12-2013

Quantifying groundwater-surface water interactions to improve the outcomes of human activities

Ryan Parish Gordon

Abstract

Interactions between surface water and groundwater impact both the quality and quantity of water resources. This dissertation is focused on the interactions between surface water in streams and groundwater in hyporheic zones and shallow fluvial aquifers, and how human beings influence these interactions and are influenced by them. The three chapters of this dissertation span scales from cm to km and locations from Upstate New York to the Peruvian Andes, but are united by the goal to improve the scientific understanding of groundwater-surface water interactions (GWSWI) using novel techniques, in order to improve the outcomes of human activities for people and ecosystems.

Heat is a useful tracer for quantifying GWSWI, but analyzing large amounts of raw thermal data has many challenges. Chapter 1 presents a computer program named VFLUX for processing raw temperature time series and calculating vertical water flux in shallow sub-surface-water systems. The workflow synthesizes several recent advancements in signal processing, and adds new techniques for calculating flux rates with large numbers of temperature records from high-resolution sensor profiles. The program includes functions for quantitatively evaluating the ideal spacing between sensor pairs, and for performing error and sensitivity analyses for the heat transport model due to thermal parameter uncertainty. The new method is demonstrated by processing two field temperature time series datasets collected using discrete temperature sensors and a high-resolution DTS profile. The analyses of field data show vertical flux rates significantly decreasing with depth at high-spatial resolution as the sensor profiles penetrate shallow, curved hyporheic flow paths, patterns which may have been obscured without the unique analytical abilities of VFLUX.

Natural channel design restoration projects in streams often include the construction of cross-vanes, which are stone, dam-like structures that span the active channel, and are often thought to increase local hyporheic exchange. In Chapter 2, vertical hyporheic exchange flux (HEF) and redox-sensitive solutes were measured in the streambed around 4 cross-vanes with different morphologies. Observed patterns of HEF and redox conditions are not dominated by a single, downstream-directed hyporheic flow cell beneath cross-vanes. Instead, spatial patterns of moderate (<0.4 md^{-1}) upwelling and downwelling are distributed in smaller cells around pool and riffle bed forms upstream and downstream of structures. Patterns of biogeochemical cycling are controlled by dissolved oxygen concentrations and resulting redox conditions, and are also oriented around secondary bedforms. Total hyporheic exchange volumes are ~0.4% of stream discharge in restored reaches of 45-55 m. Results show that shallow hyporheic flow and associated biogeochemical cycling near cross-vanes is primarily controlled by secondary bed forms created or augmented by the cross-vane, rather than by the cross-vane itself. This study suggests that cross-vane restoration structures benefit the stream ecosystem by creating heterogeneous patches of varying HEF and redox conditions in the hyporheic zone, rather than by processing large amounts of nutrients to alter in-stream water chemistry.

Melting tropical glaciers supply approximately half of dry season stream discharge in glacierized valleys of the Cordillera Blanca, Peru. The remainder of streamflow originates as groundwater stored in alpine meadows, moraines and talus slopes. A better understanding of the dynamics of alpine groundwater, including sources and contributions to streamflow and GWSWI, is important for making accurate estimates of glacial inputs to the hydrologic budget, and for our ability to make predictions about future water resources as glaciers retreat. The field study described in Chapter 3 focused on two high-elevation meadows in valleys of the Blanca.

Tracer measurements of stream and spring discharge and groundwater-surface water exchange were combined with synoptic sampling of water isotopic and geochemical composition, in order to characterize and quantify contributions to streamflow from different geomorphic features. In a valley headwaters study site, groundwater supplied approximately half of stream discharge from the meadow, with most originating in a debris fan adjacent to the meadow and little from the meadow itself (6%); however, in at a mid-valley study site, where meadows are extensive, local groundwater has a large impact on stream flow and chemistry through large net discharge and fractional hydrologic turnover.

Quantifying groundwater-surface water interactions to improve the outcomes of human activities

by

Ryan P. Gordon

B.A. Williams College, 2005

A dissertation, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Earth Sciences.

Syracuse University
December 2013

Acknowledgements

I would like to thank Laura Lautz, my advisor, for introducing me to the world of hydrology, for being an excellent mentor and friend over the last five years, and for showing me the way in both work and family life. Thank you to Don Siegel, for your wisdom and humor, Jeff McKenzie for your technical expertise and lots more humor, Bryan Mark for your openness and sense of adventure, and Ted Endreny for your infectious curiosity. Thanks are also due to the hydrology group at Syracuse University, especially Martin Briggs and Timothy Daniluk for fruitful collaboration and camaraderie, and AnneMarie Glose for accompanying me all the way to Peru. Thanks to my parents Suzanne and Geoff for a lifetime of support, experience and example. Finally, thank you, Naomi. This work is dedicated to Willa Adelle.

Table of Contents

List of Tables

List of Figures

Introduction

The interactions between surface water and groundwater are increasingly being recognized for their impact on both the quality and quantity of water resources (Winter *et al.*, 1998; Poole, 2010). My graduate research and this dissertation are focused on the interactions between surface water in streams and groundwater in hyporheic zones and shallow fluvial aquifers, and how human beings influence these interactions and are influenced by them. Broadly, my goal is to improve the scientific understanding of groundwater-surface water interactions (GWSWI) using novel techniques, in order to improve the outcomes of human activities for people and ecosystems. This dissertation is divided into three distinct research projects, described in Chapters 1 through 3. The three chapters address GWSWI at a range of increasing scale from the centimeter to kilometer, and describe research conducted in different geographic settings with different immediate concerns; however, all three are united by similar hydrologic processes and the methods used study and understand them, as well as by my broader goal to improve the outcomes of human activities on water resources.

Chapters 1 and 2 are focused on the impacts of stream restoration on hyporheic exchange fluxes and associated biogeochemical cycling in the hyporheic zone, and on the methods used to quantify hyporheic exchange in these systems. Hyporheic exchange is the transfer of water between stream channels and the adjacent hyporheic zone, which is comprised of shallow saturated sediments of the streambed and banks where water from the active channel resides for a time before returning to the channel (Triska *et al.*, 1989; White, 1993; Harvey and Wagner, 2000; Gooseff, 2010). Stream restoration is a growing billion-dollar industry in the United States (Bernhardt *et al.*, 2005); however, hyporheic connectivity is very rarely a target of stream restoration projects, despite recent calls by several researchers for a more complete approach to

restoration that includes hyporheic processes (Boulton, 2007; Hester and Gooseff, 2010). The hyporheic zone is an important component of ecological restoration because it provides several functions of healthy streams, including habitat heterogeneity, temperature regulation, and redox processing of nutrients and anthropogenic pollutants (Hester and Gooseff, 2010).

Hyporheic exchange itself is difficult to measure in the field, especially at small (cm to m) scale, but recently heat has been used successfully to trace and quantify the movement of water in the hyporheic zone by a variety of methods (Anderson, 2005; Lautz and Ribaudo, 2012). Chapter 1 introduces a method and computer program for processing raw temperature records and calculating vertical water flux at cm resolution in the hyporheic zone. The method, named VFLUX, improves the temporal and spatial resolution of flux estimates using heat, and also provides an accessible and complete computer program to the hydrologic community to make heat transport modeling more approachable, methodologically consistent, and quantifiably precise. The VFLUX method is used extensively in the research described in Chapter 2, during which I investigated patterns of hyporheic exchange and redox-sensitive nutrient processing in restored stream reaches in Central New York.

While modeling studies have predicted that certain built restoration structures may promote the exchange of water and solutes between the stream and the hyporheic zone (e.g. Hester and Doyle, 2008), few studies have investigated this connection in the field (notable exceptions include Kasahara and Hill, 2006; Kasahara and Hill, 2006; Fanelli and Lautz, 2008; Lautz and Fanelli, 2008). In Chapter 2, I used heat and diurnal temperature variations as tracers (including the VFLUX method of Chapter 1) in order to quantify the vertical fluxes of water between the stream and subsurface (Conant, 2004; Hatch *et al.*, 2006; Gordon *et al.*, 2012) around built restoration structures. I also collected shallow pore water samples from piezometers

in the streambed, analyzed the water samples for redox-sensitive nutrients and metals, used statistical methods to quantify patterns of redox cycling, and matched those patterns to the patterns of water flux. My results show that stream restoration does create flow cells of hyporheic exchange, but that single large restoration structures do not facilitate as much nutrient processing as predicted, and that smaller bedforms created by channel redesign are responsible for most of the patterns of redox cycling over the bed. I hope that the results of this research will change the way that scientists, engineers, and restoration practitioners envision and construct stream restoration projects (Gordon *et al.*, 2013).

