T. Green, B. A. Bekins, and J. K. Böhlke (2012), Factors controlling nitrate fluxes in groundwater in agricultural areas, *Water Resour. Res.*, *48*(6), n/a–n/a, doi:10.1029/2011WR011008.

Lindsey, B. D., S. W. Phillips, C. A. Donnelly, G. K. Speiran, L. N. Plummer, J. K. Böhlke, M. J. Focazio, W. C. Burton, and E. Busenberg (2003), Residence times and nitrate transport

in ground water discharging to streams in the Chesapeake Bay watershed, *Water-Resources Investigations Report 03-4035*, U.S. Geol. Surv., Reston, Va.

McMahon, P. B., L. N. Plummer, J. K. Böhlke, S. D. Shapiro, and S. R. Hinkle (2011), A comparison of recharge rates in aquifers of the United States based on groundwater-age data, *Hydrogeol. J.*, *19*, 779–800, doi:10.1007/s10040-011-0722-5.

Meals, D. W., S. A. Dressing, and T. E. Davenport (2010), Lag time in water quality response to best management practices: A review, *J. of Environ. Qual.*, *39*(1), 85, doi:10.2134/jeq2009.0108.

Mew, H. E., D. K. Hirth, D. Van Lewis, R. B. Daniels, and A. J. Keyworth (2002), Methodology for compiling ground water recharge maps in the Piedmont and Coastal Plain provinces of North Carolina, *Ground Water Bulletin Number 25*, NC Dept. of Environment and Natural Resources, Raleigh, NC.

Modica, E., H. T. Buxton, and L. N. Plummer (1998), Evaluating the source and residence times of groundwater seepage to streams, New Jersey Coastal Plain, *Water Resour. Res.*, *34*, 2797, doi:10.1029/98WR02472.

NCDENR. (2012) *2012 Annual progress report on the Neuse River Agricultural Rule*. NC Dept. of Environment and Natural Resources, Division of Water Resources, Raleigh, NC. [available at: http://portal.ncdenr.org/web/wq/ps/nps/neuse]

Nolan, B. T., A. L. Baehr, and L. J. Kauffman (2003), Spatial variability of groundwater recharge and its effect on shallow groundwater quality in southern New Jersey, *Vadose Zone J.*, *2*(4), 677, doi:10.2136/vzj2003.6770.

Osmond, D. L., D. L. K. Hoag, A. E. Luloff, D. W. Meals, and K. Neas (2015), Farmers' use of nutrient management: lessons from watershed case studies, *J. Environ. Qual.*, *44*(2), 382–390, doi:doi:10.2134/jeq2014.02.0091.

Poreda, R. J., T. E. Cerling, and D. K. Salomon (1988), Tritium and helium isotopes as hydrologic tracers in a shallow unconfined aquifer, *J. Hydrol.*, *103*(1–2), 1–9, doi:10.1016/0022-1694(88)90002-9.

Puckett, L. J., C. Zamora, H. Essaid, J. T. Wilson, H. M. Johnson, M. J. Brayton, and J. R. Vogel (2008), Transport and fate of nitrate at the ground-water/surface-water interface, *J. Environ. Qual.*, *37*(3), 1034, doi:10.2134/jeq2006.0550.

Ruddy, B.C., Lorenz, D.L., and Mueller, D.K. (2006). County-level estimates of nutrient inputs to the land surface of the conterminous United States, 1982–2001*, U.S. Geological Survey Scientific Investigations Report 2006-5012*, U.S. Geol. Surv., Reston, Va.

Schlosser, P., M. Stute, H. Dörr, C. Sonntag, and K. O. Münnich (1988), Tritium/^{3}He dating of shallow groundwater, *Earth Planet. Sc. Lett.*, *89*(3–4), 353–362, doi:10.1016/0012-821X(88)90122-7.

Solomon, D. K., S. L. Schiff, R. J. Poreda, and W. B. Clarke (1993), A validation of the ^{3}H/^{3}He method for determining groundwater recharge, *Water Resour. Res.*, *29*(9), 2951–2962, doi:10.1029/93WR00968.

Solomon, D. K., P. G. Cook, and L. N. Plummer (2006), Models of groundwater ages and residence times, in *Use of Chlorofluorocarbons in Hydrology: A Guidebook*, edited by E. Busenberg, pp. 73–78, International Atomic Energy Agency, Vienna, Austria.

Solomon, D. K., D. P. Genereux, L. N. Plummer, and E. Busenberg (2010), Testing mixing models of old and young groundwater in a tropical lowland rain forest with environmental tracers, *Water Resour. Res.*, *46*(4), W04518, doi:10.1029/2009WR008341.

Solomon, D. K., T. E. Gilmore, J. E. Solder, B. A. Kimball, and D. P. Genereux (2015), Evaluating a groundwater flow system using SF_6 in streamflow, *in prep*.

Spruill, T. B., A. J. Tesoriero, H. E. Mew, K. M. Farrell, S. L. Harden, A. B. Colosimo, and S. R. Kraemer (2004), Geochemistry and characteristics of nitrogen transport at a confined animal feeding operation in a coastal plain agricultural watershed, and implications for nutrient loading in the Neuse River basin, North Carolina, 1999–2002, *Scientific Investigations Report 2004-5283*, U.S. Geol. Surv., Reston, Va.

Stelzer, R., D. Drover, S. Eggert, and M. Muldoon (2011), Nitrate retention in a sand plains stream and the importance of groundwater discharge, *Biogeochemistry*, *103*(1), 91–107, doi:10.1007/s10533-010-9449-y.

Stewart, M. K., U. Morgenstern, and J. J. McDonnell (2010), Truncation of stream residence time: how the use of stable isotopes has skewed our concept of streamwater age and origin, *Hydrol. Process.*, *24*(12), 1646–1659, doi:10.1002/hyp.7576.

Tesoriero, A. J. (2005), Nitrogen transport and transformations in a coastal plain watershed: Influence of geomorphology on flow paths and residence times, *Water Resour. Res.*, *41*(2), doi:10.1029/2003WR002953.

Tesoriero, A. J., T. B. Spruill, and J. L. Eimers (2004), Geochemistry of shallow ground water in coastal plain environments in the southeastern United States: implications for aquifer susceptibility, *Appl. Geochem.*, *19*(9), 1471–1482, doi:10.1016/j.apgeochem.2004.01.021.

Tesoriero, A. J., J. H. Duff, D. A. Saad, N. E. Spahr, and D. M. Wolock (2013), Vulnerability of streams to legacy nitrate sources, *Environ. Sci. Technol.*, *47*(8), 3623–3629, doi:10.1021/es305026x.

Usry, B. P. (2006), Nitrogen loading in the Neuse River Basin, North Carolina: The Rivernet Monitoring Program, M.S. thesis, Dept. of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC.

Welch, H. L., Green, C. T., & Coupe, R. H. (2011). The fate and transport of nitrate in shallow groundwater in northwestern Mississippi, USA. *Hydrogeol. J.*, *19*(6), 1239–1252. http://doi.org/10.1007/s10040-011-0748-8

Winner, M. D., and R. W. Coble (1996), Hydrogeologic framework of the North Carolina Coastal Plain, *U.S. Geological Survey Professional Paper 1404-I*, U.S. Geol. Surv., Reston, Va.

Wolock, D. M. (2003), Base-flow index grid for the conterminous United States. *U.S. Geological Survey Open-File Report 03−263*, U.S. Geol. Surv., Reston, Va. [available at: http://water.usgs.gov/lookup/getspatial?bfi48grd]

Visser, A., Broers, H. P., & Bierkens, M. F. P. (2007). Dating degassed groundwater with $^{3}H/^{3}He$. *Water Resources Research*, *43*(10), W10434. doi:10.1029/2006WR005847
Vogel, J.C. (1967), Investigation of groundwater flow with radiocarbon, in *Isotopes in Hydrology (Proc. Int. Conf. Vienna, 1966)*, pp. 355-369, IAEA, Vienna, Austria.

APPENDICES

Appendix A – Uncertainty in Reach Mass-balance Calculations

A1. Uncertainty in groundwater flux

Assuming stream discharge is at steady state, groundwater flux is calculated as

$$v_{RMB} = \frac{Q_{down} - Q_{up} - \sum Q_{trib}}{wL} \tag{A1}$$

where v_{RMB} is groundwater flux into the stream (m/day) and the subscript RMB indicates the estimate is from reach mass-balance, Q_{down} and Q_{up} *are* the stream discharge at the downstream and upstream ends of the stream reach, respectively, ΣQ_{trib} is the inflow from any tributaries contained in the reach, L is the length of the reach, and w is the average width of the stream.

Standard propagation of error through a function $y = f(x_1, x_2, \ldots)$ requires estimates of uncertainty (W, Table A1) in each variable x_i and partial derivatives of y with respect to each variable:

$$W_y = \sqrt{\sum_i \left(\frac{\partial y}{\partial x_i} W_{x_i} \right)^2} \tag{A2}$$

The partial derivatives for equation A1 are:

$$\frac{\partial v}{\partial Q_{down}} = \frac{1}{wL} \tag{A3}$$

$$\frac{\partial v}{\partial Q_{trib}} = \frac{\partial v}{\partial Q_{up}} = \frac{-1}{wL} \tag{A4}$$

$$\frac{\partial v}{\partial w} = \frac{-\left(Q_{down} - Q_{up} - \sum Q_{trib}\right)}{w^2 L} \tag{A5}$$

$$\frac{\partial v}{\partial L} = \frac{-\left(Q_{down} - Q_{up} - \sum Q_{trib}\right)}{wL^2} \quad \text{(A6)}$$

Note that equation A4 was calculated for each tributary in the summation in equation A1.

Table A1. Parameters (with units) used to calculate uncertainties in groundwater flux, groundwater nitrate flux, and groundwater nitrate concentrations for the July 2012 and March 2013 field campaigns.

Parameter, variable	Units	July 2012	March 2013
		uncertainty, W^a	uncertainty, W^a
stream discharge, Q, from dilution of bromide	L/sec	3.3%	3-4%
stream width, w	meters	8%	8%
stream length, L	meters	1%	1%
stream bromide conc., [Br^-]	mg/L	3%	2-4%
mass flow of injectate, M_{inj}	mg/sec	1% (assumed)	1%
stream anion conc.	µM	5% for NO_3^- and Cl^-	5% for NO_3^- and Cl^-
Flow Tracker stream discharge, Q	L/sec	not applicable	2-5%

[a]All uncertainty estimates are for the 95% confidence level. Note that percent uncertainty must be multiplied by the parameter estimate to calculate the W used in uncertainty equations presented here.

Uncertainty in stream discharge is based on uncertainty in the injection rate for the conservative tracer (Br^-) and analytical uncertainty for [Br^-] in stream water. Discharge was calculated as:

$$Q = \frac{Q_{inj}[\mathrm{Br}^-]_{inj}}{[\mathrm{Br}^-]_{down} - [\mathrm{Br}^-]_0} = \frac{M_{inj}}{[\mathrm{Br}^-]_{down} - [\mathrm{Br}^-]_0} \quad \text{(A7)}$$

where Q is the stream discharge at any measurement location downstream of the Br injection, $[Br^-]_{inj}$ and $[Br^-]_0$ are the bromide concentration in the injectate and background bromide concentration in the stream (mg/L), respectively, Q_{inj} is the flow rate of the injectate (L/sec), and *Minj* is rate of injection of Br (mg/s).

Partial derivatives of equation A7, for use in equation A2, are:

$$\frac{\partial Q}{\partial M_{inj}} = \frac{1}{[\mathrm{Br}^-]_{down} - [\mathrm{Br}^-]_0} \tag{A8}$$

$$\frac{\partial Q}{\partial [\mathrm{Br}^-]_0} = \frac{M_{inj}}{\left([\mathrm{Br}^-]_{down} - [\mathrm{Br}^-]_0\right)^2} \tag{A9}$$

$$\frac{\partial Q}{\partial [\mathrm{Br}^-]_{down}} = \frac{-M_{inj}}{\left([\mathrm{Br}^-]_{down} - [\mathrm{Br}^-]_0\right)^2} \tag{A10}$$

In March 2013, uncertainty in the mass flow of injectate was 1%, based on the 95% confidence interval for repeated measurements, and was assumed to be 1% in July 2012. Uncertainty in $[Br^-]_{down}$ and $[Br^-]_0$ were set equal to the analytical uncertainty (3%) in July 2012. In March 2013, $[Br^-]$ from automated samplers was averaged over 4-5 hours (n = 6-9), and $W_{[Br]down}$ was determined from variability in the observed stream water $[Br^-]$ (Table A1). Uncertainty in stream discharge (W_Q) was then calculated by substituting equations A8 – A10 into equation A2.

Uncertainty in each tributary discharge to West Bear Creek (W_{Qtrib}) in March 2013 was calculated by adding in quadrature the uncertainties in Flow Tracker ADV measurements (or discharge calculated from Br^- dilution) taken in the main channel upstream and downstream of the tributary:

$$W_{\Delta Q} = \sqrt{{W_{Q_1}}^2 + {W_{Q_2}}^2} \tag{A11}$$

where Q_1 and Q_2 are discharges at the upstream and downstream ends, respectively, of the reach with the tributary (from Flow Tracker ADV measurements or Br^- dilution), and $\Delta Q = Q_2 - Q_1$.

In the case of Trib 1 and Trib 2, stream discharge was measured by Flow Tracker at 810 m and 860 m in March 2013 (530.6 ± 14.3 and 543.1 ± 13.0 L/sec, respectively), which

bracketed both tributaries (Fig. 1.1 and 1.3, main text of this article). Thus we assumed the sum of the flow in the two tributaries was 543.1 – 530.6 = 12.5 L/sec. Also, we assumed that discharge in Trib 1 was 4x that in Trib 2, based on tributary channel size and other observations (Table A2; Trib 2 is a small vegetated drainage ditch, Trib 1 has wider sandy un-vegetated streambed). Uncertainty in Trib 1 and Trib 2 estimates were set equal to the uncertainty in ΔQ between the 810 m and 860 m Flow Tracker measurements (155%, based on ΔQ = 12.5 ± 19.4 L/sec).

We effectively assumed zero (negligible) groundwater inflow between the Flow Tracker measurements taken above and below Trib 1 and Trib 2. Flow Tracker measurements were about 50 m apart and groundwater discharge from points at the nearest transect (715 m) suggested negligible groundwater flux between the Flow Tracker measurements was a reasonable assumption (v = 0.31 m/day multiplied by 6.5x50 m^2, the area of the streambed in this 50 m reach, gives 1.2 L/sec, a value much less than the 19.4 L/sec uncertainty in ΔQ). Similarly, for the streambed area between stream discharge measurements at 1700 m and 1800 m, points at 1700 m suggested groundwater inflow of about 1.2 L/sec. We could have incorporated these groundwater discharge estimates from nearby point transects in reach mass-balance calculations, but given that the purpose of this study was a comparison of methods, we calculated the reach mass-balance results independently of point data.

Table A2. Tributary estimates and uncertainties for March 2013 reach mass-balance.

Tributary, location	Q_{trib} (L/sec)	W_{Qtrib}	basis for tributary and uncertainty estimate (ΔQ in L/sec)
Trib 1, 830 m	10.0	15.5	Flow Tracker at 810 m and 860 m, ΔQ = 12.5 ± 20.2; 10 L/sec estimate based on field estimate and channel size roughly 4x larger than Trib 2
Trib 2, 850 m	2.5	3.9	Flow Tracker at 810 m and 860 m, ΔQ = 12.5 ± 20.2
Trib 3, 1790 m	15.7	23.2	Flow Tracker at 1700 m and Br^- dilution at 1800 m
Trib 4, 1840 m	47.7	25.0	Br^- dilution at 1800 m and Flow Tracker at 1700 m
Trib 5, 2400 m	29.3	30.1	Flow Tracker at 2380 m and 2420 m

In July 2012, Flow Tracker measurements were not taken upstream and downstream of tributaries, as tributary inputs were negligible with the exception of Trib 1 (830 m) and Trib 5 (2400 m). Stream discharge was calculated at high resolution (100 m), so ΔQ was calculated at 100 m intervals, and the two 100 m reaches where tributaries entered the stream had clearly elevated water inputs (Fig. 1.3, main text, where there is an abrupt increase in Q at 900 m and 2500 m, just below Trib 1 and Trib 5, respectively). For each 100 m reach containing a tributary (800-900 m for Trib 1, 2300-2400 m for Trib 5), groundwater input was assumed to be equal to the mean ΔQ for the 100 m reaches immediately upstream and downstream (e.g., groundwater input to 800-900 m was taken as the average of groundwater input to 700-800 m and 900-1000 m). The assumed groundwater input was then subtracted from ΔQ for the section with the tributary, and the remainder was used as an estimate of tributary inflow (Table A3):

$$Q_{trib} = \Delta Q_2 - 0.5\left(\Delta Q_1 + \Delta Q_3\right) \tag{A12}$$

where ΔQ_2, ΔQ_1, and ΔQ_3 are ΔQ values for the reach containing the tributary, the reach immediately upstream, and the reach immediately downstream, respectively. Tributary inflow estimates were also verified by measuring tributary inflow at a later date, under similar stream discharge. Uncertainty in each July 2012 tributary input (W_{Qtrib}) was calculated as:

$$W_{Qtrib} = \sqrt{W_{\Delta Q_2}{}^2 + \left(-0.5W_{\Delta Q_1}\right)^2 + \left(-0.5W_{\Delta Q_3}\right)^2} \quad \text{(A13)}$$

Table A3. Tributary estimates and uncertainty for July 2012 reach mass-balance.