My work on stream restoration in the US shows how a better understanding of GWSWI can improve the results of human activities, in this case by making stream restoration more effective at rehabilitating important ecosystem services such as pollution remediation and biodiversity. The research project described in Chapter 3 expands my study of humans and GWSWI to a different type of human-natural system: the interactions between surface water resources and shallow proglacial aquifers in a semi-arid mountainous area of Peru, where diverse livelihoods depend on consistently available stream water.

The Cordillera Blanca of Peru is the most glacierized tropical mountain range in the world. The glaciers of the Cordillera are retreating rapidly due to climate change, which, along with population growth, threatens water resources for the quarter-million inhabitants of the upper Santa valley and many more downstream (Bury *et al.*, 2011). The high seasonal variation in precipitation also contributes to the unique vulnerability of this location to climate change and water stress (Mark *et al.*, 2010). Recent studies have shown that glacial melt supplies approximately half of dry season stream discharge in Cordillera Blanca valleys (Mark *et al.*, 2005), and melt water buffers streamflow throughout the range, making it less susceptible to

variations in precipitation (Mark and Seltzer, 2003). The remainder of dry-season streamflow is supplied by groundwater. In the future, when glacier loss has reduced the influence of melt water on the streams of the Cordillera, groundwater discharge will be the sole dry-season source of stream water for irrigation, municipalities, and hydropower in the Santa watershed, putting livelihoods and economies at risk (Bury *et al.*, 2011). A better understanding of the dynamics of alpine groundwater, including sources and exchange fluxes, is therefore vitally important for future planning in this region. Understanding these groundwater-surface water interactions is important for making accurate estimates of meltwater contributions to the hydrologic budget, and for our ability to make predictions about future water resources under deglaciated conditions.

In the research described in Chapter 3, I measured gains and losses of stream water from and to shallow aquifers in alpine meadows, moraines, talus slopes, and an alluvial fan in two proglacial vallies of the Cordillera Blanca. I used reach-length chemical tracer experiments to quantify the net water exchange between selected stream reaches and underlying aquifers. I also performed synoptic sampling of stream water, groundwater springs, and glacial melt end-members to determine the impact of GWSWI on surface water chemistry and on water supply. The results of this study show that groundwater is an influential element in the proglacial hydrologic landscape. Groundwater makes up a large percentage of incipient stream flow in the highest headwaters of glacierized catchments, and in mid-valley meadow systems, GWSWI are substantial and influence stream discharge rates and geochemistry, with sizeable groundwater contributions to streamflow that represent an important dry season water resource. I hope that my results will expand the current state of knowledge about groundwater storage and its interaction with surface water resources in tropical glacierized catchments and high-elevation watersheds in general.

Chapter 1.

Automated calculation of vertical pore-water flux from field temperature time series using the VFLUX method and computer program

Chapter 1 has been published as:

Gordon, R.P., L.K. Lautz, M.A. Briggs, and J.M. McKenzie, 2012, Automated calculation of vertical pore-water flux from field temperature time series using the VFLUX method and computer program, *Journal of Hydrology*, 420-421:142-158, doi: 10.1016/j.jhydrol.2011.11.053.

Abstract

Heat is a useful tracer for quantifying groundwater-surface water interaction, but analyzing large amounts of raw thermal data has many challenges. We present a computer program named VFLUX, written in the MATLAB computing language, for processing raw temperature time series and calculating vertical water flux in shallow sub-surface-water systems. The step-by-step workflow synthesizes several recent advancements in signal processing, and adds new techniques for calculating flux rates with large numbers of temperature records from high-resolution sensor profiles. The program includes functions for quantitatively evaluating the ideal spacing between sensor pairs, and for performing error and sensitivity analyses for the heat transport model due to thermal parameter uncertainty. The program synchronizes and resamples temperature data from multiple sensors in a vertical profile, isolates the diurnal signal from each time series and extracts its amplitude and phase angle information using Dynamic Harmonic Regression (DHR), and calculates vertical water flux rates between multiple sensor pairs using heat transport models. Flux rates are calculated every one-to-two hours using four similar analytical methods. One or more "sliding analysis windows" can be used to automatically identify any number of variably spaced sensor pairs for flux calculations, which is necessary when a single vertical profile contains many sensors, such as in a high-resolution fiber-optic distributed temperature sensing (DTS) profile. We demonstrate the new method by processing two field temperature time series datasets collected using discrete temperature sensors and a high-resolution DTS profile. The analyses of field data show vertical flux rates significantly decreasing with depth at high-spatial resolution as the sensor profiles penetrate shallow, curved hyporheic flow paths, patterns which may have been obscured without the unique analytical abilities of VFLUX.

1. Introduction

While heat has been used for many years as a tracer of groundwater-surface water interaction (Anderson, 2005), several recent developments have improved the process of estimating vertical water flux to or from surface water using temperature time series, including the introduction of one-dimensional analytical solutions to the heat transport equation (Hatch et al., 2006; Keery et al., 2007), improved methods for signal processing of raw temperature time series (Keery et al., 2007; Young et al., 1999), and the deployment of many sensors in high-resolution profiles (Briggs et al., in revision; Vogt et al., 2010). These advancements have made the estimation of pore-water flux easier, more accurate, and more useful, but their combined benefits have not yet been compiled into a step-by-step workflow that can be easily automated for processing large amounts of thermal data or performing complicated sensitivity analyses.

In streams, heat is a naturally occurring, non-reactive tracer that is easy to measure and model, making it a practical tool for studying water fluxes through the streambed (Constantz, 2008). Numerical models of fluid and heat flow (e.g. Healy and Ronan, 1996; Vogel et al., 2010) have often been calibrated to fit temperature patterns at specific boundaries (e.g. Niswonger and Prudic, 2003; Ronan et al., 1998; Vogel et al., 2011). Simplified analytical models have also been developed for idealized boundary conditions, such as the steady state model of Bredehoeft and Papadopulos (1965) and the transient model of Stallman (1965). Steady state models were originally employed in the geothermal zone, but have since been used to quantify exchange with surface water (e.g. Schmidt et al., 2007). Lapham (1989) used vertical temperature profiles beneath streams to quantify the rate of vertical water flow using a numerical approximation of the one-dimensional, analytical heat transport equation of Stallman (1965). Analytical extensions to the Stallman model have since been developed (Hatch et al., 2006;

Keery et al., 2007) that solve for the one-dimensional flux of fluid between two vertically-spaced temperature sensors using the phase and amplitude changes of transient thermal signals.

Analytical modeling of transient heat transport takes advantage of the daily fluctuation in temperature that takes place in stream water. As streams heat and cool during the diurnal cycle, a quasi-sinusoidal temperature signal with a 24-hour period propagates into the streambed. The depth, speed, and strength of the propagation of the thermal front depends on the heat capacity and conductivity of water and sediment, as well as the volume of fluid moving vertically through the sediment (Stonestrom and Constantz, 2003). By measuring the attenuation of amplitude and the increase in phase angle of the temperature signal with depth, it is possible to quantify the magnitude and direction of the vertical component of water flux (Hatch et al., 2006; Keery et al., 2007).

Temperature is easy and inexpensive to measure in streams and the streambed, and the analytical models of Hatch et al. (2006) and Keery et al. (2007) are relatively simple. However, processing raw temperature time series and extracting the information necessary to compute flux, namely amplitude and phase, is not trivial. Fluctuations in stream temperature are caused not only by daily variation in solar incidence, but by seasonality, weather patterns, precipitation and snowmelt, stream shading, long-wave radiation, changes in latent heat, upstream watershed characteristics, and even wind (Caissie, 2006). Furthermore, temperature sensors have different noise characteristics, depending on the technology used. Real-world temperature time series therefore contain a variety of signals and noise, and filtering techniques are required to isolate the signal of interest, in this case the 24-hour, diurnal oscillation (Hatch et al., 2006; Keery et al., 2007). After filtering, objective procedures are needed to extract signal amplitudes and phase angles to compute flux rates.