Reach (m)	US[a] Q L/sec	DS[a] Q L/sec	ΔQ L/sec	Q_{trib} L/sec	W_{Qtrib} L/sec
700-800	72.14	74.05	1.91	n.a.	n.a.
800-900	74.05	80.02	5.98	4.79	4.40
900-1000[b]	80.02	80.49	0.465	n.a.	n.a.
2300-2400	96.51	96.94	0.430	n.a.	n.a.
2400-2500	96.94	103.4	6.44	6.09	5.73
2500-2600	103.4	103.6	0.252	n.a.	n.a.

[a]"US" = upstream, "DS" = downstream

[b]discharge at 1000 m was interpolated from 900 m and 1100 m discharge, because no Br^- data were available

Uncertainty in stream width was based on variability in stream width measurements taken in July 2012 and March 2013. Uncertainty in stream length was assumed to be small (Table A1) because the stream was manually marked off using a 100m measuring tape.

Total uncertainty in groundwater seepage rate, *v*, was calculated by substituting equations A3-A6 into equation A2.

A2. Uncertainty in groundwater nitrate flux from reach mass-balance

Using the reach mass-balance approach, nitrate flux out of the groundwater system and into a stream reach was calculated as:

$$f_{RMB} = \frac{Q_{down}C_{down} - \sum Q_{trib}C_{trib} - Q_{up}C_{up}}{wL} \quad \text{(A14)}$$

The variable *C* is the concentration of the nutrient of interest (e.g., nitrate) at the location indicated by the subscript. Concentrations and discharges were measured or estimated individually for each tributary in the summation. The partial derivatives of equation A14, for use in equation A2, are:

$$\frac{\partial f_{RMB}}{\partial Q_{down}} = \frac{C_{down}}{wL} \tag{A15}$$

$$\frac{\partial f_{RMB}}{\partial Q_{trib}} = \frac{-C_{trib}}{wL} \tag{A16}$$

$$\frac{\partial f_{RMB}}{\partial Q_{up}} = \frac{-C_{up}}{wL} \tag{A17}$$

$$\frac{\partial f_{RMB}}{\partial C_{down}} = \frac{Q_{down}}{wL} \tag{A18}$$

$$\frac{\partial f_{RMB}}{\partial C_{trib}} = \frac{-Q_{trib}}{wL} \tag{A19}$$

$$\frac{\partial f_{RMB}}{\partial C_{up}} = \frac{-Q_{up}}{wL} \tag{A20}$$

$$\frac{\partial f_{RMB}}{\partial w} = \frac{-\left(Q_{down}C_{down} - \sum Q_{trib}C_{trib} - Q_{up}C_{up}\right)}{w^2 L} \tag{A21}$$

$$\frac{\partial f_{RMB}}{\partial L} = \frac{-\left(Q_{down}C_{down} - \sum Q_{trib}C_{trib} - Q_{up}C_{up}\right)}{wL^2} \tag{A22}$$

Uncertainty in stream water concentrations of nitrate and chloride was 5%. In July 2012, background nitrate samples (7/16/2012) were collected and treated (acidified and chilled in the field). The primary synoptic sampling occurred on 7/18/2012, however, and only Br^- samples (unacidified, not chilled) were collected. Thus, to calculate groundwater nitrate concentrations, there were two options: (1) use the stream water [NO_3^-] from the 7/16/2012 samples and try to estimate what the stream discharge would have been on that day from the [Br^-] data collected on 7/18/2012 and the overall slow drop in stream discharge (e.g., at the USGS gauge downstream on Bear Creek), or (2) use the stream water nitrate [NO_3^-] from the 7/18/2012 samples, which were directly linked to the 7/18/2012 [Br^-] and discharge data

(i.e., discharge was calculated from [Br^-] in the same samples). We took the latter approach and felt it more appropriate for reach mass-balance, because the [NO_3^-] and stream discharge were based on the same set of samples, and it's very unlikely there was significant degradation or production of nitrate in the Br^- samples (for example, in March 2013 all [NO_3^-] fell within ±5% of the regression line, and all but two were within ±5% of the 1:1 line, when [NO_3^-] from "nitrate samples" was compared with [NO_3^-] from "Br^- samples", Fig. A1). Ultimately, the major findings of the study, including the observed trend of decreasing groundwater nitrate concentration with increasing sampling integration scale, were not sensitive to which of the two approaches was used. In March 2013, averaged stream water [NO_3^-] from automated samplers were used (roughly 5-hour sampling window, n=5-6 at each sampler) and uncertainty in [NO_3^-] was determined from variability during the sampling window.

In March 2013, nitrate samples were collected from tributaries during the field campaign. In July 2012, the tributary near 2400 m was sampled for nitrate during the field campaign, while the other tributary (at roughly 830 m) was sampled on three occasions (Jan., March, and Nov. 2013) and showed little variation (W was ±64 μM), so the mean concentration (652 μM) was used in the July 2012 analysis.

The total uncertainty in groundwater nitrate flux was then calculated by substituting equations A15 – A22 into equation A2.

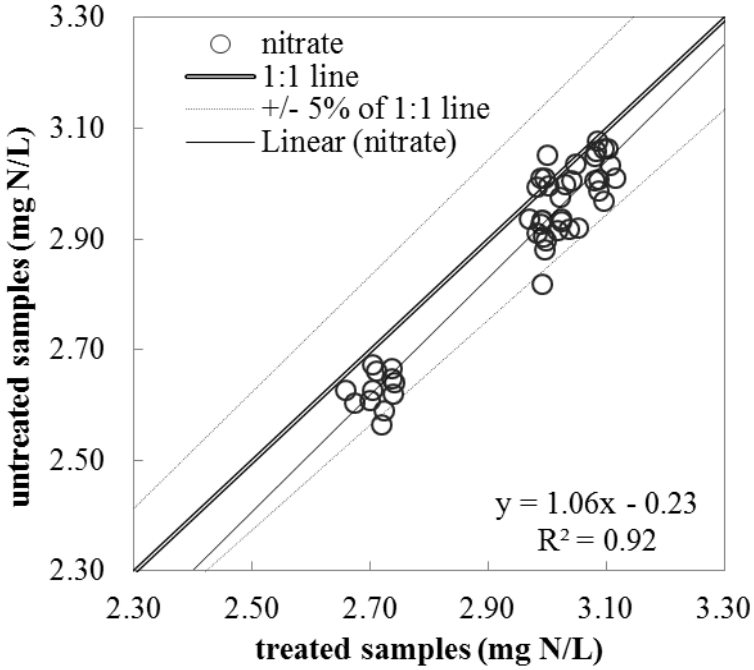

Figure A1. Comparison of concentrations in "treated" (acidified, chilled) nitrate samples analyzed on a LACHAT instrument at the North Carolina State Soil Science Analytical Services Lab (Raleigh, NC) with "untreated" (unacidified, stored at ambient temperature) nitrate samples analyzed by ion chromatograph at the USGS Water Science Center in Salt Lake City, Utah. Samples were collected during the March 2013 field campaign.

A3. Uncertainty in groundwater nitrate concentration from reach mass-balance

Assuming stream discharge is at steady state, groundwater nitrate concentration is calculated as

$$C_{RMB} = \frac{Q_{down}C_{down} - Q_{up}C_{up} - \sum Q_{trib}C_{trib}}{Q_{down} - Q_{up} - \sum Q_{trib}} \tag{A23}$$

The partial derivatives of equation A23, for use in equation A2, are:

$$\frac{\partial C_{RMB}}{\partial Q_{down}} = \frac{C_{down}}{Q_{down} - \sum Q_{trib} - Q_{up}} - \frac{Q_{down}C_{down} - \sum Q_{trib}C_{trib} - Q_{up}C_{up}}{\left(Q_{down} - \sum Q_{trib} - Q_{up}\right)^2} \tag{A24}$$

$$\frac{\partial C_{RMB}}{\partial Q_{trib}} = \frac{-C_{trib}}{Q_{down} - \sum Q_{trib} - Q_{up}} + \frac{Q_{down}C_{down} - Q_{trib}C_{trib} - Q_{up}C_{up}}{\left(Q_{down} - \sum Q_{trib} - Q_{up}\right)^2} \quad \text{(A25)}$$

$$\frac{\partial C_{RMB}}{\partial Q_{up}} = \frac{-C_{up}}{Q_{down} - \sum Q_{trib} - Q_{up}} + \frac{Q_{down}C_{down} - \sum Q_{trib}C_{trib} - Q_{up}C_{up}}{\left(Q_{down} - \sum Q_{trib} - Q_{up}\right)^2} \quad \text{(A26)}$$

$$\frac{\partial C_{RMB}}{\partial C_{down}} = \frac{Q_{down}}{Q_{down} - \sum Q_{trib} - Q_{up}} \quad \text{(A27)}$$

$$\frac{\partial C_{RMB}}{\partial C_{trib}} = \frac{-Q_{trib}}{Q_{down} - \sum Q_{trib} - Q_{up}} \quad \text{(A28)}$$

$$\frac{\partial C_{RMB}}{\partial C_{up}} = \frac{-Q_{up}}{Q_{down} - \sum Q_{trib} - Q_{up}} \quad \text{(A29)}$$

Total uncertainty in groundwater nitrate concentration for a given reach was calculated by substituting equations A24-A29 into equation A2.

A4. Uncertainty in reach mass-balance calculations of mean groundwater concentrations for N_2, Ar

The reach mass-balance calculation of dissolved gas concentration in the groundwater discharging into a stream reach was:

$$C_{RMB} = \frac{C_{down}Q_{down} - C_{up}Q_{up} + zkA\left(0.5\left(C_{up} + C_{down}\right) - C_{atm}\right)}{Q_{down} - Q_{up}} \quad \text{(A30)}$$

where C_{RMB} is the groundwater concentration of a given gas, C_{down} and C_{up} are the downstream and upstream concentrations in stream water, respectively, C_{atm} is the concentration the gas would have in stream water if the stream water were in equilibrium with the atmosphere at the stream temperature, k is the gas transfer velocity (m/day), and A is the area of the stream water surface ($A = wL$, in m^2). Units for Q_{up} and Q_{down} are L/s, and z is a unit conversion factor (0.0116 Lday/m^3 sec).

The partial derivatives of equation A30 for use in equation A2 are:

$$\frac{\partial C_{RMB}}{\partial C_{down}} = \frac{Q_{down} + 0.5zkA}{Q_{down} - Q_{up}} \tag{A31}$$

$$\frac{\partial C_{RMB}}{\partial C_{up}} = \frac{-Q_{up} + 0.5zkA}{Q_{down} - Q_{up}} \tag{A32}$$

Table A4. Parameters (with correct units for use in equation A30) used to calculate uncertainties in groundwater concentrations of dissolved gases for the July 2012 and March 2013 field campaigns.

Parameter, variable	Units	July 2012	March 2013
		uncertainty, W[a]	uncertainty, W[a]
stream discharge, Q,from dilution of bromide	L/sec	3.3%	3-4%
stream surface area, A	m^2	29%	16%
stream N_2 conc., [N_2]	μM	16%	3%
stream Ar conc., [Ar]	μM	3%	3%
gas transfer velocity, k	m/day	8%	8%

[a]All uncertainty estimates are for the 95% confidence level, with the exception of stream surface area (uncertainty in length was assumed) and gas transfer velocity (value from Baulch et al. (2010), their Appendix 1). Note that percent uncertainty must be multiplied by the parameter estimate to calculate the W used in equation A2.

$$\frac{\partial C_{RMB}}{\partial C_{atm}} = \frac{-zkA}{Q_{down} - Q_{up}} \tag{A33}$$

$$\frac{\partial C_{RMB}}{\partial Q_{down}} = \frac{C_{down}}{Q_{down} - Q_{up}} - \frac{Q_{down}C_{down} - Q_{up}C_{up} + zkA\left(0.5\left(C_{up} + C_{down}\right) - C_{atm}\right)}{\left(Q_{down} - Q_{up}\right)^2} \tag{A34}$$

$$\frac{\partial C_{RMB}}{\partial Q_{up}} = -\frac{C_{up}}{Q_{down} - Q_{up}} + \frac{Q_{down}C_{down} - Q_{up}C_{up} + zkA\left(0.5\left(C_{up} + C_{down}\right) - C_{atm}\right)}{\left(Q_{down} - Q_{up}\right)^2} \tag{A35}$$

$$\frac{\partial C_{RMB}}{\partial k} = \frac{zA\left(0.5\left(C_{up} + C_{down}\right) - C_{atm}\right)}{Q_{down} - Q_{up}} \tag{A36}$$

$$\frac{\partial C_{RMB}}{\partial A} = \frac{zk\left(0.5\left(C_{up} + C_{down}\right) - C_{atm}\right)}{Q_{down} - Q_{up}} \tag{A37}$$

Total uncertainty in groundwater concentration of volatile tracers for a given reach was calculated by substituting equations A31-A37 into equation A2 and using uncertainty values in Table A4.

REFERENCES

Baulch, H. M., J. J. Venkiteswaran, P. J. Dillon, and R. Maranger (2010), Revisiting the application of open-channel estimates of denitrification, *Limnol. Oceanogr. Meth.*, *8*, 202–215, doi:10.4319/lom.2010.8.202.

Appendix B – Dissolved Gas Modeling and N_2-den Results

B1. Excess N_2 model

The CE model formulation we used was:

$$C_{i-\mathrm{mod}} = \frac{C_i^{eq}(1 + AH_{i-rech})}{(1 + BH_{i-sam})} \tag{B1}$$

where $C_{i\text{-mod}}$ is the modeled concentration of gas i (i = N_2, Ne, Ar), C_i^{eq} is the solubility equilibrium concentration of gas i at recharge conditions (recharge temperature, atmospheric pressure and recharge salinity), $H_{i\text{-}rech}$ and $H_{i\text{-}samp}$ are the Henry's Law coefficients for gas i at recharge conditions and sampling conditions (two potentially different temperatures), respectively, and A and B are gas to water volume ratios in the porous medium at recharge and sampling, respectively (Aeschbach-Hertig et al., 2008). Solubility equilibrium concentrations were calculated using solubility equations for Ne (Weiss, 1971), Ar (Weiss, 1970) and N_2 (Benson and Krause, 1976; Andrews, 1992). The recharge temperature used to calculate C_i^{eq} (12.8 C) was based on noble gas thermometry (e.g., Aeschbach-Hertig and Solomon, 2013) using dissolved gas data from well nests shown in Figure 2.1B in the main text.

The CE model was applied in one of two modes: (1) an "excess air" mode, or (2) a "degassing" mode. Excess air is the condition in which groundwater is supersaturated relative to solubility equilibrium with the atmosphere at the recharge temperature (Heaton and Vogel, 1981). Degassing is the condition in which dissolved gases have been stripped from solution due to the production of gas bubbles in the groundwater system (Aeschbach-Hertig et al., 2000, 2008, Visser et al., 2007, 2009). ΔNe was used to choose the appropriate mode of the CE model for each groundwater sample from a streambed sampling point or blanket. ΔNe is calculated as $([Ne]_{meas}/ [Ne]^{eq}-1)100\%$, where $[Ne]^{eq}$ is the concentration of Ne in solubility equilibrium with the atmosphere at the recharge temperature and the subscript *meas* indicates a measured concentration. The excess air mode was applied when

ΔNe was positive (indicating "excess" Ne dissolved in the groundwater), and the degassing mode was applied when ΔNe was negative.

The parameter A in equation B1 accounts for excess air, while the B parameter accounts for degassing. When ΔNe was positive (excess air) the parameter B was set to zero and the parameter A was optimized in the model; in this case, the model simplifies to a simple unfractionated excess air model, or the UA model described by Aeschbach-Hertig et al. (2008). When ΔNe was negative, the parameter A was set equal to the average A determined from well nests (2.3 mL of air/L of water) and the B parameter was optimized. Where only Ar was available (1 point in July 2012, 3 points in March 2013) and for all reach mass-balance analyses, ΔAr was used to select the appropriate model (excess air or degassing) and [Ar] was used to directly solve for either the A or B parameter in equation B1.