A variety of methods have been used by previous researchers to filter and analyze temperature time series to calculate fluid flux rates. Hatch et al. (2006) isolated the diurnal signal using a cosine taper band-pass filter and then selected daily temperature maxima and minima (using a semi-automated computer program) to calculate amplitude attenuation and time lag with depth. Fanelli and Lautz (2008) similarly identified daily extrema manually, but did not filter the raw temperature data before applying the Hatch (2006) method, producing quantifiable errors (Lautz, 2010). Keery et al. (2007) used Dynamic Harmonic Regression (DHR) (Young et al., 1999) both to isolate the diurnal signal and to extract amplitude and phase information. Vogt et al. (2010) similarly used the amplitude and phase output from a DHR analysis of temperature time series to solve for flux using the analytical solution of Stallman (1965). Swanson and Cardenas (2010) avoided filtering altogether by fitting a stationary sine function to individual day-long periods of temperature records, and using the amplitude and phase of the sine wave in the analytical equations of Hatch et al. (2006) and Keery et al. (2007).

Recently, scientists have used heat modeling methods in research on such diverse topics as streambed chlorobenzene concentrations (Schmidt et al., 2011), wetland phosphorus mobility (Maassen and Balla, 2010), aquifer recharge from perennial pools (Rau et al., 2010), potential culvert removal (Anderson et al., 2010), hyporheic exchange around geomorphic features (Crispell and Endreny, 2009; Fanelli and Lautz, 2008; Lautz et al., 2010), streambed infiltration (Vogt et al., 2010), groundwater discharge at stream cross-sections (Jensen and Engesgaard, 2011), and contaminated groundwater contribution to streams (Lautz and Ribaudo, in review). However, these studies have all used different methods for collecting, filtering, and analyzing temperature time series and estimating vertical flux, some with distinct flaws. Each researcher has had to devise individual solutions to—or simply ignore—a host of challenges created by

field temperature records, and they have been forced to develop their own computational workflows to deal with the volume of thermal data collected. Furthermore, researchers are increasingly collecting very large datasets from many sensors in high-resolution temperature profiles using fiber-optic distributed temperature sensing (DTS) methods (e.g. Briggs et al., in revision; Vogt et al., 2010). This proliferation in data gives researchers greater flexibility when selecting the ideal spacing between temperature sensors (Hatch et al., 2006), and allows them to detect changes in vertical flux with depth to greater precision, but also creates a computational problem in which the number of potential flux calculations increases quadratically with the number of sensors.

We believe that heat transport modeling would be used more often by a wider range of researchers if an automated, step-by-step program were freely available to easily and consistently calculate vertical fluid flux from raw temperature time series, without resorting to shortcuts in signal processing. Such an automated workflow in the form of a computer program would make processing very large amounts of data from many temperature sensors more practical and flexible. The wide availability and use of such a program would support a more consistent application of time series analysis methods among different studies by hydrologists, ecologists, geochemists, engineers, and others. By automating the calculation of vertical fluid flux, the program may also help avoid computational error.

In this paper, we present a computer program that automates the entire process of calculating vertical flux rates from raw temperature time series in the shallow beds of streams or other surface water bodies. The method builds upon previous work (Hatch et al., 2006; Keery et al., 2007; Vogt et al., 2010) and adds new techniques for processing large numbers of temperature records from high-resolution sensor profiles with greater spatial and temporal

resolution. We automate the new method using a computer program named VFLUX (Vertical Fluid Heat Transport Solver). The program formats, synchronizes, and resamples temperature data from multiple sensors in a vertical temperature profile (TP), then isolates the diurnal signal from each time series using DHR, extracts amplitude and phase information, and calculates vertical flux rates between multiple sensor pairs using a "sliding analysis window" and the analytical models of Hatch et al. (2006) and Keery et al. (2007). Flux rates are calculated every one-to-two hours, depending on the sampling rate of the original data, or at every sample interval if the original data was sampled more coarsely than every two hours. The program is written in the MATLAB computing language, and is designed to integrate with other command-based MATLAB functions or scripts, allowing it to be easily modified and incorporated into custom workflows. To illustrate this point, we demonstrate two MATLAB functions that run VFLUX iteratively in order to perform analyses of error due to thermal parameter uncertainty, as well as sensitivity of the model to each thermal parameter. The VFLUX program will be useful to other researchers who need a flexible and robust method for automatically calculating vertical flux rates from profiles of temperature sensors, including complex error analysis, and especially for practitioners of high-resolution DTS sensors. We demonstrate our method and the functions of VFLUX by processing two field temperature time series datasets collected using different sensor technologies. These examples illustrate the challenges associated with modeling real-world time series, and how they can be overcome using the VFLUX method and program.

At the time of publication, only one automated method for processing raw temperature time series and calculating vertical flux has been published as a computer program, named Ex-Stream (Swanson and Cardenas, 2011). VFLUX differs from the program of Swanson and Cardenas in several ways. Ex-Stream is operated through a graphical user interface, while

VFLUX is a command-line tool executed using well-documented commands, which allows it to be called by custom scripts and easily incorporated into other MATLAB programs. VFLUX employs DHR for extracting the diurnal signal from a time series and estimates changing flux at hourly time scales, while Ex-Stream operates by fitting a static sine function to the raw data day-by-day (Swanson and Cardenas, 2011), which may not fit well to parts of the time series that deviate significantly from a perfect sinusoid, such as during weather pattern shifts, and only allows flux calculations at daily intervals. Ex-Stream also uses two steady-state solutions (Bredehoeft and Papadopulos, 1965; Schmidt et al., 2007), in addition to the analytical solutions of Hatch et al. (2006) and Keery et al. (2007). Finally, VFLUX is designed for calculating flux between many variably-spaced sensor pairs in a single profile of temperature sensors, which is useful for detecting changing flux with depth and for identifying the ideal sensor spacing for a particular experiment, and is necessary when employing many sensors, such as in a high-resolution DTS profile. Ex-Stream only calculates flux between one pair of sensors in a profile at a time.

To our knowledge, VFLUX is the first published computer program for estimating water flux that automates a complete approach to data pre-processing of raw temperature time series, including multiple sensor synchronization, resampling, and filtering of a single tracer frequency with a robust method like DHR. It is also the first method that uses a "sliding window" to provide flux results at many depths with high spatial resolution. The program includes new tools for evaluating the sensitivity of the heat transport model to sensor spacing, which allows the user to intelligently determine the ideal sensor spacing for a particular application. The VFLUX package also includes two separate functions that perform analyses of model sensitivity to thermal parameters and estimates of error due to parameter uncertainty. Automated error

analysis and sensitivity programs such as these have not yet been published, and many recently published papers do not include error or uncertainty estimates for reported flux values (e.g. Schmidt et al., 2011; Vogt et al., 2010). The analysis of field temperature data in this paper is also the first high-resolution presentation, to our knowledge, of vertical flux rates that decrease with depth due to penetration through curved hyporheic flow cells.

2. Methods

2.1 Field data collection

A set of thermal records was collected in the bed of Ninemile Creek, a fourth-order stream (contributing area to site: 177 km^2) in Marcellus, New York between August 27 and September 18, 2009. The streambed at the site was composed of cobbly, sandy gravel with some silt. Thermal parameters of the saturated sediment were estimated through field observation and the guidelines provided by Lapham (1989), and are summarized in Table 1. Seven iButton Thermochron band-gap temperature sensors and loggers (Maxim Integrated Products, Inc., Sunnyvale, CA) were embedded in a steel rod to make a vertical TP (Figure 1a). The steel TP rod was driven into the streambed at the head of a riffle so that one sensor was 0.05 m above the streambed interface, and the others were positioned at 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30 m below the streambed interface. The sensors have a resolution of 0.0625º C, a manufacturer-reported accuracy of ±0.5º C and a manufacturer-reported thermal response time of 130 s, much shorter than our sampling interval. In a laboratory test of 80 iButtons, including those used in the present study, all sensors were precise within a range of 0.12º C in an ice bath and at room temperature. The sensors were programmed to record temperature every 10 minutes, for a total of 3070 measurements at each sensor. All seven sensors were synchronized prior to deployment,

so that measurements were taken on the same time schedule. There were no breaks in sampling, so the measurements were evenly spaced in time throughout.