[N_2-den] was calculated according to equation 17 of Aeschbach-Hertig et al. (2008), re-written here for N_2 in equation B2, which accounts for degassing of samples. This assumes that all N_2 produced in the groundwater was produced prior to degassing, corresponding to for example the groundwater degassing as it moves upward toward discharge at the streambed and the bulk of denitrification occurring upgradient of, and perhaps causing, this degassing. For samples where no degassing occurred, B = 0, and [N_2-den] is simply the difference between measured and modeled N_2.

$$\left[N_2 - den\right] = \left(\left[N_2\right]_{meas} - \left[N_2\right]_{\mathrm{mod}}\right)\left(1 + BH_N\right) \qquad \text{(B2)}$$

B2. Recharge conditions modeled from well data

Two well nests (BC1, BC2) were installed (Figure 2.1B in main text) in early April 2013. Depths from ground surface to the tops of screens at the BC1 well nest were 5.9 m, 10.0 m, and 16.3 m, with the deepest well located at the top of the regional confining unit (Black Creek confining unit). At the BC2 nest, well depths were 5.5 m, 7.9 m, and 14.3 m (Table B1). Screens were 45 cm in length. During sampling (June 19-21, 2013), water table depth was about 1.8 m and 2.7 m for the BC1 and BC2 well nests, respectively (ground surface to the water level in the shallowest well at each nest).

Groundwater samples for analysis of dissolved noble gases were collected from the BC1 and BC2 well nests on June 19-21, 2013. Recharge temperature (RT) and excess air (A) were modeled using noble gas concentrations in an unfractionated excess air model (UA). Maximum and minimum allowable temperature was set to 30 and 8 deg C, respectively, while A was permitted to range from 0 to 0.030 (i.e., 0 to 30 mL of gas/L of water). The parameter A and recharge temperature were optimized to fit the noble gas data, resulting in a mean A of 0.0023 (vol/vol, s.d.=0.009, n = 6) and mean recharge temperature of 12.8 C (s.d. = 1.7 C).

Determination of values for two modeled parameters (RT and A) using four dissolved gases (Ne, Ar, Kr and Xe) gives 2 degrees of freedom. All χ^2 for recharge well noble gas models were less than $\chi^2_{0.01,2} = 9.2$, indicating that the misfit in individual models can be explained by the uncertainty in the noble gas concentrations (uncertainty due to analytical and sampling uncertainty was estimated at 3% for all gases). The combined data set gives a mean squared weighted deviate (MSWD) value of 2.9, which is slightly higher than $MSWD_{0.01,12} = 2.2$, suggesting a statistically significant misfit (Ballentine and Hall, 1999). This misfit suggests that factors other than excess air and recharge temperature may have influenced the observed noble gas concentrations. However, only a modest increase in estimated uncertainty in noble gas concentrations (to 3.5%) is necessary to achieve MSWD = 2.2. Although analytical uncertainty is well-known, sampling uncertainty is more ambiguous, thus an overall uncertainty of 3.5% in noble gas concentrations is not unrealistic.

Table B1. Noble gas data and model results for well nests.

Sample	Well Depth	Ne $x10^3$	Ar	Kr $x10^3$	Xe $x10^4$	RT	A	ΔNe	χ^2
(-)	(m)	(μM)	(μM)	(μM)	(μM)	(C)	(mL/L)	(%)	(-)
BC1S	5.9	9.48	15.5	3.33	5.20	15.5	1.1	11.2	7.8
BC1M	10.0	11.1	17.3	3.70	5.52	13.2	3.0	28.0	2.7
BC1D	16.3	11.5	17.9	3.89	6.19	10.8	3.1	29.6	8.1
BC2S	5.5	10.6	17.7	3.78	5.98	11.3	2.1	19.7	7.0
BC1M	7.9	9.86	16.5	3.58	5.42	13.5	1.4	13.7	3.3
BC2D	14.3	11.4	17.5	3.72	5.80	12.4	3.1	30.0	6.5

S = shallow, M = intermediate, and D = deep; "RT" = recharge temperature

B3. Uncertainty from derived recharge temperature and A parameter

Estimates of uncertainty in E_{den}, $[N_2]_{den}$, $[NO_3]^0_{FWM}$, and f_{NO3}^0 were determined using a sensitivity analysis where extreme values of recharge parameters (likely upper and lower bounds, based on variability in recharge parameters from modeling dissolved gases in well samples) were used in the CE model for each point sample. Effectively, these upper and lower bounds in recharge parameters translated to high and low gas content at recharge.

To induce high gas concentrations, low recharge temperature was used (9.3 C, which is 2σ lower than mean recharge temperature from well samples, where lower temperature results in higher solubility equilibrium concentrations) and for degassed samples (where excess air is set equal to a constant), high excess air was used (0.0041, which is 2σ higher than the mean of 0.0023). For the low gas concentration case, higher recharge temperature (16.3 C, 2σ higher than mean, which is very close to mean annual air temperature) and no excess air was assumed (mean A minus 2σ was very small, at 0.0005). Literature values for excess air near this study site (Kennedy et al. 2009a, their Table 2) suggest that upper and lower bounds of excess air used in our models are reasonable. Recharge temperatures derived from old waters (7.8-10.9, Kennedy and Genereux (2007)) are in the range of our lower recharge temperature constraint. Our upper constraint on recharge temperature is just below mean annual air temperature (16.5 C, Kennedy et al. (2009a)), which is arguably the highest

reasonable recharge temperature, given that about 60% of recharge in the North Carolina coastal plain occurs during winter months (Coes et al., 2007).

Uncertainty in E_{den}, [N_2-den], $[NO_3]^0_{FWM}$, and $f_{NO3}{}^0$ from using an estimated recharge temperature and amount of excess air derived from well nests was modest for point samples. The upper and lower bounds from the sensitivity analysis were no more than ± 13% different than our best estimate of the mean for each variable (Table B2). We view these differences as modeling uncertainties, which are likely minimum uncertainties in mean values derived from N_2 modeling. In particular, we note that 26 of 36 point sample [N_2] measurements for July 2012, and 20 of 30 point samples for March 2013 were from copper tube samples that were stored for months at room temperature. While there is some potential for denitrification in water samples stored in copper tubes, this appears not to have occurred, judging from regression of [N_2] in water samples from copper tubes (stored at ambient temperature at the University of Utah Noble Gas Lab in Salt Lake City, UT) versus [N_2] in water samples from glass bottles (stored in the dark at 4° C prior to analysis at the USGS Dissolved Gas Lab in Reston, VA) (Fig. B1). The regression slope was slightly less than one (0.93, $p \ll 0.01$) and the small intercept (53 μM) was not statistically significant ($p = 0.48$). Out of 45 samples (surface water, blankets, points, and well nests) where both [N_2] measurements were available, only about six samples (3 from points, 3 from well nests) show potential for biased-high copper tube [N_2].

Table B2. Results from dissolved gas models for point samples using end member recharge scenarios.

Recharge case	$[N_2]_{den}$	$[NO_3]^0_{FWM}$	E_{den}	f_{NO30}
	(µM)	(µM)	(%)	(mmol m^{-2} d^{-1})
		July 2012		
high gas[a]	368	1391	52%	510
low gas[b]	286	1209	49%	444
deviation[c]	13%	7%	4%	7%
		March 2013		
high gas	193	486	77%	220
low gas	193	543	78%	245
deviation	2%	6%	2%	6%

[a]The "high gas" case assumed low recharge temperature and high excess air. [b]The "low gas" case assumed high recharge temperature and no excess air. [c]Percent deviations from the base case means are also shown.

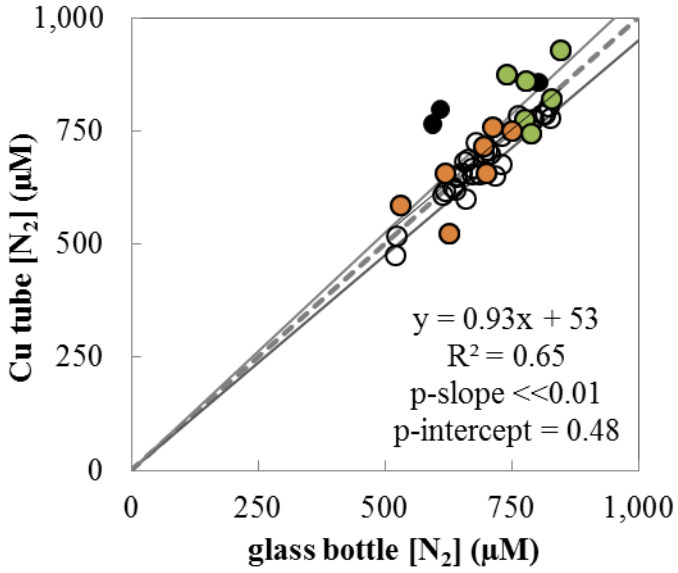

Figure B1. $[N_2]$ for samples stored at ambient temperature in Cu tubes versus samples stored at 4° C in glass bottles. Shaded circles are concentrations in blankets (orange) and well nests (green). Black circles show point samples which may have been biased-high due to storage in copper tubes. The dashed line shows 1:1 relationship and thin solid lines show ± 5% of the 1:1 line.

B4. Potential uncertainty in streambed N_2 from groundwater mixing

The majority of groundwater samples from streambed point sampling were at or near anaerobic conditions (43 of 62 samples <16 µM, Fig. B2), suggesting that mixing of anaerobic and aerobic waters was generally not occurring at the depth where point samples were collected. Two samples where significant mixing of aerobic and anaerobic groundwater may have occurred (300LB, 474L) gave moderate estimates of [N_2-den] and were not considered outliers in our analysis. Other studies have observed denitrification in samples with [O_2] in the range of 50-60 µM (Green et al. [2008] and Böhlke et al. [2002], respectively).

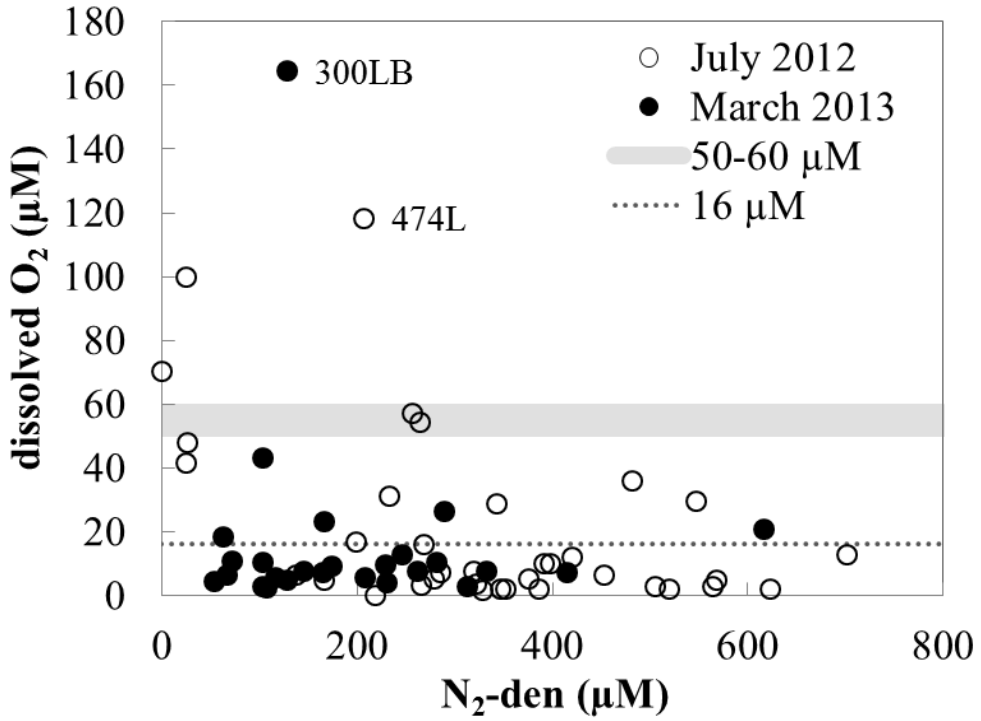

Figure B2. Dissolved oxygen versus [N_2-den] for point samples. Samples where the [N_2-den] estimate was < 0 were set equal to zero.

B5. Streambed point and streambed blanket gas data and [N_2-den]

Table B3. Dissolved gas data and [N_2-den] results.

transect (m)	site (-)	N_2 (μM)	Ar (μM)	Ne ($x10^3$) (μM)	N_2-den (μM)	χ^2 (-)
July 2012 streambed points						
466	RB	803	12.8	5.50	622	1.61
466	R	736	13.6	7.01	319	6.27
466	C	723	13.0	6.66	352	11.9
466	L	749	15.0	7.49	269	0.12
466	LB	776	14.4	8.22	257	7.41
474	RB	830	13.3	6.45	519	4.84
474	R	794	14.3	7.60	329	3.62
474	C	848	14.6	7.28	421	0.74
474	L	834	15.5	9.35	207	3.89
474	LB	n.a.	n.a.	n.a.	n.a.	n.a.
481	RB	732	14.5	6.71	322	0.01
481	R	713	15.2	n.a.	218	n.a.
481	C	727	12.8	5.45	506	1.11
481	L	718	13.5	6.08	391	0.78
481	LB	642	15.3	9.13	24	4.04
491	RB	807	13.6	6.42	482	2.40
491	R	n.a.	n.a.	n.a.	n.a.	n.a.
491	C	732	13.2	5.81	452	0.92
491	L	821	13.1	6.03	565	3.13
491	LB	821	14.4	7.32	386	1.62
499	RB	705	13.5	6.32	347	2.34
499	R	727	11.8	4.49	701	0.34
499	C	798	14.5	7.11	375	0.52
499	L	709	14.7	7.81	199	2.33
499	LB	696	16.4	10.3	26	1.01
508	RB	946	17.0	6.78	568	19.9
508	R	736	13.6	6.24	398	1.10
508	C	732	14.5	7.36	266	1.52
508	L	669	13.9	10.3	25	44.8
508	LB	n.a.	n.a.	n.a.	n.a.	n.a.
516	RB	665	14.2	7.60	166	5.13
516	R	700	14.2	6.15	343	0.21
516	C	707	14.3	7.01	265	0.86

Table B3 Continued

516	L	696	14.6	7.16	233	0.32
516	LB	647	15.5	10.2	0	8.13
524	RB	705	13.2	7.13	278	14.3
524	R	785	13.4	5.68	548	0.04
524	C	750	14.4	7.42	285	2.17
524	L	n.a.	n.a.	n.a.	n.a.	n.a.
524	LB	674	14.0	8.25	138	13.4
July 2012 corrected blankets						
516	RB	610	17.4	6.36	155	35.4
516	R	658	13.4	5.65	349	0.00
516	C	488	13.1	7.74	0	26.7
516	L	238	7.4	6.49	0	424
516	LB	630	20.2	8.01	45	56.8
481	RB	614	12.5	4.94	392	0.04
481	R	699	9.4	3.86	931	24.8
481	C	627	14.4	5.16	345	9.91
481	L	622	13.8	6.94	164	3.85
481	LB	527	13.6	7.83	0	16.5
July 2012 uncorrected blankets						
516	RB	607	17.3	6.43	147	31.9
516	R	660	13.5	5.67	352	0.01
516	C	522	13.0	7.79	0	29.5
516	L	440	11.0	7.37	0	109
516	LB	593	17.0	7.94	23	8.22
481	RB	607	12.6	5.21	344	0.77
481	R	629	11.1	5.75	376	40.7
481	C	620	14.3	5.49	295	4.32
481	L	763	13.8	6.98	354	3.85
481	LB	531	13.5	7.86	0	18.3
March 2013 streambed points						
300	RB	726	17.0	10.2	53	0.05
300	R	866	16.5	9.38	230	0.01
300	C	773	16.4	9.59	128	0.04
300	L	1080	15.5	7.90	616	0.02
300	LB	680	15.5	7.99	129	0.05
715	RB	824	14.9	6.81	415	0.12
715	R	801	15.7	n.a.	262	n.a.
715	C	763	16.0	10.0	104	2.09

Table B3 Continued

715	L	759	16.1	n.a.	166	n.a.

715	LB	795	16.5	10.7	104	1.50
715	RB	828	17.6	9.74	174	3.42
1260	R	818	17.0	9.75	165	0.61
1260	C	813	16.9	10.1	145	0.07
1260	L	812	16.1	n.a.	229	n.a.
1260	LB	803	16.2	8.34	246	0.49
1700	RB	785	17.0	10.1	116	0.08
1700	R	799	17.4	10.7	107	0.32
1700	C	775	17.3	10.9	72	0.02
1700	L	775	15.5	8.47	208	0.72
1700	LB	839	15.2	8.40	290	2.14
1910	RB	841	18.3	11.7	104	1.29
1910	R	802	18.4	11.6	67	1.88
1910	C	863	15.9	8.40	312	0.00
1910	L	688	13.4	6.45	282	6.75
1910	LB	n.a.	n.a.	n.a.	n.a.	n.a.
2530	RB	n.a.	n.a.	n.a.	n.a.	n.a.
2530	R	n.a.	n.a.	n.a.	n.a.	n.a.
2530	C	455	13.2	5.15	62	0.03
2530	L	799	15.3	7.28	332	0.19
2530	LB	n.a.	n.a.	n.a.	n.a.	n.a.
March 2013 corrected blankets						
715	RB	1062	17.9	7.55	615	22.0
715	R	794	16.8	8.63	211	2.60
715	C	849	17.8	11.2	134	0.68
715	L	758	17.7	11.0	51	0.69
715	LB	n.a.	n.a.	n.a.	n.a.	n.a.
March 2013 uncorrected blankets						
715	RB	701	16.8	8.96	80	1.62
715	R	751	16.8	8.79	138	1.80
715	C	714	17.1	9.97	50	0.48
715	L	697	17.2	10.1	26	0.53
715	LB	n.a.	n.a.	n.a.	n.a.	n.a.

"n.a." = no sample analyzed

REFERENCES

Aeschbach-Hertig, W., and D. K. Solomon (2013), Noble gas thermometry in groundwater hydrology, in *The Noble Gases as Geochemical Tracers*, edited by P. Burnard, pp. 81–122, Springer, Heidelberg.

Aeschbach-Hertig, W., F. Peeters, U. Beyerle, and R. Kipfer (2000), Palaeotemperature reconstruction from noble gases in ground water taking into account equilibration with entrapped air, *Nature*, *405*(6790), 1040–1044, doi:10.1038/35016542.

Aeschbach-Hertig, W., H. El-Gamal, M. Wieser, and L. Palcsu (2008), Modeling excess air and degassing in groundwater by equilibrium partitioning with a gas phase, *Water Resour. Res.*, *44*(8), W08449, doi:10.1029/2007WR006454.

Andrews, J. N. (1992), Mechanisms for noble gas dissolution by groundwaters, in *Isotopes of Noble Gases as Tracers in Environmental Studies*, pp. 87–110, International Atomic Energy Agency, Vienna.

Ballentine, C. J., and C. M. Hall (1999), Determining paleotemperature and other variables by using an error-weighted, nonlinear inversion of noble gas concentrations in water, *Geochim. Chosmochim. Acta.*, *63*(16), 2315–2336, doi:10.1016/S0016-7037(99)00131-3.

Benson, B. B., and D. Krause (1976), Empirical laws for dilute aqueous solutions of nonpolar gases, *J. of Chem. Phys.*, *64*(2), 689–709, doi:doi:10.1063/1.432215.

Böhlke, J. K., R. Wanty, M. Tuttle, G. Delin, and M. Landon (2002), Denitrification in the recharge area and discharge area of a transient agricultural nitrate plume in a glacial outwash sand aquifer, Minnesota, *Water Resour. Res.*, *38*(7), 10–1, doi:10.1029/2001WR000663.

Coes, A., T. Spruill, and M. Thomasson (2007), Multiple-method estimation of recharge rates at diverse locations in the North Carolina Coastal Plain, USA, *Hydrogeol. Journal*, *15*(4), 773–788, doi:10.1007/s10040-006-0123-3.

Green, C. T., L. J. Puckett, J. K. Böhlke, B. A. Bekins, S. P. Phillips, L. J. Kauffman, J. M. Denver, and H. M. Johnson (2008), Limited occurrence of denitrification in four shallow aquifers in agricultural areas of the United States, *J. of Environ. Qual.*, *37*(3), 994, doi:10.2134/jeq2006.0419.

Kennedy, C. D., and D. P. Genereux (2007), ^{14}C groundwater age and the importance of chemical fluxes across aquifer boundaries in confined Cretaceous aquifers of North Carolina, USA, *Radiocarbon*, *49*(3), 1181–1203.

Kennedy, C. D., D. P. Genereux, D. R. Corbett, and H. Mitasova (2009), Relationships among groundwater age, denitrification, and the coupled groundwater and nitrogen fluxes through a streambed, *Water Resour. Res., 45*, W09402, doi:10.1029/2008WR007400.

Visser, A., H. P. Broers, and M. F. P. Bierkens (2007), Dating degassed groundwater with $^{3}H/^{3}He$, *Water Resour. Res.*, *43*(10), W10434, doi:10.1029/2006WR005847.

Visser, A., J. D. Schaap, H. P. Broers, and M. F. P. Bierkens (2009), Degassing of $^{3}H/^{3}He$, CFCs and SF_6 by denitrification: Measurements and two-phase transport simulations, *J. Contam. Hydrol.*, *103*(3–4), 206–218, doi:10.1016/j.jconhyd.2008.10.013.

Weiss, R. F. (1970), The solubility of nitrogen, oxygen and argon in water and seawater, *Deep-Sea Res. and Oceanographic Abstracts*, *17*(4), 721–735, doi:10.1016/0011-7471(70)90037-9.

Weiss, R. F. (1971), Solubility of helium and neon in water and seawater, *J. Chem. Eng. Data*, *16*(2), 235–241, doi:10.1021/je60049a019.

Appendix C – Reach Mass-balance Approach for N_2

C1. Dissolved gas concentrations in West Bear Creek

July 2012 stream water [N_2] was just above the "WEA concentration", the concentration expected for water in equilibrium with air (485 μM), except at 200 and 2100 m, and did not increase downstream (Fig. C1). In March 2013, dissolved N_2 in stream water increased in the downstream direction and was significantly above the WEA concentration (617 μM, Fig. C1). Also, changes in [N_2] (and [Ar]) were strongly linked to changes in stream water temperature.

Based on stream water temperatures measured at the 0 m, 1000 m, and 2700 m stream locations, the stream followed a diel temperature cycle in March 2013. Based on observed stream water [Ar] concentrations and [Ar] WEA concentrations computed from stream temperature (average of the three locations mentioned above), there was a 4.7 hour lag between stream temperature and stream water [Ar] (Fig. C2). For example, as the stream warmed after sunrise, [Ar] dropped but with a lag of about 4.7 hours, such that measured [Ar] remained above the WEA [Ar]. A similar time lag was observed for [N_2], but with a noticeable upward shift of about 27 μM in concentration, i.e., for [N_2] the same sinusoidal model that fit the WEA curve also fell through most of the observed data points when shifted forward in time by 4.0 hours and upward by 27 μM. The upward shift in stream [N_2] was likely due to excess [N_2] in groundwater inflow to the stream and/or production of [N_2] in the stream channel.

The fact that observed [N_2] and [Ar] decreased at a similar rate compared to WEA suggests that change in temperature affected observed stream water concentrations. To account for changing stream temperature, a simple correction was applied to the March 2013 stream concentration data before applying equations C1 and C2 below (Table C1). When calculating groundwater concentrations or in-channel production, the upstream concentration (C_{up}, in equation C1 or C2) was corrected for the change in WEA concentration during the time lapse between collection of upstream and downstream samples, and the observed

downstream concentration was used for C_{down}. This way, C_{up} and C_{down} refer to the same moment in time.

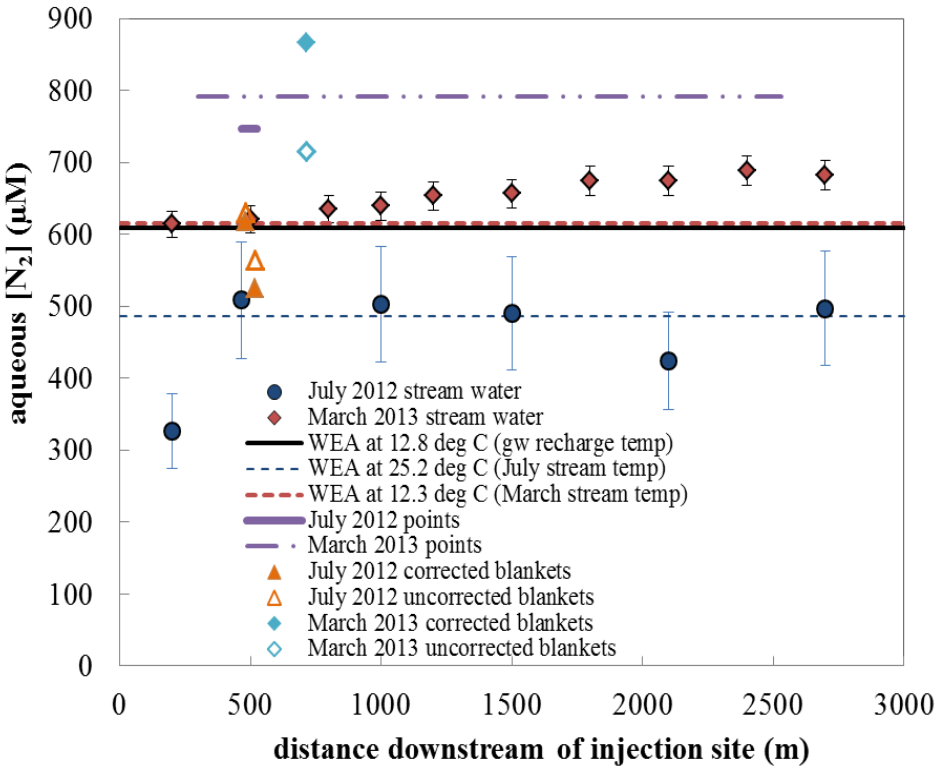

Figure C1. Aqueous N_2 concentrations, [N_2], in West Bear Creek during July 2012 and March 2013. Horizontal dashed lines indicate the expected [N_2] in "water in equilibrium with air" (WEA), at 12.3°C, 12.8°C, and 25.2°C (mean stream temperature in March 2013, estimated groundwater recharge temperature, and mean stream temperature in July 2012, respectively). Mean groundwater [N_2] from July 2012 point samples is shown by the short (466-524m) purple bar and mean [N_2] from March 2013 point samples is shown by the purple line broken by two dots (300-2350m). Mean [N_2] from July 2012 and March 2013 blanket transects are shown; open symbols are for raw blankets and shaded symbols are for corrected blankets.

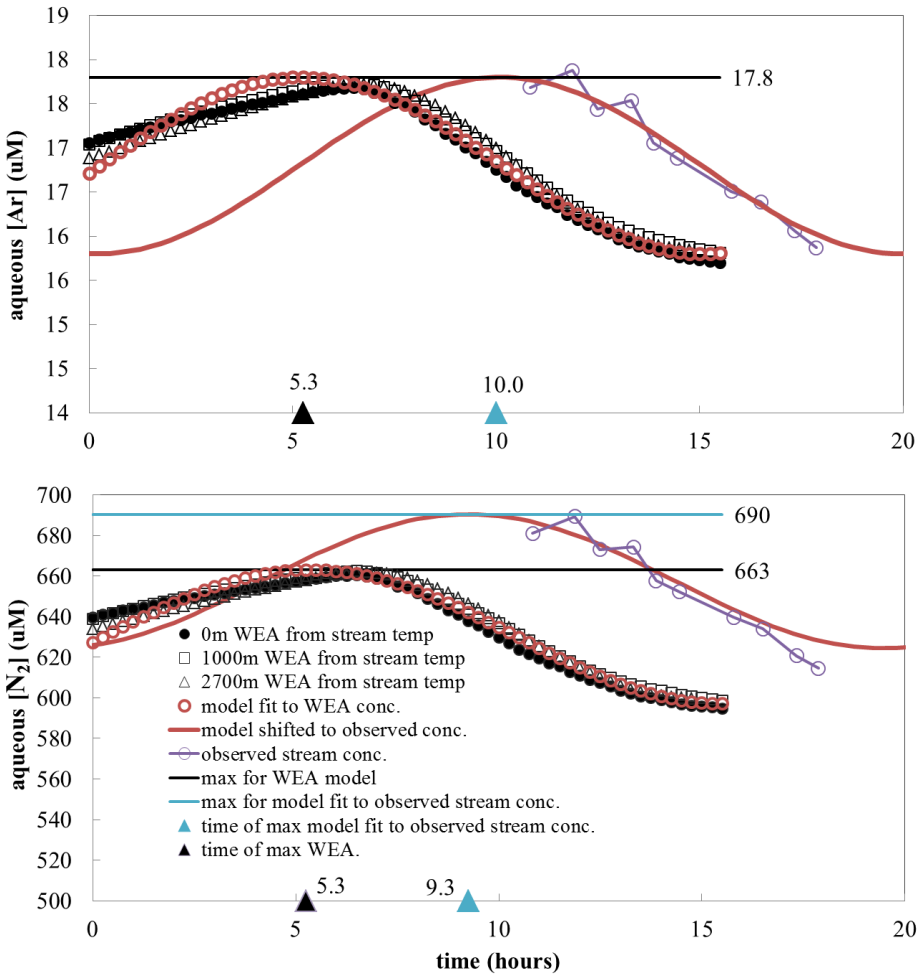

Figure C2. Stream water concentrations predicted from stream water temperature (WEA) and observed concentrations in the stream in March 2013. Time = 0 hours was at midnight on 15 March 2013. Sinusoidal curves were fit to the computed WEA concentrations (red circles, wavelength = 20 hours for Ar, 20.5 hours for N_2). Horizontal black lines show the maximum concentrations from curves fit to WEA values. For [Ar], this maximum agrees with that suggested by a curve fit to the measured stream water [Ar], but for [N_2] the stream water data point to a higher maximum (horizontal blue line) than suggested by WEA values. Black triangles indicate the time of maximum concentration for WEA curves and blue triangles the time of maximum concentration suggested by the stream water data.

Table C1. Observed and corrected $[N_2]$ and [Ar] for March 2013.

Reach	time between samples (hr)	observed		corrected	
		upstream $[N_2]$ (µM)	upstream [Ar] (µM)	upstream $[N_2]$ (µM)	upstream [Ar] (µM)
200-1000m	2.08	615	15.9	630	16.4
1000-1800m	2.47	640	16.5	664	17.3
1800-2700m	2.48	674	17.5	693	18.0
200-2700m	7.03	615	15.9	673	17.6

C2. Models for groundwater concentrations and in-stream production of N_2

Elevated $[N_2]$ in stream water (i.e., measured $[N_2]$ > WEA $[N_2]$) could result from either input of excess N_2 from groundwater or production of N_2 in the channel (or both). To assess this, we calculated both (1) apparent groundwater $[N_2]$ from a reach mass-balance approach, to see if the groundwater $[N_2]$ needed to explain elevated stream $[N_2]$ were feasible (i.e., similar in magnitude to the $[N_2]$ measured at streambed point and blankets), and (2) apparent in-stream N_2 production (assuming groundwater $[N_2]$ was represented by streambed point measurements), to see if reasonable production rates alone could explain elevated stream $[N_2]$.

Mean concentrations of N_2 and Ar in the groundwater discharge into a stream reach were calculated as:

$$C_{RMB} = \frac{C_{down}Q_{down} - C_{up}Q_{up} + zkA\left(0.5\left(C_{up} + C_{down}\right) - C_{WEA}\right)}{Q_{down} - Q_{up}} \quad \text{(C1)}$$

where

- C_{RMB} = dissolved gas concentration in groundwater (µM)
- C_{WEA} = dissolved gas concentration for water in equilibrium with the atmosphere at the stream temperature (µM)

- C_{up} = dissolved gas concentration in stream water at the upstream station (μM)
- C_{down} = dissolved gas concentration in stream water at the downstream station (μM)
- Q_{up} = stream discharge at the upstream station (m^3/sec)
- Q_{down} = stream discharge at the downstream station (m^3/sec)
- k = gas transfer velocity (m/day)
- A = the surface area of the stream water (m^2)
- $z = 0.0116 \times 10^{-5}$ is for unit conversion (day/sec).

In-channel production of N_2 within a stream reach would elevate C_{down} which in turn would elevate C_{RMB}, raising the estimate of excess N_2 in the groundwater input to the reach and thus the estimate of [N_2-den].