The second set of thermal records was collected in the bed just upstream of a small beaver dam in Cherry Creek, a second-order stream (contributing area to site: 31.1 km^2) in Lander, Wyoming between July 11 and August 11, 2010 (see Briggs et al., in revision, for further details). The streambed was composed of interbedded organic-rich silt, sand, and gravels. Thermal parameters of the saturated sediment were again estimated through field observation and the relationships presented by Lapham (1989), and are summarized in Table 1. Temperature measurements were collected using a fiber-optic distributed temperature sensing (DTS) method, in which the temperature-dependent backscatter of a laser pulse is used to continuously estimate temperature along a glass fiber (Dakin et al., 1985; Selker et al., 2006). The fiber used at Cherry Creek was housed in stainless-steel cladding and wrapped in a tight coil around a one-meter long, threaded PVC rod (Figure 1b). An Agilent Distributed Temperature Sensor (model N4386A, Agilent Technologies, Inc., Santa Clara, CA) was used to measure temperature every meter along the cable length, yielding a temperature measurement every 0.014 m along the length of the vertical TP rod. The rod was driven vertically into the streambed and positioned such that temperature was measured at 57 locations between 0.014 and 0.78 m depth below the streambed interface. The DTS instrument in this experimental configuration had a precision of ±0.2 ºC, estimated using the manufacturer's software, and was calibrated using mixed thermal baths monitored with iButton Thermochrons (± 0.5 ºC accuracy, 0.0625 ºC resolution, 130 s thermal response time). The DTS system recorded temperature every 20 minutes, for a total of 2182 measurements at each of the 57 depths along the TP. Due to the continuous nature of DTS, measurements of temperature at different depths were all

synchronized in time; however, because the battery of the DTS unit needed to be changed daily, some sample times were missed. There were 30 gaps in the data that were less than 81 minutes in length, and 3 gaps due to equipment malfunctions that were between 203 and 524 minutes in length. The missing samples were linearly interpolated from the original data and inserted into the time series, so that the time series processed with VFLUX was evenly spaced in time.

2.2 The VFLUX program

The temperature time series from Ninemile Creek and Cherry Creek were processed with the VFLUX program to calculate vertical water flux rates at multiple depths in the streambeds of the study sites. VFLUX is distributed as a MATLAB toolbox, a set of functions written in the MATLAB computing language that are designed to run in the MATLAB environment. The program was developed in MATLAB 7.10.0 (R2010a) (The MathWorks, Inc., Natick, MA). The VFLUX functions are executed by typing commands at the MATLAB command prompt, and can also be called by external functions or integrated into custom scripts. The functions are distributed as open-source code (in MATLAB M-files) that is free to use, easily modified, and well-commented. The VFLUX program files are included in supplementary material available with the online version of this article, and the most up-to-date version may be downloaded from the following web site: http://hydrology.syr.edu/lautz_group/vflux.html. Full documentation is provided along with the program, which explains in detail the program's commands, inputs, outputs, options, and switches.

Given a set of temperature time series from a single vertical sensor profile and the depth of each sensor, the VFLUX program will perform the following six major steps (see Figure 2): (1) format and synchronize all the time series to a single vector of sampling times; (2) low-pass filter and resample the time series; (3) isolate the fundamental signal (the signal of interest,

typically diurnal) using DHR; (4) extract amplitude and phase information for the fundamental signal using DHR; (5) identify pairs of sensors based on one or more “sliding analysis windows”; and (6) calculate vertical water flux rates between the identified sensor pairs. Additional programs included with the VFLUX package perform sensitivity analyses on the thermal input parameters, and calculate confidence intervals for flux estimates based on thermal parameter uncertainty.

2.2.1 Format and synchronize time series

Time series data for each temperature sensor are input as a vector (a one-dimensional array) of evenly-spaced sample times, in days, and a corresponding vector of observed temperatures, in degrees C. If all the sensors were sampled at the same time schedule (i.e., the time vectors for all sensors have the same start time, end time, and sampling rate), then VFLUX simply formats the time series into a MATLAB structure array to be processed by the remainder of the program. However, if different sensors in the temperature profile were sampled at different time schedules (i.e., they have different start or end times, and/or different sampling rates), then VFLUX synchronizes all the time series by reducing them to the “lowest common denominator”; that is, it trims all the input series to the shortest time range common to all and, if necessary, linearly interpolates all the input series at the same sampling rate. This creates a single time vector that is common to all the temperature data. The interpolation method can easily be changed to a spline interpolation or another method.

2.2.2 Low-pass filter and resample time series

The second step in the VFLUX program is to reduce the sampling rate to approximately 12 to 24 samples per fundamental cycle, if the original sampling rate was greater. This step

improves the filtering results because VFLUX uses the standard frequency-domain optimization method within the Dynamic Harmonic Regression (DHR) model. This optimization method is sensitive to oversampling, and functions best when the sampling rate is not much higher than the frequency of the oscillation of interest (Young et al., 1999; Włodzimierz Tych, personal communication, March 3, 2011). If the time series are oversampled, then the signal of interest becomes compressed in the frequency domain, and the optimization method is ineffective at identifying the model hyperparameters. Although it is beneficial in this case to resample at a lower rate before applying DHR, oversampling in the collection of raw data is still desirable, because it allows for the identification and reduction of noise by low-pass filtering during the resample process, and it avoids the greater problem of undersampling. If the raw data is undersampled during collection (fewer than approximately 12 samples per cycle), then important information about the signal may be irrecoverably lost (Box et al., 1994).

VFLUX first calculates an appropriate integer factor by which to reduce the sampling rate (called *rfactor*). The program then decimates the time vector by this rate, keeping every *rfactor*-th sample time and discarding the rest. Alternately, *rfactor* can be specified by the user when calling the VFLUX program. VFLUX then resamples each temperature vector using an anti-aliasing, lowpass FIR (finite impulse response) filter designed with a Kaiser window and decimates the filtered signal by *rfactor*, discarding samples that are not needed. VFLUX reduces the edge effects of the filtering process by reflecting and mirroring the temperature vector at both ends as padding, and then discarding this padding after running the resample command.

The low-pass filter reduces high-frequency noise that is inherent in the natural temperature signal or is an artifact of the sensor technology. The anti-aliasing properties of the

filter prevent noise from reflecting around the Nyquist frequency and disrupting the signal identification in the next step.

2.2.3 Isolate the fundamental tracer signal

The third step in the VFLUX program is to isolate a tracer signal of a single fundamental frequency (typically the diurnal signal). This is performed using DHR as implemented in the Captain Toolbox (Young et al., 2010), a set of MATLAB functions developed at Lancaster University. DHR is a method for non-stationary time series analysis that is particularly useful for extracting harmonic signals from dynamic environmental systems (Young et al., 1999). DHR was first used in hydrologic heat modeling by Keery et al. (2007). The driving temperature oscillation in a stream system changes in time due to weather and seasonality, and vertical water flux rates may also change in time; therefore, the amplitude and phase of the diurnal signal in the streambed are also time-varying, and a non-stationary approach to signal extraction is necessary. DHR, a simplification of the Unobserved Component model, has the following form:

$$y_t = T_t + C_t + e_t \tag{1}$$

where y_t is the observed time series, T_t is a trend or zero-frequency component, C_t is a cyclical component, and e_t is an irregular, white-noise component (Young et al., 1999). The cyclical term is modeled as a sum of the fundamental signal and its associated harmonics:

$$C_t = \sum_{i=1}^{N} \left[a_{i,t} \cos(\omega_i t) + b_{i,t} \sin(\omega_i t) \right] \tag{2}$$

where $a_{i,t}$ and $b_{i,t}$ are stochastic time-varying parameters (TVPs) and $\omega_1, \omega_2, \ldots, \omega_N$ are the fundamental frequency (ω_1) and its harmonics ($\omega_i = i\,\omega_1$) up to the Nyquist frequency (ω_N). The trend component T_t can also be considered as a zero-frequency term ($\omega_0 = 0$) incorporated into the cyclical term sum. This DHR model can be thought of as a non-stationary extension of the

discrete Fourier transform, where the amplitude and phase of each time series component themselves change with time. Identification of the TVPs is achieved in a stochastic state space formulation using two-step Kalman filtering and fixed-interval smoothing (Young et al., 1999).

VFLUX calls the DHR function for each time series to separate the fundamental cyclical signal from the trend, noise, and harmonics. In practice, only the harmonic components that are present in the original temperature data need to be identified, typically ω_1, ω_2, and ω_3. Therefore, by default VFLUX attempts to identify a trend, the fundamental signal (ω_1), and the first and second harmonics (ω_2 and ω_3) using an auto-regression (AR) frequency spectrum created with the Captain Toolbox. VFLUX displays a diagnostic plot of the AR spectrum for each time series, allowing the user to evaluate whether the appropriate number of harmonics was identified. The specific harmonics, and the method that VFLUX uses to identify them, can be modified by the user. VFLUX then fits the DHR model to the AR spectrum, by optimization to a non-linear least-squares objective function using the Captain Toolbox (Young et al., 1999). The AR spectrum and the model fit are plotted and displayed to the user. Finally, VFLUX filters and isolates the trend, fundamental signal, identified harmonics, and noise components. These components are then plotted and displayed to the user.

2.2.4 Extract amplitude and phase information

The fourth step in the VFLUX program is performed during the DHR analysis above. The amplitude and phase of any harmonic component at any discrete time can be calculated by the equations:

$$A_{i,t} = \sqrt{a_{i,t}^2 + b_{i,t}^2} \tag{3}$$

$$\phi_{i,t} = \tan^{-1}\left(\frac{a_{i,t}}{b_{i,t}}\right) \quad (4)$$

where $A_{i,t}$ is the amplitude and $\phi_{i,t}$ is the phase angle for the component with frequency ω_i at time t (Vogt et al., 2010). The Captain Toolbox function 'dhr' supplies the amplitude and phase angle of the fundamental frequency at each time step in the resampled temperature time series, producing an amplitude time series and a phase time series for each sensor. From Equation 4, the phase angle reported by DHR must be between $-\pi/2$ and $\pi/2$; it is therefore possible that the phase angle of a time series reported by DHR may jump from $\pi/2$ to $-\pi/2$ or from $-\pi/2$ to $\pi/2$ during the duration of the time series. VFLUX converts the phase angle into a total phase offset for each sample of each time series by adding π whenever the phase time series jumps from $\pi/2$ to $-\pi/2$, and subtracting π whenever the phase time series jumps from $-\pi/2$ to $\pi/2$.

2.2.5 Identify sensor pairs

The analytical methods of Hatch et al. (2006) and Keery et al. (2007) (the Hatch and Keery methods) calculate water flux between a single pair of vertically-spaced temperature sensors. If more than two sensors are deployed in the same profile, multiple combinations of sensors can be used to calculate vertical flux. The multiple sensor pairs may have different midpoint depths, different separation distances, and/or they may overlap in space (for an example, see Jensen and Engesgaard, 2011). When a TP contains many sensors, such as in a high-resolution DTS profile, the number of possible combinations of sensors can become very large (equal to $N(N\text{-}1)/2$ in a profile of N sensors). The main benefit of using DTS technology in a vertical temperature profile is the high spatial resolution and therefore the large number of effective sensors available, which allow for many possible sensor spacings at different depths. A large number of possible sensor pairs with a range of separation distances can be very useful to

help identify the ideal sensor spacing for a particular flux rate and set of field conditions (Hatch et al., 2006). It is also important to have control over the depth of the sensor pairs in order to detect changes in the vertical component of hyporheic flux, which may vary significantly with depth (Buffington and Tonina, 2009).

The large number of potential pairs in a high-resolution DTS profile leads to a very large number of calculations. VFLUX is designed to automatically identify any number of sensor pairs through the use of one or more "sliding analysis windows". The window(s) are specified by the user in units of "sensor-spacings" when calling the VFLUX program. The actual distance between sensors is not used for identifying pairs, so sensors do not need to be evenly-spaced. For example, a window of 2 sensor-spacings is depicted in Figure 1a. As this analysis window slides from the top of the TP to the bottom, it identifies all the sensor pairs that are separated by 2 spacings. With a total of 6 sensors in the profile, this window would identify sensors 1 and 3, 2 and 4, 3 and 5, and 4 and 6 as pairs. Multiple windows can be defined in a single VFLUX run; if we included a window of 1 sensor-spacing in the above example, then VFLUX would also identify sensors 1 and 2, 2 and 3, 3 and 4, 4 and 5, and 5 and 6 as pairs. The maximum number of possible pairs in this example is 15, which would all be identified by defining windows of 1, 2, 3, 4, and 5 sensor-spacings.

When processing the data from Ninemile Creek with VFLUX, every possible window was used, i.e., 1, 2, 3, 4, 5, and 6 sensor spacings, which correspond to sensor separation distances of 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30 m. This leads to 21 total sensor pairs. In the case of Cherry Creek, the high-resolution DTS profile was processed with windows of 8, 9, 10, 11, and 12 sensor spacings, which correspond to sensor separation distances of 0.110, 0.124, 0.138, 0.152, and 0.166 m, for a total of 235 sensor pairs. VFLUX performed flux calculations

at each two-hour time step for every pair of sensors that was separated by any of these distance values. A total of 4,809 flux calculations were made with the data from Ninemile Creek, and 85,540 flux calculations with the data from Cherry Creek.

2.2.6 Calculate vertical flux

The sixth and final step in the VFLUX program is to calculate vertical water flux between the identified sensor pairs. The flux value calculated for each pair is assigned to a point equidistant between the sensors, referred to as the center-of-pair depth. The Hatch and Keery methods are both analytical solutions to the one-dimensional heat transport equation:

$$\frac{\delta T}{\delta t} = \kappa_e \frac{\delta^2 T}{\delta z^2} - q \frac{C_w}{C} \frac{\delta T}{\delta z} \tag{5}$$

where T is temperature (ºC), t is time (s), κ_e is the effective thermal diffusivity of the saturated sediment ($m^2\ s^{-1}$), z is depth (m), q is fluid flux ($m\ s^{-1}$), C is the volumetric heat capacity of the saturated sediment ($J\ m^{-3}\ ºC^{-1}$), and C_w is the volumetric heat capacity of the water (Goto et al., 2005; Stallman, 1965). C is calculated as the mean of C_w and C_s, the volumetric heat capacity of the sediment grains, weighted by total porosity. Both Hatch et al. (2006) and Keery et al. (2007) developed methods to solve for q in Equation 5 by measuring the attenuation of the amplitude of a quasi-sinusoidal temperature signal as it propagates vertically through the streambed, or by measuring the speed at which it propagates.

The analytical solution provided by Hatch et al. (2006) solves for the vertical water flux between two sensors as a function of either amplitude or phase differences between the sensors' temperature signals:

$$q = \frac{C}{C_w}\left(\frac{2\kappa_e}{\Delta z}\ln A_r + \sqrt{\frac{\alpha + v^2}{2}}\right) \tag{6}$$

$$|q| = \frac{C}{C_w}\sqrt{\alpha - 2\left(\frac{4\pi\,\Delta t\,\kappa_e}{P\Delta z}\right)^2} \tag{7}$$

where q is vertical fluid flux in the downward direction (m s^{-1}), A_r is the ratio of amplitudes (a measure of amplitude attenuation) between the lower sensor and the upper sensor (dimensionless), Δz is the distance between the two sensors in the streambed (m), v is the velocity of the thermal front (m s^{-1}), Δt is the time lag (a measure of the speed of signal propagation) between the two temperature signals (s), P is the period of the temperature signal (s), and α is defined by:

$$\alpha = \sqrt{v^4 + \left(\frac{8\pi\kappa_e}{P}\right)^2} \tag{8}$$

Effective thermal diffusivity, κ_e, which incorporates conductivity and dispersivity, is given by:

$$\kappa_e = \left(\frac{\lambda_o}{C}\right) + \beta|v_f| \tag{9}$$

where λ_o is the baseline thermal conductivity (i.e., thermal conductivity without the dispersive effect of fluid flow) of the saturated sediment (J $s^{-1}m^{-1}$ $^oC^{-1}$), β is thermal dispersivity (m), and v_f is the linear particle velocity (m s^{-1}) (Hatch et al., 2006).