In-channel production of N_2 can be modeled for a steady-state stream reach as:

$$P = \left[\frac{d}{\tau}\left(C_{down} - C_{up}\right) + \frac{q}{w}\left(C_{ave} - C_{gw}\right) + Kd\left(C_{ave} - C_{WEA}\right) \right] * 3600 \qquad \text{(C2)}$$

where

- P = rate of N_2 production (mmol m^{-2} hr^{-1})
- τ = transit time of stream water through the reach (sec)
- C_{ave} = mean of C_{up} and C_{down} (mmol m^{-3})
- C_{gw} = concentration from streambed point samples in the stream reach (mmol m^{-3})
- q = rate of groundwater inflow to the reach (m^3 sec^{-1} m^{-1})
- w = width of stream (m)
- d = depth of stream (m)

- K = first-order gas transfer rate constant (sec^{-1})
- 3600 = conversion factor from seconds to hours

Equation C2 was derived from a 1-D advection-dispersion equation (e.g., *Genereux and Hemond* [1990]) but with an additional term for production of N_2. Starting with equation A4 of *Genereux and Hemond* [1990]:

$$U\frac{dC_{ave}}{dx}+\frac{q}{A_y}C_{ave}=r \tag{C3}$$

where

- U = stream velocity [L/T]
- x = distance in downstream direction [L]
- A_y = cross-sectional area, perpendicular to x [L^2]
- r = source/sink term [M/L^3T]

With r accounting for groundwater input, gas exchange, and production:

$$r=\frac{qC_{gw}}{A_y}-K\left(C_{stream}-C_{WEA}\right)+p \tag{C4}$$

where p = in-stream production rate in units of [M/L^3T] and C_{stream} is the concentration in stream water. Then the steady-state governing equation is:

$$U\frac{dC_{stream}}{dx}+\frac{q}{A_y}C_{stream}=\frac{qC_{gw}}{A_y}-K\left(C_{stream}-C_{WEA}\right)+p \tag{C5}$$

Solving for the production term, p gives:

$$p=U\frac{dC_{stream}}{dx}+\frac{q}{A_y}\left(C_{stream}-C_{gw}\right)+K\left(C_{stream}-C_{WEA}\right) \tag{C6}$$

which can be approximated in finite difference form as:

$$p = U\frac{C_{down} - C_{up}}{\Delta x} + \frac{q}{A_y}\left(C_{ave} - C_{gw}\right) + K_{ave}\left(C_{ave} - C_{WEA}\right) \quad \text{(C7)}$$

Where all variables except C_{down}, C_{up}, and Δx are reach-averaged values. Substituting the transit time through the reach ($\tau = \Delta x/U$) into equation C7 gives the equation for N_2 production in the stream channel in units of [M/(L^3T)], which if divided by the average stream depth (d) gives equation C2 in units of [M/(L^2T)].

C3. Groundwater [Ar], [N_2], and [N_2-den] by reach mass-balance

The overall approach to determining [N_2-den] using reach mass-balance data was as follows:

1. Calculate groundwater [N_2] and [Ar] using equation C1.

2. Use ΔAr (rather than ΔNe, which was used for point and blanket samples, Appendix B) to choose the appropriate Closed-Equilibrium (CE) model formulation and recharge parameters.

3. Use the calculated groundwater [Ar] to solve the CE model for the value of the appropriate model parameter (either A or B) (Appendix B).

4. Use the calculated groundwater [N_2] from the reach mass-balance in the CE model (with A or B calculated from step 3) to estimate [N_2-den] in groundwater. In other words, once the A or B parameter is determined in step 3, groundwater [N_2] from reach mass-balance calculations (equation C1) is used in the same manner as measured [N_2] from points or blankets to estimate [N_2-den] (Appendix B).

Calculation of [N_2] and [Ar] by reach mass-balance requires estimates of gas transfer velocities (k, m/day) for N_2 and Ar. Estimates for k were determined for the injected tracers, Xe and Kr, using the RADIN13 spreadsheet model (obtained from Peter Cook, CSIRO Land and Water, Australia). We assumed that the ratio of k_{Xe}/k_{Kr} was equal to $(D_{Xe}/D_{Kr})^n$, where D is the aqueous diffusion coefficient for the gas at the mean stream temperature [*Jähne et al.*

1987b]. Kr and Xe data from July 2012 suggested *n* was very close to 1.0 (0.96); a value of 1 was used for all calculations. The relationship $k_i/k_{Kr} = (D_i/D_{Kr})^n$ (with n= 1) was used to determine *k* values for N_2 and Ar. All diffusion coefficients were calculated from *Jähne et al.* [1987a], with the exception of N_2 and Ar [*Wise and Houghton* 1966] (Table C2). Width and depth models for July 2012 and March 2013 are detailed in *Solomon et al.* [2015].

Table C2. Diffusion coefficients (*D*) and gas transfer velocities (*k*) at July 2012 and March 2013 mean stream temperature (25.7° C and 12.3° C, respectively).

Gas (-)	*D* at 25.7° C (cm^2/sec)	*k* at 25.7° C (m/d)	*D* at 12.3° C (cm^2/sec)	*k* at 12.3° C (m/d)
Kr	$1.9x10^{-5}$	1.46	$1.3x10^{-5}$	1.05
Xe	$1.5x10^{-5}$	1.17	$1.0x10^{-5}$	0.82
N_2	$2.9x10^{-5}$	2.27	$2.0x10^{-5}$	1.65
Ar	$2.6x10^{-5}$	2.03	$1.8x10^{-5}$	1.45

The groundwater concentrations from reach mass-balance (equation C1) generally were lower than flow-weighted concentrations from point measurements for the July 2012 field campaign (Table C3). In one July 2012 reach (466-2700m) there was an apparent denitrification signal, [N_2-den] = 529 μM. However, the signal was not so much due to a high reach mass-balance estimate of groundwater [N_2], (624 μM, which is just above WEA at recharge temperature, 610 μM, but lower than the minimum [N_2] in points 642 μM). Rather, it was because the groundwater [Ar] used to interpret the groundwater [N_2] signal was low (Table C3). Low [Ar] gives the impression of a deeply degassed sample and produces a large estimate of [N_2-den] from the small excess [N_2]. In addition, uncertainty in [Ar] and [N_2] from reach mass-balance was 21% and 91%, respectively which results in unacceptably high uncertainty in the [N_2-den] estimate. The largest reach mass-balance [N_2-den] estimate for March 2013 (551 μM, 1800-2700m) was also associated with relatively low estimates of [N_2] and [Ar] from reach mass-balance, and also had high uncertainty due to large uncertainties in [N_2] and [Ar] (>1000%). In short, we believe the apparent [N_2-den] estimates from reach

mass-balance calculations are as likely to be a product of random variation in stream water concentrations as they are likely to be a product of groundwater discharge.

Table C3. Flow-weighted groundwater [N_2], [Ar], and [N_2-den] from points and reach mass-balance (RMB) approaches.

approach/reach	[N_2] (μM)	[Ar] (μM)	[N_2-den] (μM)
July 2012 points	728	14.2	320
July 2012 blankets[a]	596	14.0	262
July RMB 466-1000m	587	16.5	-18
July RMB 466-2700m	624	11.8	529
March 2013 points	820	16.6	188
March 2013 blankets[a]	811	17.2	182
March RMB 200-1000m	908	19.3	36
March RMB 1000-1800m	821	22.5	-256
March RMB 1800-2700m	751	13.2	551
March RMB 200-2700m	1013	22.7	-145
WEA at 12.8° C	610	16.1	n.a.

"n.a." = not applicable
"WEA" = water in equilibrium with air
[a]only corrected blanket concentrations are shown

C4. N_2 production in the stream channel and total N_2 budget

Stream water [N_2] and flow-weighted [N_2] from points were used in equation C2 to estimate N_2 production in the stream channel (*P*, Table C4). The corrected stream [N_2] (Table C1) were used for March 2013 calculations. Because observed groundwater concentrations are explicitly used in equation C2, the estimates of *P* are a proxy for how well groundwater concentrations predict stream concentrations (i.e., $P = 0$ suggests that observed groundwater from points fully explains stream water concentrations). In the cases where $P > 0$, the estimates of *P* are of reasonable magnitude compared to literature values compiled by Baulch et al. (2010) (range of 0 – 8 mmol N_2 m^{-2} hr^{-1} their Appendix 1), although most of those estimates were for rivers much larger than West Bear Creek. However, given that point measurements showed elevated [N_2] (due to denitrification in the surficial aquifer) and if

additional denitrification was occurring in the stream channel (we believe it was), then stream reaches with a strong groundwater signal should show both an [N_2-den] signal (calculated as explained in Section C.3) and a stream channel production signal (P, as calculated from equation C2). Both of the reaches with large [N_2-den] estimates by reach mass-balance (466-2700m in July, 1800-2700m in March) have no in-stream N_2 production ($P < 0$). The result of $P < 0$ is consistent with the relatively low groundwater [N_2] calculated from equation C1, and highlights how the large [N_2-den] estimates for these reaches (Table C3) are largely a product of low [Ar] rather than actual detection of excess N_2 in groundwater.

Table C4. Flow-weighted groundwater [N_2] from points, and calculated production (P) from equation C2.

reach	point [N_2] (μM)	P (mmol m^{-2} hr^{-1})
July RMB 466-1000m	728	-4.0
July RMB 466-2700m	728	-2.3
March RMB 200-1000m	850	2.4
March RMB 1000-1800m	812	3.0
March RMB 1800-2700m	779	-2.6
March RMB 200-2700m	820	5.2

Ultimately, the fact that excess N_2 in stream water (either from groundwater or in-stream production) is a small fraction of the total N_2 flux through a stream reach is likely the primary driver for the limited sensitivity of the reach mass-balance approach and the variability in [N_2-den] results (Table C5). For example, the average stream water flux of N_2 was about 90 mmol m^{-2} hr^{-1} in the 466-1000 m reach in July 2012. In contrast, the flux of groundwater [N_2] from point measurements was only about 11 mmol m^{-2} hr^{-1} (12% of stream water N_2 flux), and the flux of [N_2-den] from points was only about 5 mmol m^{-2} hr^{-1} (6%). Based on this analysis of the relative magnitudes of stream water and groundwater-derived N_2 fluxes, it appears that only the 466-2700 m reach in July 2012 has an N_2-den flux (based

on points) that is significant enough relative to the stream water N_2 flux (~20%) to feasibly be detected using the reach mass balance approach.

Table C5. Fluxes of N_2 from stream water, N_2-den, *P*, and point measurements.

reach	US N_2 flux[a]	DS N_2 flux[a]	RMB N_2-den flux[b]	*P*	point N_2-den flux[c]	point N_2 flux[d]
July RMB 466-1000m	99.3	81.4	<0	<0	4.9	11.1
July RMB 466-2700m	30.8	19.5	6.3	<0	4.9	11.1
March RMB 200-1000m	452	418	0.8	2.4	4.0	13.8
March RMB 1000-1800m	198	177	<0	3.0	3.6	20.6
March RMB 1800-2700m	202	203	2.5	<0	2.6	9.6
March RMB 200-2700m	202	164	<0	5.2	3.5	15.4

all units are mmol m^{-2} hr^{-1}

[a]stream discharge times concentration then divided by total reach area. "US" = upstream, "DS" = downstream.

[b]$[N_2$-den] estimate multiplied by groundwater discharge ($Q_{down} - Q_{up}$) then divided by total reach area

[c]mean of $[N_2$-den]*v for points in the given reach

[d]mean of $[N_2]$*v for points in the given reach

The variability in $[N_2$-den] and *P* results, the large estimated uncertainties in groundwater dissolved gas concentrations from reach mass-balance, and the relatively small fluxes for the parameters of interest (N_2 and N_2-den from groundwater, and *P*), suggest that these methods would require further development to be highly useful in streams similar to West Bear Creek.

Based on the results from our two field campaigns, we conclude that assessment of $[N_2$-den] by reach mass-balance is likely to be successful only in stream reaches where the groundwater flux of N_2 into the reach is a large proportion of stream water N_2 fluxes, $[N_2]$ in groundwater and surface water is measured with high precision, and the flux of N_2-den is large relative to in-channel production and the stream gas exchange flux.

REFERENCES

Baulch, H. M., J. J. Venkiteswaran, P. J. Dillon, and R. Maranger (2010), Revisiting the application of open-channel estimates of denitrification, *Limnol. Oceanogr. Meth.*, *8*, 202–215, doi:10.4319/lom.2010.8.202.

Genereux, D. P., and H. F. Hemond (1990), Naturally occurring Radon 222 as a tracer for streamflow generation: Steady state methodology and field example, *Water Resour. Res.*, *26*(12), 3065–3075, doi:10.1029/WR026i012p03065.

Jähne, B., G. Heinz, and W. Dietrich (1987a), Measurement of the diffusion coefficients of sparingly soluble gases in water, *J.of Geophy. Res.: Oceans*, *92*(C10), 10767–10776, doi:10.1029/JC092iC10p10767.

Jähne, B., K. O. Münnich, R. Bösinger, A. Dutzi, W. Huber, and P. Libner (1987b), On the parameters influencing air-water gas exchange, *J. Geophys. Res.*, *92*(C2), 1937–1949, doi:10.1029/JC092iC02p01937.

Solomon, D. K., T. E. Gilmore, J. E. Solder, B. A. Kimball, and D. P. Genereux (2015), Evaluating a groundwater flow system using SF6 in streamflow, *in prep*.

Wise, D.L., and G. Houghton. (1966), The diffusion coefficients of ten slightly soluble gases in water at 10–60°C, *Chemical Engineering Science 21* (11): 999–1010. doi:10.1016/0009-2509(66)85096-0.

Appendix D – Dissolved Gas Modeling and Apparent Age Results

D1. Sensitivity of mean apparent age and mean transit time to recharge parameters and R_{terr}

As described in the Chapter 2 and in *Gilmore et al.* [2015, supplemental material], recharge parameters (mean recharge temperature (RT) and excess air (A)) were determined from noble gas concentrations (Ne, Ar, Xe, Kr) in groundwater collected from six wells in the Bear Creek watershed (Fig. 2.1, main text). The recharge parameters were then used to constrain models of dissolved gas concentrations in groundwater collected in the streambed. Estimates of uncertainty in mean apparent age and mean transit time (*MTT*) from the SF_6 and $^3H/^3He$ age-dating methods were determined using a sensitivity analysis where limiting values of recharge parameters were used in the Closed Equilibrium (CE) model (Table D1). These limiting values represent the "high gas" and "low gas" cases described in *Gilmore et al.* [2015, supplemental material] and correspond to the likely maximum and minimum amounts of atmospherically derived gases in the groundwater at recharge. The limiting cases were compared to the "base case" (i.e., values reported in the main text) to determine the case that caused maximum deviation. The maximum deviation (%) was considered an estimate of uncertainty in mean apparent age and *MTT*.

Terrigenic He was observed in some samples, as indicated by low (< 1) R_{meas}/R_{atm} ratios (where R is the ratio of $[^4He]/[^3He]$ for the given subscript; "meas" = measured, "atm" = atmospheric). Large amounts of terrigenic ^4He ($^4He_{terr}$) have been observed in the confined aquifers (e.g., Black Creek aquifer) underlying the surficial aquifer near West Bear Creek [*Kennedy and Genereux*, 2007], and may diffuse through the Black Creek confining unit and into older water in the surficial aquifer. Small amounts of $^3He_{terr}$ are produced at the same time as $^4He_{terr}$ and must be accounted for in the calculation of $^3He_{trit}$ (Equation 4, main text). The theoretical ratio for $[^4He_{terr}/^3He_{terr}]$ ($R_{terr} = 2 \times 10^{-8}$) was used as part of the calculation. Larger values for R_{terr} have been estimated (e.g., 4×10^{-7} [*Robertson et al.*, 2013]), so the sensitivity of apparent age from $^3H/^3He$ was also tested using a higher R_{terr} (Table D1).

Table D1. Values for recharge variables and R_{terr} for sensitivity analyses.

Recharge case	**RT**	**A**	**R_{terr}**
	(°C)	(mL/L)	(-)
base case	12.8	23	2×10^{-8}
high gas	9.3	41	2×10^{-8}
low gas	16.2	0	2×10^{-8}
high gas, high R_{terr}	9.3	41	1×10^{-7}
low gas, high R_{terr}	16.2	0	1×10^{-7}

An estimate of "high" R_{terr} was derived using data from the sample with the lowest R_{meas}/R_{atm} (point sampled at 1910R in March 2013; $R_{meas}/R_{atm} = 0.64$), given the likelihood that He_{terr} was a significant fraction of the He budget for that sample. When the value of $R_{terr} = 1 \times 10^{-7}$ was used to model apparent age for the sample from 1910R, $^3He_{terr}$ was roughly equal to $^3He_{trit}$ (i.e., a significant proportion of the ^3He budget), so 1×10^{-7} was assumed as a reasonable estimate for "high" R_{terr}.