Note that flux, q, is a true volumetric flux rate (volume per area per time), while linear particle velocity, v_f, is a true velocity (distance per time), and both are distinct from the thermal front velocity, v, despite all having the same units. The relationships between these terms are:

$$v_f = \frac{q}{n_e} \tag{10}$$

$$q = \frac{C}{C_w}v \tag{11}$$

where n_e is the effective porosity (dimensionless) (Hatch et al., 2006).

The analytical solution to the Stallman (1965) equation provided by Keery et al. (2007) is similar to the Hatch method, but does not include thermal dispersivity. Vertical water flux is again calculated as a function of amplitude or phase differences (Keery et al., 2007):

$$\left(\frac{H^3 \ln A_r}{4\Delta z}\right) q^3 - \left(\frac{5H^2 \ln^2 A_r}{4\Delta z^2}\right) q^2 + \left(\frac{2H \ln^3 A_r}{\Delta z^3}\right) q + \left(\frac{\pi C}{\lambda_o P}\right)^2 - \frac{\ln^4 A_r}{\Delta z^4} = 0 \qquad (12)$$

$$|q| = \sqrt{\left(\frac{C\Delta z}{\Delta t C_w}\right)^2 - \left(\frac{4\pi\Delta t\lambda_o}{P\Delta z C_w}\right)^2} \qquad (13)$$

where $H = C/\lambda_o$. Note that in both Equations 7 and 13, the phase methods only give the magnitude of flux, and not its direction.

In order to apply the Hatch and Keery methods, VFLUX first must calculate the amplitude ratio and time lag at each time step (i.e., at each sample) in the resampled and filtered fundamental signal. At each sample (time t), the time lag, Δt, is defined by:

$$\Delta t = \frac{P}{2\pi}\left(\phi_{z+\Delta z,t+\Delta t} - \phi_{z,t}\right) \qquad (14)$$

where $\phi_{z,t}$ is the phase angle of the sensor at depth z and at time t (see Equation 4), and $\phi_{z+\Delta z,t+\Delta t}$ is the phase angle of the sensor deeper by Δz at time $t+\Delta t$ (Vogt et al., 2010). However, because Δt must be known to determine the phase angle at $t+\Delta t$, and because the phase angle typically changes slowly in time on a daily scale, we assume it is generally safe to calculate the time lag using the estimate:

$$\Delta t \approx \frac{P}{2\pi}\left(\phi_{z+\Delta z,t} - \phi_{z,t}\right) \qquad (15)$$

which is the formulation used in VFLUX. The program then calculates the amplitude ratio, A_r, by:

$$A_r = \frac{A_{z+\Delta z,t+\Delta t}}{A_{z,t}} \qquad (16)$$

where $A_{z,t}$ is the amplitude of the sensor at depth z and at time t (see Equation 3), and $A_{z+\Delta z,t+\Delta t}$ is the amplitude of the sensor deeper by Δz at time $t+\Delta t$ (Keery et al., 2007). The amplitude ratio and time lag are illustrated in Figure 3.

VFLUX calculates vertical water flux, q, for each sample in the resampled time series using the Hatch amplitude and phase methods (Equations 6 and 7) and the Keery amplitude and phase methods (Equations 12 and 13). The sediment and thermal properties total porosity, dispersivity, baseline thermal conductivity, and volumetric heat capacity of the sediment and water must be input when running VFLUX, and the other parameters in Equations 6, 7, 12, and 13 are calculated by the program.

Because of the relationship in Equation 11, the thermal front velocity is found on both sides of Equations 6 and 7 and they must be solved iteratively. VFLUX solves Equations 6 and 7 by subtracting q from both sides of the equation and then finding the value of q at which the right side of the equation equals zero. Equation 12 must be solved by finding the roots of the third-order polynomial on the left side. At least one root must be real, and the other two may be real or complex, although only a result with one real root represents a true physical flux rate (Keery et al., 2007). VFLUX solves the third-order polynomial and only accepts a single real root as the value for q. If all three roots are real, then VFLUX warns the user and automatically assigns a null value (Keery et al., 2007). Equation 13 is solved for $|q|$ explicitly as written. If any value for A_r, Δt, or q cannot be calculated for any reason during a VFLUX run, then a null value is assigned, a warning is displayed to the user and the reason is described. The most common reasons that VFLUX fails to calculate flux are that A_r is greater than or equal to one or Δt is less than or equal to zero (both usually caused by non-ideal sensor spacing), or an attempt to find roots failed to converge on a real value.

There is an optional switch in the VFLUX program that can be set to also exclude flux values (write a null value) if the calculated flux estimate is outside a specified range of optimal sensitivity. This sensitivity check is only available for the Hatch amplitude method, and must be enabled in the VFLUX code. As described in Results and Discussion below, Hatch et al. (2006) identified a range of flux values for which the analytical heat transport model is most sensitive. This range depends on the sensor spacing Δz and the specific thermal properties, and is based on the slope of the amplitude ratio versus flux curve, or the derivative dA_r/dq (Hatch et al., 2006). If this slope is too low, which occurs at very high and low A_r values, then small errors in the amplitude ratio can cause large errors in flux. If this option is enabled, VFLUX will calculate dA_r/dq and assign a null value if the derivative is below a certain value, 0.001 d m^{-1} by default (after Hatch et al., 2006). Ultimately, violations of this sensitivity range are caused by sensors that are spaced too close together (the amplitude ratio is too close to 1) or too far apart (the amplitude ratio is too close to 0), therefore, this low-derivative check acts like an additional constraint on the ideal sensor spacing.

When the VFLUX program runs to completion, it creates four data matrices containing the results of the flux calculations: one matrix each for the Hatch amplitude method, Hatch phase method, Keery amplitude method, and Kerry phase method, plus a vector containing the time at which each flux calculation was made. VFLUX also creates metadata matrices that describe the sensor pairs used for each set of flux calculations, and the thermal parameters used in the heat transport models. At the end of a successful VFLUX run, the program can optionally enter a results visualization routine. If selected by the user, the visualization routine displays several plots, containing raw and filtered time series, amplitude and phase information, and final flux results through time for all four methods. Default MATLAB zooming, panning, and editing

tools can be used to customize the plots. The visualization routine also can display bar charts of successful flux calculations for different sensor pairs and window sizes, to help in the analysis of ideal sensor spacing, as discussed below. These plots show the percent of time steps for which flux was successfully calculated for each sensor depth and each window size. This visualization illustrates which sensor pair combinations are within the ideal sensitivity range of the models. Other combinations of results can be displayed after exiting VFLUX by plotting flux through time using the 'plot' commands in MATLAB.

It is recommended that the first two or three days of flux data from the beginning and end of the data collection periods be discarded, due to the edge effects of digital filtering (Hatch et al., 2006; Keery et al., 2007), although that choice is left up to the user and is not implemented in VFLUX. In the case of the field data from Ninemile Creek and Cherry Creek used in this study, the first and last two days of flux results were discarded.

2.2.7 Sensitivity and uncertainty estimation programs

The VFLUX package contains two additional programs, "vfluxsens", for performing sensitivity analyses, and "vfluxmc", for performing Monte Carlo error estimations. These programs both run the VFLUX method iteratively in order to produce different estimates of flux from different thermal parameters. The sensitivity analysis program is run using similar inputs to the standard VFLUX, but instead of single thermal and sediment input parameters, the user inputs a low and high value for each parameter in addition to the estimated or measured value ("base" value). The program first runs VFLUX only for the base values, then subsequently runs VFLUX twice for each parameter, using the input high and low values, while maintaining all other parameters at the base values. At the end, for each sensor pair, the program displays a series of plots, each showing maximum and minimum values of flux through time for each of the

varied parameters. The sensitivity analysis was performed for the data collected at Ninemile Creek, with sediment and thermal parameter ranges as shown in Table 1. The high and low values in Table 1 represent the full range of typical physical values for coarse-grained sediment and water between 0 and 25 degrees C from Lapham (1989), Fetter (2001), and CRC (2011).

The Monte Carlo error estimation program can be used to create statistical confidence intervals around flux estimates through time. Such confidence intervals are needed to determine if flux estimates are significantly different from zero, or to compare differences in flux between different sites or even different depths. Many recent publications that report flux values calculated with analytical heat transport methods do not include confidence intervals (e.g. Crispell and Endreny, 2009; Engelhardt et al., 2011; Schmidt et al., 2011; Vogt et al., 2010), making it difficult to determine if flux estimates are significant. The Monte Carlo estimation program has similar inputs to the standard VFLUX and sensitivity analysis programs, but takes as inputs mean and standard deviation values for each thermal and sediment parameter, which can be based on estimations or actual laboratory or field measurements. The program performs 1000 realizations (by default). For each realization, the program selects a set of random parameters from normal distributions with mean and standard deviation, as input. All but two of the parameters are uncorrelated, so that the random values selected in each realization are unrelated. Thermal conductivity and porosity are inversely correlated around their respective means, following guidelines in Lapham (1989), which show that thermal conductivity varies with dry bulk density. Porosity also affects the calculated value of C, the volumetric heat capacity of the sediment-water matrix. In each of the realizations, the program runs the standard VFLUX analysis with the randomly generated parameter values. Flux values are calculated for each sensor pair and each point in time during each of the 1000 realizations, and then standard

deviation values are calculated through all 1000 realizations. Finally, approximate 95-percent confidence intervals are created around the mean flux values by adding and subtracting two standard deviations from the mean flux values. This Monte Carlo analysis was performed for the data collected at Ninemile Creek, with parameter means and standard deviations as shown in Table 1.

3. Results and Discussion

Two examples of raw temperature time series from Ninemile Creek in New York and Cherry Creek in Wyoming are plotted in Figure 4. A diurnal temperature signal with a 24-hour period of oscillation is obvious at 0.05 and 0.30 m depths at Ninemile Creek (Figure 4a) and at 0.097 m depth at Cherry Creek, and it can also be made out by eye despite the noise at 0.51 m depth in Cherry Creek (Figure 4b). In both cases, the deeper time series has smaller amplitude and is shifted forward in time relative to the shallower time series, as we would expect from the models of Hatch et al. (2006) and Keery et al (2007). In both raw data sets, the time series exhibit non-sinusoidal features such as noise, asymmetry, and changes in trend from day-to-day (i.e., the daily mean temperature changes from one day to the next). The DTS data particularly appear to contain significant high-frequency noise, due to the relatively low precision (±0.2 ºC) of the DTS system used in this specific configuration. Both datasets have enough noise, natural or sensor-derived, to occasionally create local maxima and minima that do not correspond to the daily maximum and minimum temperatures. In both stream systems, temperature rises faster in the morning and falls more slowly in the evening, leading to asymmetric or saw-toothed patterns. This periodic asymmetry gives significant amplitude to the first few harmonics of the fundamental diurnal signal, which will be seen during the DHR analysis below. In both data sets there are also variations in the temperature trend (which can be thought of as a moving average

daily temperature), which are caused by changing weather patterns and differing amounts of solar radiation. All of these deviations from an ideal sinusoid are common in natural temperature time series and can cause error in flux calculations (Lautz, 2010).

Both Hatch et al. (2006) and Keery et al. (2007) note that the amplitude ratio method and the phase lag method differ in their sensitivity to flux depending on sensor spacing, signal frequency, and flux rate, with the amplitude ratio being more sensitive at lower rates on the order of $1x10^{-5}$ m s^{-1} (Hatch et al., 2006), which are more typical of general stream values found in the literature (e.g. Fanelli and Lautz, 2008; Hatch et al., 2006; Keery et al., 2007; Lautz et al., 2010). Furthermore, Lautz (2010) demonstrated that the amplitude ratio method is more reliable in non-ideal conditions, which likely exist at most field sites, and Hatch et al. (2006) notes that filtering can create spurious phase shifts. Most researchers have focused on the amplitude ratio method in their research (e.g. Fanelli and Lautz, 2008; Keery et al., 2007; Schmidt et al., 2011; Vogt et al., 2010), and we do the same here by concentrating this discussion on the amplitude ratio results.

A number of researchers have investigated potential errors in water flux estimates using heat modeling methods, including the effects of error in temperature measurement, thermal diffusivity, and sensor spacing (Shanafield et al., 2011), sediment heterogeneity (Ferguson and Bense, 2011; Schornberg et al., 2010), and non-vertical flow (Lautz, 2010; Rau et al., 2010). These non-ideal conditions are certainly present to some degree at both field sites, particularly non-vertical flow (see Section 3.6, below), but the potential uncertainty is not discussed in detail. However, we do conduct a sensitivity and uncertainty analysis for errors in sediment and thermal properties, which demonstrate how VFLUX can be a useful tool for investigating uncertainty and error in modeling or field studies such as Lautz (2010) or Shanafield et al. (2011).

3.1 Sensor synchronization

In both data sets, from Ninemile Creek and Cherry Creek, the temperature sensors in each TP were synchronized by nature of the experimental setup, so synchronization of the time series using VFLUX was unnecessary. The first step performed by VFLUX was therefore simple; the program copied the time series into a MATLAB structure to be processed by the remainder of the program. In other potential cases, however, especially those that use discrete sensors and data loggers like iButtons, the sensors may not collect measurements in sync. In these cases, synchronization is an important step in processing for two reasons. First, it is important that all time series be sampled at the same rate so that the discrete-time DHR analysis can be performed consistently on all the time series in a TP. Different discrete sampling rates have different bandwidth capacities, and therefore may contain different frequency spectra (Box et al., 1994), making comparison between DHR results more difficult. Second, DHR as implemented in the Captain Toolbox is blind to the absolute time of each point in a time series, and calculates phase angles based on the position of the first data point in each time series (Young et al., 2010). If the first data point is shifted in absolute time from one time series to another, the Captain Toolbox will not recognize this and the phase shift, Δt, in Equation 15 will be calculated incorrectly. Finally, it is obviously necessary that the time series from all sensors be complete over the same time period, so that comparisons between the signals can be made for every day and every possible pair of sensors.

3.2 Oversampling and high-frequency noise

The resample routine is an important part of the VFLUX process because it reduces oversampled data and removes high-frequency noise. Oversampling reduces the effectiveness of DHR optimization and filtering, as described above in Methods. The raw data from Ninemile

Creek was sampled every 10 minutes, for a sampling rate of 144 samples per day, which is much higher than the rate of approximately 12-to-24 samples per day that is ideal for describing a perfect diurnal signal. However, oversampling in the original data collection stage has no disadvantage as long as the data are resampled prior to a DHR analysis, and oversampling has the advantage of better describing the diurnal signal in relation to high-frequency noise, which is then removed by the low-pass filter. When analyzing the Ninemile Creek time series, VFLUX chose an integer resample factor (*rfactor*) of 12, so that the reduced sampling rate was 12 samples per day. Figure 5a shows an example of data from Ninemile Creek (0.05 m depth) at its original sampling rate and after being resampled and low-pass filtered by VFLUX. Because the original data collected at Ninemile Creek was not overly noisy, the resampled time series appears to be very similar to the original series in form, but simply described by fewer data points. The time series from Cherry Creek, however, include a lot of high frequency noise (Figure 5b). Temperature at Cherry Creek was sampled every 20 minutes, for a sampling rate of 72 samples per day. The value of *rfactor* was therefore calculated to be 6, and the reduced sampling rate was again 12 samples per day. Comparing a subset of the resampled data with the raw data in Figure 5b shows that the low-pass FIR filter has reduced the high-frequency noise and smoothed the signal, in addition to lowering the sampling rate.