Results of the sensitivity analysis show that mean apparent age and *MTT* from SF_6 were relatively insensitive to the recharge variables (1-7% deviation from the base case values reported in the main text). Mean apparent age and *MTT* from $^3H/^3He$ were more sensitive to recharge variables than SF_6, but the choice of R_{terr} resulted in the maximum deviation for mean apparent age in both July (15%) and March (14%), and likewise for *MTT* (23% and 8%, Table D2). We conclude that a reasonable estimate of uncertainty in mean apparent age and *MTT* from $^3H/^3He$ is likely about 15-20% for this study. This uncertainty range includes likely uncertainty from the variability in recharge parameters and accounts for a potential bias from the choice of R_{terr}.

Table D2. Uncertainty estimates for mean apparent age and mean transit time (*MTT*).

Case I.D.	Recharge case	$^3H/^3He$ mean age[a]	$^3H/^3He$ *MTT*[a]	SF_6 mean age[a]	SF_6 *MTT*[a]
			July 2012		
A	base case	26.6	29.2	23.3	26.4
B	high gas[b]	29.2	32.4	23.7	26.5
C	low gas[c]	22.5	25.3	22.7	26.1
D	high gas, high R_{terr}[b]	30.3	29.1	n.a.	n.a.
E	low gas, high R_{terr}[c]	22.4	22.4	n.a.	n.a.
	max deviation[d]	15% (E)	23% (E)	1%	1%
			March 2013		
F	base case	29.0	31.0	22.9	31.7
G	high gas[b]	31.5	32.6	24.2	32.3
H	low gas[c]	27.3	29.5	21.3	31.2
I	high gas, high Rterr[b]	30.6	32.0	n.a.	n.a.
J	low gas, high Rterr[c]	24.9	28.6	n.a.	n.a.
	max deviation[d]	14% (J)	8% (J)	7%	2%

[a]All units are in years, unless specified as percentage.
[b]The "high gas" case assumed low recharge temperature and high excess air, "high R_{terr}" assumed $R_{terr} = 1.0 \times 10^{-7}$.
[c]The "low gas" case assumed high recharge temperature and no excess air, "high R_{terr}" assumed $R_{terr} = 1.0 \times 10^{-7}$.
[d]Maximum percent deviations from the base case means are shown, with the corresponding Case I.D. in parentheses.

D2. Assessment of apparent age sensitivity to the time of degassing

The location of degassing in the groundwater system is a potentially important consideration for estimating apparent age from $^3H/^3He$. As noted in Section 5 of the main text, degassing of samples was believed to occur near the streambed, at the end of the groundwater flow path. This was based on (1) the absence of degassing in well samples (based on noble gas concentrations), and (2) the likelihood that groundwater became supersaturated in the surficial aquifer and then degassed as groundwater flowed upward toward lower hydrostatic head conditions near the stream. In the case of degassing near the stream, $^3He_{trit}$ was produced prior to degassing, and when degassing occurred, both atmospheric and tritiogenic ^3He were removed from the groundwater (thus the need for the degassing correction factor, 1+BH, in Equation 4 of the main text). If degassing occurred at

recharge, the degassing event would remove only atmospheric ^{3}He (and no degassing correction factor is needed) because $^{3}He_{trit}$ would not be present in a groundwater sample with age of zero [e.g., *Visser et al.,* 2007].

The apparent age from $^{3}H/^{3}He$ for each point sample was calculated with and without the degassing correction factor to evaluate the sensitivity of apparent age to assumptions about the location of degassing. In July 2012, when about 85% of the groundwater samples were degassed, mean apparent age assuming degassing at recharge was 20.3 years, compared to a mean of 26.6 years if degassing occurred near the stream, while in March 2013 (43% of samples were degassed), mean ages were 28.2 years and 29.0 years, respectively. Given the high probability that degassing occurred near the stream, and that the maximum uncertainty in mean apparent age associated with the location of degassing is within the range of uncertainty linked with variability in recharge parameters (Section 1), it seems unlikely that the assumption of degassing near the stream contributed significantly to uncertainty in mean apparent age or *MTT*.

D3.Monte Carlo assessment of uncertainty in groundwater apparent age

A Monte Carlo approach was used to estimate uncertainty in groundwater apparent age estimates. Point samples from the 516 m transect were chosen as a representative sample because $^{3}H/^{3}He$ ages ranged from about 2.0 to 41.1 years for that transect. For each sample, 200 simulations were run where dissolved gas concentrations, ^{3}H concentration, and recharge temperature were randomly varied (Table D3) and a new groundwater age was determined for each simulation. Random variation of the input parameters was performed using the Excel equation norm.inv(rand(),μ, σ), where μ was set to the measured value, σ was set to analytical uncertainty (or model uncertainty in the case of recharge temperature). The norm.inv function gives the value from a normal distribution with mean = μ and standard deviation = σ and probability = rand(). The rand() function randomly generated a probability between 0 and 1 so that the norm.inv function randomly selected values from a normal distribution with mean μ and standard deviation σ. After 200 simulations were completed, the coefficient of variation (CoV) of the 200 simulated apparent ages was taken as an

approximation of uncertainty. While the analysis did not evaluate sensitivity to R_{terr} (see Section 1), the results do illustrate how apparent age from $^3H/^3He$ became much more uncertain (on a percentage basis) when apparent age was small (Table D4, Fig. D1). Large uncertainty in young apparent age is expected because [$^3He_{trit}$] is determined by subtraction (Equation 4, main text) and $^3He_{trit}$ is a small percentage of the total ^3He budget for very young groundwater. Uncertainty in modeled SF_6 age did not show a strong trend.

Table D3. Uncertainties used in Monte Carlo simulations.

parameter	uncertainty
[Ar], [Ne], [N_2]	3%
[He]	2%
R/R_a	1%
[^3H]	4-7%
recharge temperature	13%
number of simulations	200

Table D4. Uncertainty in apparent ages from Monte Carlo simulations.

sample	$^3H/^3He$ analyt.[a] age (years)	$^3H/^3He$ MC[b] age (years)	$^3H/^3He$ MC[b] CoV (%)	$^3H/^3He$ MC[b] age x CoV (years)	SF_6 analyt.[a] age (years)	SF_6 MC[b] age (years)	SF_6 MC[b] CoV (%)	SF_6 MC[b] age x CoV (years)
516RB	8.81	10.8	39%	4.2	18.0	17.3	6%	1.0
516R	30.2	31.9	12%	3.8	20.5	20.6	3%	0.6
516C	41.1	44.3	5%	2.2	29.5	29.3	2%	0.6
516L	30.3	33.2	6%	2.0	25.5	25.3	2%	0.5
516LB	2.02	1.24	486%	6.0	10.0	9.91	13%	1.3

[a]"analyt." = "analytical", or best estimate from CE or UA model
[b]"MC" = Monte Carlo, MC ages are the mean of 200 simulations

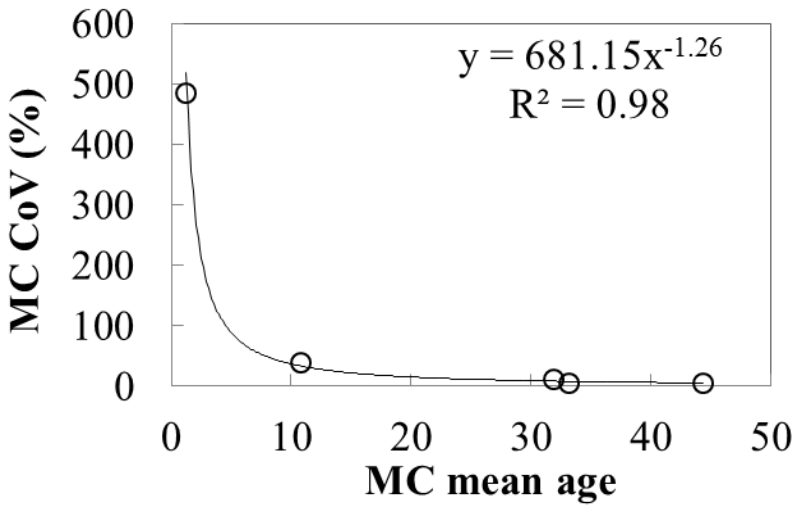

Figure D1. Trend in uncertainty as $^3H/^3He$ age increases.

Table D5. Raw data and apparent age estimates for July 2012 and March 2013 field campaigns.

trans-ect	site	[Ar] x 10^{-4}	[Ne] x 10^{-7}	[^{4}He] x 10^{-8}	R/Ra	[^{3}H]	ΔNe	[^{3}He$_{trit}$]	χ^2	^{3}H/^{3}He age	SF_6 age[a]	CFC-11 age[a]	CFC-12 age[a]	CFC-113 age[a]
(m)	(-)	(gas conc. in ccSTP)			(-)	(TU)	(%)	(TU)	(-)	(yrs)	(yrs)	(yrs)	(yrs)	(yrs)
						July 2012 points								
524	RB	2.96	1.60	3.69	1.08	2.3	-18	5.2	14.3	21.2	n.a.	n.a.	n.a.	n.a.
524	R	3.01	1.27	2.89	1.25	2.3	-35	9.3	<0.1	28.8	n.a.	n.a.	n.a.	n.a.
524	C	3.23	1.66	4.33	1.38	2.7	-15	19.1	2.2	37.6	n.a.	n.a.	n.a.	n.a.
524	L[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
524	LB	3.15	1.85	4.36	1.23	3.4	-6	10.2	13.4	24.7	n.a.	n.a.	n.a.	n.a.
516	RB[c]	3.18	1.70	4.02	1.02	4.2	-13	2.8	5.1	9.2	18.0	39.5	C	30.0
516	R[c]	3.18	1.38	3.31	1.12	1.4	-30	6.5	0.2	31.5	21.0	39.5	C	29.5
516	C	3.21	1.57	3.87	1.43	1.7	-20	18.1	0.9	43.9	29.5	48.0	36.5	37.0
516	L	3.28	1.60	3.62	1.68	4.1	-18	21.7	0.3	32.9	25.5	37.0	24.5	29.0
516	LB	3.46	2.29	5.16	1.03	3.9	17	0.5	8.1	2.0	10.0	30.5	23.0	26.0
508	RB	3.80	1.52	3.43	1.15	4.0	-22	3.2	19.9	10.6	n.a.	n.a.	n.a.	n.a.
508	R	3.05	1.40	3.44	1.21	3.0	-29	11.1	1.1	28.0	n.a.	n.a.	n.a.	n.a.
508	C	3.24	1.65	4.10	1.36	2.7	-16	16.2	1.5	35.0	n.a.	n.a.	n.a.	n.a.
508	L	3.12	2.32	5.43	1.04	4.0	18	3.4	44.8	11.0	n.a.	n.a.	n.a.	n.a.
508	LB[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
499	RB	3.02	1.42	3.17	1.29	2.9	-28	10.0	2.3	26.8	n.a.	n.a.	n.a.	n.a.
499	R	2.65	1.01	2.22	1.29	4.9	-49	10.1	0.3	20.0	n.a.	n.a.	n.a.	n.a.
499	C	3.25	1.59	4.21	1.29	2.1	-19	16.5	0.5	39.2	n.a.	n.a.	n.a.	n.a.

Table D5 Continued

499	L	3.29	1.75	4.15	1.35	4.1	-11	13.4	2.3	25.9	n.a.	n.a.	n.a.	n.a.
499	LB[d]	3.67	2.30	4.61	1.10	3.7	18	-1.9	1.0	-13.1	n.a.	n.a.	n.a.	n.a.
491	RB	3.04	1.44	3.35	1.24	4.3	-26	9.9	2.4	21.4	n.a.	n.a.	n.a.	n.a.
491	R[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
491	C	2.96	1.30	3.19	1.31	2.7	-33	14.8	0.9	33.6	n.a.	n.a.	n.a.	n.a.
491	L	2.94	1.35	3.27	1.36	3.5	-31	16.2	3.1	31.2	n.a.	n.a.	n.a.	n.a.
491	LB	3.23	1.64	3.72	1.30	3.8	-16	10.0	1.6	23.3	n.a.	n.a.	n.a.	n.a.
481	RB[c]	3.26	1.51	3.24	1.47	1.9	-23	13.1	<0.1	37.4	27.5	46.0	C	34.5
481	R[b]	3.40	n.a.	n.a.	n.a.	2.91	n.a.	n.a.	n.a.	n.a.	34.5	57.0	64.5	59.5
481	C	2.87	1.22	3.18	1.29	1.3	-38	17.4	1.1	47.8	27.5	48.0	23.0	37.0
481	L[e]	3.03	1.36	3.08	1.48	3.0	-30	16.4	0.8	33.4	24.0	43.0	**18** (0)	33.0
481	LB[c]	3.44	2.05	4.67	1.19	2.3	5	5.1	4.0	20.9	15.5	29.5	C	25.0
474	RB	2.98	1.45	3.22	1.39	3.5	-26	13.5	4.8	28.5	n.a.	n.a.	n.a.	n.a.
474	R	3.22	1.70	4.11	1.59	2.9	-13	22.9	3.6	39.4	n.a.	n.a.	n.a.	n.a.
474	C	3.27	1.63	3.98	1.36	2.9	-17	15.0	0.7	32.5	n.a.	n.a.	n.a.	n.a.
474	L	3.46	2.10	5.12	1.22	3.8	7	8.1	3.9	20.5	n.a.	n.a.	n.a.	n.a.
474	LB[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
466	RB	2.86	1.23	2.69	1.26	3.7	-37	8.7	1.6	21.7	n.a.	n.a.	n.a.	n.a.
466	R	3.06	1.57	3.51	1.28	4.0	-20	9.7	6.3	22.0	n.a.	n.a.	n.a.	n.a.
466	C	2.91	1.49	3.38	1.38	4.0	-24	14.6	11.9	27.6	n.a.	n.a.	n.a.	n.a.
466	L	3.36	1.68	4.26	1.42	1.7	-14	18.7	0.1	44.2	n.a.	n.a.	n.a.	n.a.
466	LB	3.23	1.84	4.20	1.17	3.4	-6	6.2	7.4	18.6	n.a.	n.a.	n.a.	n.a.

Table D5 Continued

July 2012 corrected blankets														
516	RB[c]	3.92	1.43	3.29	1.12	1.8	-27	3.0	35.4	17.5	20.0	38.0	C	31.5
516	R[c]	3.02	1.27	3.40	1.09	3.1	-35	10.2	<0.1	26.0	25.5	44.5	C	39.5
516	C	2.94	1.74	4.14	1.23	3.9	-11	12.0	26.7	25.2	27.0	43.5	36.5	44.5
516	L[e]	1.66	1.46	3.61	1.66	3.6	-25	97.7	424	59.6	5.0	36.0	**17** (2)	30.5
516	LB[d]	4.55	1.80	3.92	1.11	3.6	-8	-0.4	56.8	-2.0	16.0	34.5	22.0	26.0
481	RB	2.82	1.11	2.42	1.73	3.6	-43	23.1	<0.1	36.1	22.0	47.5	26.0	39.5
481	R[f]	2.12	8.67	3.10	1.85	1.5	-56	74.7	24.8	70.3	n.a.	n.a.	n.a.	n.a.
481	C	3.23	1.16	2.92	1.38	1.1	-41	16.8	9.9	50.5	16.0	41.0	20.5	36.0
481	L[f]	3.10	1.56	3.45	1.25	2.5	-20	8.2	3.8	26.0	n.a.	n.a.	n.a.	n.a.
481	LB[c]	3.05	1.76	4.13	1.06	4.4	-10	5.0	16.5	13.6	12.0	33.5	C	25.0
July 2012 uncorrected blankets														
516	RB[c]	3.89	1.44	3.33	1.12	1.9	-26	3.0	31.9	17.1	20.0	38.0	C	31.5
516	R[c]	3.02	1.27	3.40	1.09	3.1	-35	10.2	<0.1	26.0	26.0	44.5	C	38.5
516	C	2.92	1.75	4.29	1.08	4.2	-11	8.1	29.5	19.3	16.5	39.5	29.5	32.5
516	L	2.48	1.65	4.12	1.19	4.1	-16	18.9	109	30.8	9.0	37.0	25.0	29.5
516	LB	3.82	1.78	4.13	1.05	3.9	-9	0.9	8.2	3.8	14.0	36.0	25.5	30.0
481	RB	2.83	1.17	2.59	1.62	3.6	-40	20.7	0.8	34.1	20.5	46.0	26.0	37.0
481	R	2.49	1.29	3.72	1.36	2.9	-34	29.0	40.7	43.1	20.5	46.5	39.5	45.0
481	C	3.20	1.23	3.09	1.32	1.4	-37	14.3	4.3	42.8	15.0	40.5	21.5	35.5
481	L[f]	3.10	1.57	3.45	1.25	2.5	-20	8.0	3.8	25.8	n.a.	n.a.	n.a.	n.a.
481	LB	3.03	1.76	4.16	1.05	4.4	-10	4.6	18.3	12.9	11.5	34.5	C	30.0

Table D5 Continued

						March 2013 points								
300	RB[d]	3.81	2.29	5.12	0.99	2.87	17	-2.3	<0.1	-27.7	n.a.	n.a.	n.a.	n.a.
300	R	3.69	2.10	5.03	1.43	5.15	7	12.6	<0.1	22.2	n.a.	n.a.	n.a.	n.a.
300	C	3.69	2.15	8.12	1.09	3.06	10	20.7	<0.1	36.8	n.a.	n.a.	n.a.	n.a.
300	L	3.47	1.77	5.91	0.96	3.76	-10	12.8	<0.1	26.6	n.a.	n.a.	n.a.	n.a.
300	LB	3.48	1.79	4.35	1.09	3.46	-8	5.4	<0.1	16.9	n.a.	n.a.	n.a.	n.a.
715	RB	3.35	1.53	3.68	1.63	4.42	-22	24.0	0.1	33.4	28.5	46.5	41.5	41.5
715	R[b]	3.52	n.a.	n.a.	n.a.	2.93	n.a.	n.a.	n.a.	n.a.	36.0	45.5	49.5	44.5
715	C	3.59	2.24	5.14	1.15	3.44	14	3.4	2.1	12.5	15.0	34.5	28.0	27.5
715	L[b]	3.62	n.a.	n.a.	n.a.	3.98	n.a.	n.a.	n.a.	n.a.	1.5	37.0	26.0	28.0
715	LB[g]	3.70	2.41	5.74	1.03	4.02	23	1.1	1.5	4.2	n.a.	37.5	27.5	28.0
1260	RB	3.95	2.18	5.04	1.29	5.39	12	7.1	3.4	15.0	14.0	43.5	23.0	37.0
1260	R	3.82	2.19	5.07	1.65	5.09	12	17.7	0.6	26.8	23.0	50.0	32.5	49.0
1260	C	3.80	2.27	5.48	1.75	4.06	16	23.4	0.1	34.3	36.5	49.5	59.0	50.5
1260	L[b,c]	3.61	n.a.	n.a.	n.a.	2.71	n.a.	n.a.	n.a.	n.a.	22.0	39.5	C	32.5
1260	LB	3.63	1.87	5.87	0.98	2.13	-4	10.5	0.5	32.0	28.0	43.5	35.5	39.0
1700	RB	3.80	2.27	5.17	1.35	4.59	16	8.7	0.1	19.1	n.a.	n.a.	n.a.	n.a.
1700	R	3.90	2.39	5.60	1.39	2.82	22	11.4	0.3	29.0	n.a.	n.a.	n.a.	n.a.
1700	C	3.88	2.45	7.38	0.95	0.29	25	5.9	<0.1	54.7	n.a.	n.a.	n.a.	n.a.
1700	L	3.48	1.90	4.40	1.49	3.88	-3	16.0	0.7	29.3	n.a.	n.a.	n.a.	n.a.
1700	LB	3.41	1.88	4.18	1.57	3.93	-4	16.8	2.1	29.8	n.a.	n.a.	n.a.	n.a.
1910	RB	4.09	2.62	1.12	0.64	0.28	34	3.7	1.3	47.9	n.a.	n.a.	n.a.	n.a.

Table D5 Continued

1910	R	4.12	2.61	1.07	0.67	0.26	33	4.3	1.9	51.6	n.a.	n.a.	n.a.	n.a.
1910	C	3.56	1.88	5.26	1.63	3.65	-4	30.7	<0.1	40.2	n.a.	n.a.	n.a.	n.a.
1910	L	3.00	1.45	3.27	2.06	3.98	-26	36.0	6.7	41.4	n.a.	n.a.	n.a.	n.a.
1910	LB[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2530	RB[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2530	R[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2530	C	2.97	1.16	2.62	1.56	3.84	-41	20.2	<0.1	32.9	n.a.	n.a.	n.a.	n.a.
2530	L	3.44	1.63	3.68	1.53	3.90	-17	17.0	0.2	30.1	n.a.	n.a.	n.a.	n.a.
2530	LB[b]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
						March 2013 corrected blankets								
715	RB[c,f]	4.02	1.70	3.95	1.38	4.87	-13	12.2	22.0	22.5	n.a.	C	27.0	C
715	R[g]	3.78	1.94	4.85	1.28	4.37	-1	11.7	2.6	23.3	n.a.	19.5	41.0	33.0
715	C[c]	3.99	2.51	6.64	1.25	4.11	28	12.1	0.7	24.6	20.5	C	27.5	24.5
715	L[d]	3.97	2.47	5.71	1.01	4.12	26	-1.5	0.7	-8.4	8.5	C	25.5	25.0
715	LB[b,h]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
						March 2013 blankets								
715	RB[c,e]	3.78	2.01	4.63	1.09	4.07	3	1.8	1.6	6.5	9.0	C	18 (**3**)	23.5 (**0**)
715	R[g]	3.77	1.98	4.83	1.21	4.03	1	6.8	1.8	17.7	n.a.	21.5 (**18.5**)	35.5	30.0
715	C	3.84	2.24	5.48	1.13	3.99	14	4.7	0.5	14.0	11.5	24.5 (**11.0**)	23.0	25.0
715	L	3.86	2.28	5.26	1.02	4.02	17	-0.7	0.5	-3.6	6.5	C	22.5	25.0
715	LB[b,h]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

[a]SF_6 and CFC samples were collected at a subset of point transects (481 m, 516 m, 715 m, 1260 m) and at all blanket transects (481 m, 516 m, 715 m)

[b]no Cu tube sample data; used [Ar] from glass bottle (USGS) samples if available; in two cases Cu tube data was available but showed very high (1260L) or very low (2530LB) gas concentrations
[c]one or more CFC concentrations indicated contamination
[d]negative $^{3}H/^{3}He$ age was interpreted as age = zero for calculating mean age and *MTT*
[e]ambiguous CFC age; younger age shown in parentheses; we used the CFC age closest to apparent age from other age dating tracers collected at the same point or blanket, shown in bold
[f]blanket correction resulted in negative concentration for SF_6 or CFC; no age estimate was made
[g]broken SF_6 bottle
[h]samples not collected due to low discharge from blanket

REFERENCES

Gilmore, T. E., Genereux, D. P., Solomon, D. K., Solder, J. E., Kimball, B.A., Mitasova, H., & Birgand, F. (2015). Quantifying the fate of agricultural nitrogen in an unconfined aquifer: stream-based observations at three measurement scales. *In Prep*.

Kennedy, C. D., & Genereux, D. P. (2007). ^{14}C groundwater age and the importance of chemical fluxes across aquifer boundaries in confined Cretaceous aquifers of North Carolina, USA. *Radiocarbon*, *49*(3), 1181–1203.

Robertson, W. D., Van Stempvoort, D. R., Solomon, D. K., Homewood, J., Brown, S. J., Spoelstra, J., & Schiff, S. L. (2013). Persistence of artificial sweeteners in a 15-year-old septic system plume. *Journal of Hydrology*, *477*, 43–54. doi:10.1016/j.jhydrol.2012.10.048

Visser, A., Broers, H. P., & Bierkens, M. F. P. (2007). Dating degassed groundwater with $^{3}H/^{3}He$. *Water Resources Research*, *43*(10), W10434. doi:10.1029/2006WR00584

Appendix E – Modeled NO_3^- Concentrations, NO_3^- Budget Uncertainty, and Streambed and Well Sampling Data

E1. Modeled groundwater N concentrations for well nests

Lumped-parameter models (LPM) were used to calculate hypothetical composite nitrogen (N) concentrations in aquifer outflow, which were then used in the groundwater nitrate mass balance calculations (Section 3.6.5, main text) and nitrate flux predictions (Section 3.6.6, main text). The models used were an exponential model (EM, equation 3.5, main text), and exponential-piston flow model (EPM, equation 3.6, main text) and a gamma distribution (equation 2.7, main text). Cumulative age distributions for each model were calculated (Fig. E1), and the incremental fraction of aquifer discharge (q_i) was determined for each 10-year age interval (Table E1), as described below. The composite N concentration in aquifer outflow, $[N]_{LPM}$, was then calculated as $\Sigma([N]_i q_i)$, where i indicates a 10-year age bin and [N] represents $[NO_3^-]$ or $[N_2\text{-den}]$. $[N]_{LPM}$ values from the three LPMs are shown in Table E2.

The EM and EPM models are based on an assumption of uniform porosity, which means that the rate of groundwater discharge is constant over the aquifer thickness L. Thus, the ratio of any depth z to the total thickness L is equivalent to the ratio of the discharge over length z to the total aquifer discharge over thickness L. To calculate z/L for a given age, the EPM equation was solved for z/L

$$\frac{z}{L} = 1 - \exp\left(-\left(\frac{\tau R_{EPM}}{L\theta} - \frac{x^*}{x}\right)\right) \quad \text{(E1)}$$

where τ is groundwater age, R_{EPM} is recharge to the unconfined portion of the aquifer, x^* and x are the length scales of the confined and unconfined portions of the aquifer, respectively, and θ is the porosity. Equation E1 is equivalent to the EM when x^*/x is set = 0. The upper and lower age limits for each age bin (e.g., 20 and 30 years, for the 20-30 year age bin) were

entered in equation E1 and the difference between the two resulting z/L values was taken as q_i (Table E1).

Similarly, for the gamma distribution, the function GAMMA.DIST(τ, α, β, "TRUE") was used in Excel® to calculate the cumulative fraction of discharge (Fig. E1), and q_i was calculated as the difference between the cumulative fractions calculated at the upper and lower age limits for each age bin (Table E1). The variables α, and β are the shape factor and scaling factor, respectively, for the gamma function, and the quantity $\alpha\beta = MTT$ [*Kirchner et al.*, 2010]. Using the "TRUE" setting produces a cumulative gamma distribution in Excel®.

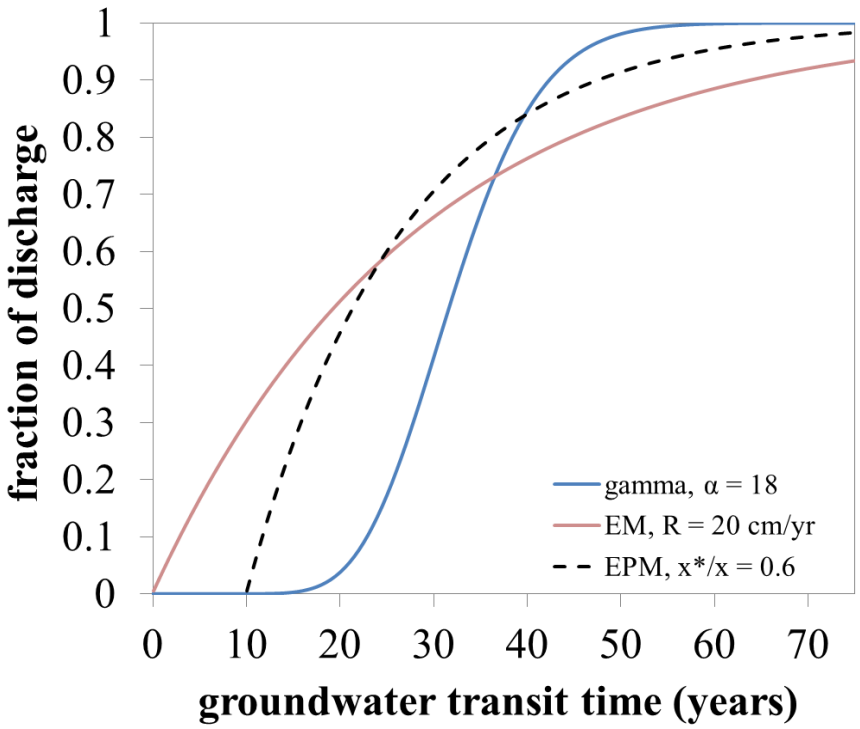

Figure E1. Curves showing the cumulative fraction of groundwater discharge versus age for the exponential model (EM), exponential-piston flow model (EPM), and the gamma distribution (gamma). Variables used to construct curves are discussed in the text.

The variables for each LPM curve in Fig. E1 were as follows. First, porosity of 0.35 and saturated thickness of 16 m were used for both EM and EPM curves. For the EM, the average recharge rate from the well nest data (20 cm/yr, Table 3.1 of main text) was used to calculate the curve in Fig. E1. The resulting *MTT* for the EM was 28 years, similar to the *MTT* from wells, which suggested the model was reasonable for calculating hypothetical $[NO_3^-]_{LPM}$ and $[N_2\text{-den}]_{LPM}$ for the nitrate mass balance. For the EPM, the model was fit through a minimum age of 10 years, the average predicted age at the upper limit of the aquifer ($z = 0$) from the EPM fits at BC1 and BC2. To calculate the curve shown in Fig. E1, the average EPM recharge rate from the BC1 and BC2 EPM fits ($R_{EPM} = 0.34$ m/yr) was entered in the EPM equation (equation 3.6, main text) and x^*/x was adjusted until the minimum predicted age was 10 years (*MTT* for this model was 26.5 years, and $R = 0.21$ cm/year). A gamma model with mean transit time of 32 years and shape parameter, α, equal to 18 was also used (these values gave the best fit for the 2.5 km reach age distribution [*Gilmore et al.*, 2015a]) to calculate hypothetical $[NO_3^-]_{LPM}$ and $[N_2\text{-den}]_{LPM}$.

Table E1. Calculated NO_3^- and N_2-den concentrations and q_i for well nests.

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	q_i, EM (-)	q_i, 10-year min. age EPM (-)	q_i, gamma distrib. σ = 18 (-)
0-10	532	150	0.30	0	0
10-20	489	422	0.21	0.46	0.04
20-30	266	336	0.15	0.25	0.39
30-40	3.5[a]	416	0.10	0.14	0.43
40-50	3.5[a]	4.9	0.07	0.07	0.13
50-60	no data[b]	no data[b]	0.05	0.04	0.02
60+	no data[b]	no data[b]	0.12	0.04	0.00

[a] $[NO_3^-]$ below detection (<7 μM) were set to 3.5 μM

[b]For groundwater with age > 50 years, $[NO_3^-]$ was assumed to be zero. In other words, for the prediction of NO_3^- flux from the aquifer (Fig. 3.7 in main text), denitrification was assumed to be complete for groundwater with apparent age > 50 years.

With q_i, $[NO_3^-]$, and $[N_2\text{-den}]$ known for each age bin (Table E1), $[NO_3^-]_{LPM}$, $[N_2\text{-den}]_{LPM}$ were calculated for the different lumped parameter models. For example, the process for calculation $[NO_3^-]_{LPM}$ was:

- for each age bin (each row in Table E1), multiply the value in the $[NO_3^-]$ column to the corresponding fraction in the "q_i, EM" column
- calculate $[NO_3^-]_{LPM} = \Sigma([NO_3^-]_i q_i)$

The resulting $[NO_3^-]_{LPM}$ and $2[N_2\text{-den}]_{LPM}$ values for the EM, EPM, and gamma distributions are shown in Table E2.

Table E2. Calculated N concentrations in aquifer discharge.

	EM	10-year min. age EPM	gamma distrib. $\sigma = 18$
	(μM)	(μM)	(μM)
$[NO_3^-]_{LPM}$	302	289	123
$2[N_2\text{-den}]_{LPM}$	226	332	325

The LPM and gamma models described above were also used to predict future fluxes of NO_3^- (f_{NO3}) from the surficial aquifer (Fig. 3.7, main text). The flux, f_{NO3}, at each 10-year time step was calculated as $v[NO_3^-]_{LPM}$, where v was set to the either a value measured in the streambed (58 m reach), or to an R converted to an equivalent v through the streambed (well

nests, 2.5 km reach). The conversion of R to v was $v = RA_R/A_S$, where A is the topographically-defined groundwater contributing area for the 2.5 km reach (subscript R, 5.91 x 10^6 m^2) or the streambed area (subscript S, 1.63 x 10^4 m^2). For this study site, the ratio of 2.5 km contributing area to the streambed area in the 2.5 km reach happened to be 364, almost the exact value of the conversion of 1 year = 365 days, meaning that the recharge rate calculated from the well data (20 cm/yr) was equal to v = 20 cm/day through the streambed for the 2.5 km reach. Similarly, the recharge rate determined from the 2.5 km reach data converted to v = 0.18 m/day. We could have used 0.18 m/day for the 2.5 km predictions and 0.20 m/day for the well nest predictions, but rather chose to use a single value (0.20 m/day) so the two prediction plots would be on the same basis. Thus 0.20 m/day was used in to calculate the NO_3^- prediction curves for the 2.5 km reach and well nest plots in Fig. 3.7 of the main text. The long-term recharge rate, rather than the v measured in the 2.5 km reach (0.4 m/day), because the latter may have been elevated during the wet conditions in March 2013 when the 2.5 km reach was sampled. In contrast, it seemed possible the v from July 2012 measurements in the 58 m reach is reasonably representative of a long-term average groundwater flux because v from a reach mass-balance experiment over the 2.5 km reach conducted at the same time (July 2012, v = 0.2 m/day) was similar to the v calculated from long-term recharge rates.

Table E3 shows example prediction calculations for $[NO_3^-]_{LPM}$ from the EPM model under the assumption of zero-order denitrification rate and no N use in the watershed. The data in each column and calculations are as follows:

- For all years after 2013, the value in reach row of the $[NO_3^-]$ column was calculated as $[NO_3^-]^0 - [N_2\text{-den}]$
- Values in the 2$[N_2\text{-den}]$ column were held constant on molar basis under the zero-order denitrification scenario (exceptions described below). For the first-order denitrification scenario the, percent removal was calculated for each row from the

2013 data (E_{den}) and then those percentages were used to recalculate the 2[N_2-den] column for each time step (2[N_2-den] = $[NO_3^-]^0 E_{den}$)

- q_i values show the fraction of aquifer discharge associated with each age bin. These values are multiplied by the $[NO_3^-]$ in the same row to create the $[NO_3^-]q_i$ column
- the $[NO_3^-]q_i$ column is summed to calculate $[NO_3^-]_{LPM}$ at each time step
- the $[NO_3^-]_{LPM}$ value is then multiplied by v to calculate f_{NO3} at each time step

For each time step, $[NO_3^-]^0$ values for each age interval were advanced forward 10 years and $[NO_3^-]$ for each age interval was calculated as $[NO_3^-]^0 - 2[N_2\text{-den}]$. Because of the assumption of no N use in the watershed, $[NO_3^-]^0 = 0$ was filled in as the observed values of $[NO_3^-]^0$ were advanced (e.g., after the first 10-year time step, $[NO_3^-]^0 = 0$ was filled in for the 0-10 year age interval, after the second time step, $[NO_3^-]^0 = 0$ was filled in for the 0-10 and 10-20 year age intervals). For the case where N use was assumed to persist at current levels, then $[NO_3^-]^0_{young}$ (e.g., 682 μM in this example) was advanced through the age intervals. In general, the value 2[N_2-den] remained unchanged for each age interval because zero-order denitrification was assumed. For this particular case, however, complete denitrification was assumed for groundwater with age > 40, due to little or no $[NO_3^-]^0$ in the original data set, and therefore the 2[N_2-den] values match the $[NO_3^-]^0$ values in those age intervals. For each simulated time step, the calculated $[NO_3^-]$ were weighted by the q_i shown in Table E1, and as described previously, $[NO_3^-]_{LPM}$ was calculated as $\Sigma([NO_3^-]q_i)$.

Table E3. Predicted N values over time for EPM, no N use, and zero-order denitrification rate.

Year[a] = 2013

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	$[NO_3^-]^0$ (μM)	q_i (-)	$[NO_3^-]q_i$ (μM)
0-10	532	150	682	0	0
10-20	489	422	911	0.46	223
20-30	266	336	602	0.25	66
30-40	3.5	416	418	0.14	0.5
40-50	3.5	5	6	0.07	0
50-60	no data	no data	no data	0.04	no data
60 +	no data	no data	no data	0.04	no data
				$[NO_3^-]_{LPM}$	289

Year = 2023

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	$[NO_3^-]^0$ (μM)	q_i (-)	$[NO_3^-]q_i$ (μM)
0-10	0	0	0	0	0
10-20	260	422	682	0.46	118
20-30	575	336	911	0.25	143
30-40	186	416	602	0.14	25
40-50	0	418	418	0.07	0
50-60	0	6	6	0.04	0
60 +	no data	no data	no data	0.04	no data
				$[NO_3^-]_{LPM}$	286

Year = 2033

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	$[NO_3^-]^0$ (μM)	q_i (-)	$[NO_3^-]q_i$ (μM)
0-10	0	150	0	0	0
10-20	0	422	0	0.46	0
20-30	346	336	682	0.25	86
30-40	495	416	911	0.14	67
40-50	0	602	602	0.07	0
50-60	0	418	418	0.04	0
60 +	0	6	6	0.04	0
				$[NO_3^-]_{LPM}$	152

Table E3 Continued

Year = 2043

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	$[NO_3^-]^0$ (μM)	q_i (-)	$[NO_3^-]q_i$ (μM)
0-10	0	150	0	0	0
10-20	0	422	0	0.46	0
20-30	0	336	0	0.25	0
30-40	265	416	682	0.14	36
40-50	0	911	911	0.07	0
50-60	0	602	602	0.04	0
60 +	0	418	418	0.04	0
				$[NO_3^-]_{LPM}$	36

Year = 2053

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	$[NO_3^-]^0$ (μM)	q_i (-)	$[NO_3^-]q_i$ (μM)
0-10	0	150	0	0	0
10-20	0	422	0	0.46	0
20-30	0	336	0	0.25	0
30-40	0	416	0	0.14	0
40-50	0	682	682	0.07	0
50-60	0	911	911	0.04	0
60 +	0	602	602	0.04	0
			$[NO_3^-]_{LPM}$		0

Year = 2063

age bin (years)	$[NO_3^-]$ (μM)	$2[N_2\text{-den}]$ (μM)	$[NO_3^-]^0$ (μM)	q_i (-)	$[NO_3^-]q_i$ (μM)
0-10	0	150	0	0	0
10-20	0	422	0	0.46	0
20-30	0	336	0	0.25	0
30-40	0	416	0	0.14	0
40-50	0	0	0	0.07	0
50-60	0	682	682	0.04	0
60 +	0	911	911	0.04	0
				$[NO_3^-]_{LPM}$	0

[a]year corresponds to year on Fig. 3.7 of main text.

E2. Uncertainty in groundwater nitrate balance

A Monte Carlo approach was used to evaluate the sensitivity of groundwater nitrate mass-balance calculations to uncertainty (W) in input variables in equation 3.7 of the main text. Upper and lower constraints ($\pm W$) were set for each variable based on estimated percent uncertainty, and then variables were randomly varied 5000 times. In Excel, this is accomplished using the formula NORM.INV(RAND(), *input variable*, σ), where the RAND() function randomly selects a probability value (0 – 1), and then the NORM.INV function returns the value associated with that probability from a normal distribution with mean equal to the input variable and standard deviation equal to σ. For each of the 5000 iterations a new dS_{NO3}/dt estimate was calculated. The coefficient of variation for the 5000 dS_{NO3}/dt estimates was used as a rough index for measurement uncertainty in dS_{NO3}/dt.

Table E4. Input variables for NO_3^- balance and estimated uncertainties.

Variable	Units	Value	% Uncertainty
		2.5 km reach	
R	cm/year	18	32%
$[NO_3^-]^0_{young}$	µM	637	25%
$[NO_3^-]_{FWM}$	µM	142	30%
$[N_2\text{-den}]_{FWM}$	µM	375	25%
		well nests	
R	cm/year	20	25%
$[NO_3^-]^0_{young}$	µM	682	25%
$[NO_3^-]_{LPM}$[a]	µM	302, 289, 123	40%
$[N_2\text{-den}]_{LPM}$[a]	µM	226, 332, 325	30%

[a]concentrations are shown in the order of EM, EPM, and gamma; see Table E2 and related text for concentrations and details of the calculations

Formal propagation of uncertainty in the $[NO_3^-]_{FWM}$ from the 2.5 km reach suggested uncertainty of only 5%, which is probably unrealistically low given the uncertainties in the measurement of *v* (about 10%), and 5% analytical uncertainty in $[NO_3^-]$, as well as uncertainty due to spatial variability in $[NO_3^-]$ in the 2.5 km reach [*Gilmore et al.*, 2015b]. To determine a more probable uncertainty, we considered the effect of removing two point measurements in the 2.5 km reach that stood out as having large f_{NO3} or $f_{N2\text{-}den}$ due to either high $[NO_3^-]$ (and $[N_2$-den]) or high seepage rate *v* relative to the rest of the data set (300L, 1260C, respectively, see Table E5 for relevant data). Removing the data collected at either 300L or 1260C data from the data set resulted in a roughly ± 50% change in $[NO_3^-]_{FWM}$ (+50% if 1260L data were removed, -50% if the 300L data were removed), suggesting that spatial variability was a significant source of uncertainty in $[NO_3^-]_{FWM}$ (removing both points resulted in a 10% change, to 130 μM). The individual *v*, $[NO_3^-]$, and $[N_2$-den] values measured at those two sampling locations (Table E5) were not unreasonable (based on the ranges of measurements from 422 measurements in West Bear Creek from *Kennedy et al.*, 2009, and from this study), but missing either sampling point during field work would have had a significant impact on the $[NO_3^-]_{FWM}$. Based on this sensitivity analysis and the much lower uncertainty estimate from formal propagation of error, we chose an uncertainty value of 30% as a rough index of the uncertainty in $[NO_3^-]_{FWM}$ measured in the 2.5 km reach (Table E4). The data collected at the 300L and 1260C points had a smaller effect on $[N_2\text{-den}]_{FWM}$, which changed by ±15% when data from either sampling location was removed from the $[N_2\text{-den}]_{FWM}$ calculation. But then, N_2-den is a modeled value with higher uncertainty than the analytical uncertainty for $[NO_3^-]$. *Gilmore et al.* [2015b] estimated that minimum uncertainty in $[N_2$-den] due to model error was probably about 13%. Based on these values, a 25% uncertainty seemed a reasonable approximation of uncertainty in $[N_2\text{-den}]_{FWM}$ (Table E4). The small sample sizes for $[NO_3^-]^0_{young}$ from wells and streambed sampling and for $[NO_3^-]_{LPM}$ and $[N_2\text{-den}]_{LPM}$ from well sampling make the assessment of uncertainty more difficult. The difference in the flow-weighted $[NO_3^-]^0_{young}$ from streambed sampling and from well nests was about 7% (637 μM versus 682 μM, respectively), but $[NO_3^-]^0_{young}$ is based on

measured $[NO_3^-]$ plus modeled 2$[N_2$-den] which collectively have uncertainty on the order of 15-20%; uncertainty in $[NO_3^-]_{young}$ was assumed to be 25%. Finally, uncertainty in $[NO_3^-]_{LPM}$ and $[N_2\text{-den}]_{LPM}$ was gauged by comparing the different concentrations derived from the different lumped-parameter models (LPMs) (Table E4). The coefficient of variation was 42% for $[NO_3^-]_{LPM}$, and 20% for $[N_2$-den]; those values were used as an estimate of uncertainty (Table E4). For *R*, the uncertainty was about 32% and 25% for values calculated from streambed data and well sampling data, respectively (Table 3.2, main text).

E3. Streambed and well nest sampling data and results

Table E5 shows groundwater concentrations, *v*, and apparent age results from streambed sampling. Concentrations and apparent ages are also shown for well samples.

Table E5. Streambed and well nest sampling data

transect (m)	site (-)	*v* (m/d)	$[NO_3^-]^a$ (μM)	$[N_2$-den] (μM)	$[NO_3^-]^0$ (μM)	app. age[b] (yrs)
		July 2012 streambed points				
466	RB	0.06	784	622	2029	22
466	R	0.12	1017	319	1655	22
466	C	0.09	474	352	1178	28
466	L	0.12	< 7	269	541	44
466	LB	0.13	1869	257	2382	19
474	RB	0.20	998	519	2036	28
474	R	0.72	< 7	329	661	39
474	C	0.10	911	421	1752	33
474	L	0.10	1823	207	2237	21
474	LB[c,e]	n.a.	n.a.	n.a.	n.a.	n.a.
481	RB	0.06	388	322	1032	37
481	R	1.44	< 7	218	440	n.a. (35)
481	C	0.94	685	506	1696	48
481	L	1.08	826	391	1609	33
481	LB	0.43	1593	24	1642	21
491	RB	0.09	1048	482	2012	21

Table E5 Continued

491	R[c]	0.52	418	n.a.	n.a.	n.a.
491	C	0.37	< 7	452	908	34
491	L	0.31	624	565	1753	31
491	LB	0.11	1467	386	2240	23
499	RB	0.06	1019	347	1713	27
499	R	0.67	144	701	1547	20
499	C	0.52	< 7	375	754	39
499	L	0.54	1480	199	1878	26
499	LB	0.97	1252	26	1305	-13[d]
508	RB	0.02	599	568	1735	11
508	R	0.93	281	398	1076	28
508	C	0.30	636	266	1168	35
508	L	0.16	1351	25	1400	11
508	LB[c]	0.02	1600	n.a.	n.a.	n.a.
516	RB	0.06	849	166	1182	9
516	R	0.84	996	343	1682	31
516	C	0.64	511	265	1040	44
516	L	0.69	1058	233	1525	33
516	LB	0.12	1266	0	1266	2
524	RB	0.05	846	278	1402	21
524	R	0.03	765	548	1860	29
524	C	0.06	< 7	285	574	38
524	L[c]	0.03	761	n.a.	n.a.	n.a.
524	LB	0.09	1136	138	1413	25
		March 2013 streambed points				
300	RB	0.06	67	53	173	-28[d]
300	R	0.80	< 7	230	464	22
300	C	0.72	< 7	128	260	37
300	L	0.73	1143	616	2375	27
300	LB	0.05	2041	129	2299	17
715	RB	0.13	98	415	928	33
715	R	0.14	23	262	548	n.a. (36)
715	C	0.74	604	104	812	12
715	L	0.14	402	166	734	n.a. (1.5)
715	LB	0.38	460	104	667	4
715	RB	0.04	59	174	406	15
1260	R	0.56	< 7	165	334	27
1260	C	4.64	< 7	145	294	34

Table E5 Continued

1260	L	0.04	338	229	797	n.a. (22)
1260	LB	0.00	116	246	608	32
1700	RB	0.02	< 7	116	236	19
1700	R	0.77	< 7	107	218	29
1700	C	0.02	< 7	72	148	55
1700	L	0.00	472	208	889	29
1700	LB	0.00	219	290	798	30
1910	RB	0.02	< 7	104	211	48
1910	R	0.68	< 7	67	137	52
1910	C	0.29	< 7	312	629	40
1910	L	0.47	34	282	598	41
1910	LB[c]	0.07	175	n.a.	n.a.	n.a.
2530	RB[c]	0.12	< 7	n.a.	n.a.	n.a.
2530	R[c]	0.02	< 7	n.a.	n.a.	n.a.
2530	C	0.01	< 7	62	129	33
2530	L	0.30	< 7	332	668	30
2530	LB[b]	0.01	< 7	n.a.	n.a.	n.a.
			well nest samples			
n.a.	BC1S	n.a.	113	209	531	15 (16)
n.a.	BC1M	n.a.	266	168	602	26 (20)
n.a.	BC1D	n.a.	< 7	2	6	49 (50)
n.a.	BC2S	n.a.	532	75	682	1 (12)
n.a.	BC2M	n.a.	865	213	1292	18 (14)
n.a.	BC2D	n.a.	< 7	208	418	35 (32)

[a]limit of detection for [NO_3^-] was 7 µM, samples < 7 µM were set to 3.5 µM in all calculations

[b]apparent ages are from $^3H/^3He$, with the exception of values shown in parentheses, which were from SF_6. SF_6 values were used in the calculation of age distributions, but not the calculation of *MTT* (only $^3H/^3He$ apparent ages were used for *MTT* calculations).

[c]no Cu tube sample data; used [Ar] from glass bottle (USGS) samples if available; in two cases Cu tube data was available but showed very high (1260L) or very low (2530LB) gas concentrations

[d]negative $^3H/^3He$ age was interpreted as age = zero for calculating mean age, *MTT*, and age distributions

[e]no sample collected

REFERENCES

Gilmore, T. E., D. P. Genereux, D. K. Solomon, and J. E. Solder (2015a), Groundwater age and mean transit time from streambed sampling in an agricultural coastal plain watershed, North Carolina, USA, *in prep*.

Gilmore, T. E., D. P. Genereux, D. K. Solomon, J. E. Solder, B. A. Kimball, H. Mitasova, and F. Birgand (2015b), Quantifying the fate of agricultural nitrogen in an unconfined aquifer: stream-based observations at three measurement scales, *in prep*.

Kirchner, J. W., D. Tetzlaff, and C. Soulsby (2010), Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments, *Hydrol. Process.*, *24*(12), 1631–1645, doi:10.1002/hyp.7676.