3.3 Non-sinusoidal raw temperature signals

Departure from an ideal sinusoidal temperature signal creates several challenges when calculating flux rates from real-world temperature time series using the Hatch and Keery methods, especially if relying on daily extrema to calculate amplitude attenuation and time lag (e.g., Fanelli and Lautz, 2008; Hatch et al., 2006; Lautz, 2010) or if fitting a static sine function to the raw time series (e.g., Swanson and Cardenas, 2010). First, the analytical solutions to the

one-dimensional heat transport equation are based on the assumption that the temperature oscillations at the surface and at depth are sinusoidal waves of a single frequency (Hatch et al., 2006; Keery et al., 2007; Stallman, 1965). However, real-world temperature time series contain many harmonic and trend components, as previously discussed, requiring the frequency of interest to be isolated by filtering. Second, the presence of local maxima and minima (such as in the second and third days of Figure 4a) make the detection of daily extrema more challenging with an automated peak-detection program, although this can certainly be overcome with good programming. Third, the saw-toothed pattern found in many natural temperature series (including those from Ninemile and Cherry Creeks) mean that maxima and minima do not occur twelve hours apart; therefore, a different phase angle would be inferred from the maximum time than from the minimum time. Furthermore, the time lag (Δt) between a shallow and deep sensor is not the same when computed with the maximum and the minimum of an irregular signal, because the slower temperature change before each minimum allows the streambed to respond faster, with a shorter time lag (Lautz, 2010). Fourth, when the temperature trend changes through time, estimates of the signal amplitude (made by halving the difference between daily extrema) can be very inaccurate, and can change dramatically from one day to the next despite a constant flux rate. Using Figure 4a as an example, the temperature difference between the maximum at 3.1 days and the following minimum is higher than twice the true diurnal signal amplitude because the trend is decreasing quickly over that time period. Alternately, if one were to use the difference between the minimum at 2.9 days and the maximum at 3.1 days, the estimated amplitude would be too low, for the same reason. This example also shows that calculating the amplitude using a maximum and the following minimum can often give a different amplitude estimate for the same day than using a minimum and the following

maximum. Abrupt changes in trend, such as those that occur with frontal weather patterns, can entirely mask the presence of a daily maximum or minimum, making it impossible on that day to calculate flux using extrema. When the raw time series deviates substantially from a pure sinusoid, it may also be very difficult to fit the raw data with a static sine function (Swanson and Cardenas, 2011), leading to substantial errors in the determination of amplitude and phase using that method, as well (for example, a static sinusoid fits poorly to any of the 24-hour periods shown in Figure 4a between days 2 and 5).

These challenges illustrate how an accurate filtering mechanism is important for accurate time series analysis, and how an objective, consistent method of extracting amplitude and phase is needed to calculate flux rates. DHR is a robust method of filtering that can identify harmonic components in a complex time series with greater accuracy than a band-pass filter (Keery et al., 2007). DHR can successfully identify the diurnal signal and separate it from trend and harmonics even when the trend is changing quickly, so flux can be calculated on days when it might otherwise be impossible. DHR is also fully automated and can function objectively and consistently on a wide range of datasets without the need for human selection of extrema or other temperature patterns (Keery et al., 2007). VFLUX has been programmed to make the application of DHR straightforward for the user for this particular application.

VFLUX processed the time series from each Ninemile Creek sensor using the same DHR settings. An auto-regression (AR) spectrum of data from the 0.05 m sensor is shown in Figure 6a, showing the peaks in power for various components of the temperature record. A similar figure is produced by VFLUX for each time series using the ‘arspec’ function from the Captain Toolbox and is shown to the user during program execution. Note the strong peaks in power at the fundamental frequency (the diurnal oscillation at a period of 12 samples per cycle) and its

first two harmonics (with periods of 6 and 4 samples per cycle). Two lesser peaks can be seen in the spectrum near the third and fourth harmonics (3 and 2.4 samples per cycle), but their power is very low and their identification by periodicity is not clear. The half-peak in the spectrum at the zero frequency (infinite period) to the left of the fundamental peak represents the trend component in Equation 1, or a zero-frequency component in Equation 2. The strong harmonic response at 2 and 3 times the fundamental frequency is due to the asymmetric but periodic saw-tooth temperature patterns, as any periodic function that deviates from a pure sinusoid can be decomposed into a sum of sinusoidal harmonics at multiples of the fundamental frequency (for details of basic Fourier theory, see, e.g., Boas, 1983). Figure 6b shows the fit of the DHR model (Equations 1 and 2) to the time series AR spectrum, optimized by VFLUX using the DHR optimization function 'dhropt' from the Captain Toolbox. The fit of the model to the data appears very good along the trend and the fundamental and first harmonic peaks, and good for the second harmonic, as well. The optimized value of the non-linear (logarithmic) least squares objective function is 171.5. DHR has successfully differentiated the fundamental diurnal frequency from its trend and asymmetric components. The results of the DHR analysis for the Ninemile Creek 0.05 m sensor are shown in Figure 7. The DHR model matches the actual data well, and has extracted the diurnal signal even in difficult sections with local extrema and changing trend, such as between 1 and 5 days.

The AR spectrum and the DHR model fit to data from 0.097 m depth in Cherry Creek are shown in Figure 8a. The trend and diurnal signal are strong, with a weaker first harmonic and a very weak second harmonic. The optimized DHR model fits the data well over the trend and diurnal peak, but not as well over the two harmonics. The non-linear least squares objective function is 166.4. A similar figure for data from 0.51 m depth in Cherry Creek is shown in

Figure 8b. The diurnal signal is much weaker at this depth compared to the sensor noise, and the DHR optimization is not as effective at modeling the trend or fundamental signal. As a result, the fitted spectrum in Figure 8b does not fit the actual spectrum well over the trend or diurnal frequency peaks. The value of the objective function, at 355.4, is more than twice the value for the fit to the time series at 0.097 m. At all sensors below 0.51 m depth at Cherry Creek, the mean amplitude of the diurnal signal is less than 0.2 ºC, the precision of the DTS equipment. As a result, the diurnal signal cannot be easily filtered from the noise. We suggest in general that signal components with amplitudes below the sensor precision should be used with caution. The filtered components from 0.51 m depth are shown in Figure 9.

3.4 Sample-by-sample flux calculation

After filtering the temperature time series with DHR, it would certainly be possible to identify daily maxima and minima and thereby calculate daily amplitudes and phase angles; however, another significant advantage of using DHR is that it allows VFLUX to calculate amplitude and phase at each sub-daily time step in the resampled data (by Equations 3 and 4), rather than estimating them daily. There are several benefits to calculating flux at the sub-daily time scale.

We have already discussed how DHR avoids the subjective human identification of extrema, but sample-by-sample calculations also have the benefit of being a more precise-in-time estimate of the character of a time series. Note how in Figures 7 and 9, the amplitude of the filtered diurnal signal changes with time. We know from the DHR model (Equations 1 and 2) that the amplitude of a single-frequency component is time-varying. When an estimate of amplitude is calculated as half of the difference between a daily maximum and minimum, the amplitude is effectively averaged over the time period between the extrema; however, with

DHR, the amplitude and phase of each frequency component is determined at each sample, which represents a period of time only as long as the sampling period (two-to-four hours in VFLUX). By employing DHR, VFLUX can calculate flux rates between a sensor pair from the amplitude attenuation and time lag at any sample in the time series, and it thereby avoids relying heavily on the time and temperature of only two measurements per day, i.e., the daily extrema. The high temporal resolution of sample-by-sample flux calculations is most useful during days with unusual weather patterns, when it may provide flux results for fractions of the day, while the daily extrema method may fail to give any flux measurement at all.

Sub-daily flux calculations might be needed in systems where flux rates change quickly, such as in tidal estuaries, dam-controlled reservoirs or rivers, or in flashy streams during discrete precipitation events. However, it should be noted that, to our knowledge, no research has been performed on the applicability of the Hatch or Keery analytical methods to systems where flux rates change significantly at sub-daily scales. Keery et al. (2007) did note that the Stallman equation is based on the assumption that flux does not change at scales shorter than the diurnal period, and "it is therefore neither necessary nor appropriate to calculate vertical flux for every time step." However, VFLUX would be an ideal tool in any field or modeling investigation of the Hatch or Keery method in dynamic-flux environments.

One clear value of making many independent estimates of a daily flux rate is that the mean or median of such estimates, which have inherent noise, may best estimate the value in question. We recommend that estimates of physical vertical water flux at a field site be made by finding the central tendency of several days of flux calculations (as in Lautz, 2010). Shown in Figure 10 are the amplitude and phase angle of a pair of sensors from Ninemile Creek through time, and the flux rates that were calculated by VFLUX at each time step. The slight differences

between flux calculated with the Keery and Hatch methods in Figure 10c are due to the differing treatment of thermal dispersion by the two models (see the discussion in Keery et al., 2007).

3.5 Sensor pair identification and ideal sensor spacing

The Hatch and Keery methods both require concurrent information from two vertically-spaced temperature sensors to make a single determination of vertical water flux. Choosing the best spacing between the sensors (Δz) requires balancing several considerations. The flux calculated for a single sensor pair is an estimate of the volume of water per area moving vertically over the distance Δz, or the average vertical flux over Δz. VFLUX assigns the flux estimate to a point equidistant between the sensors, referred to as the center-of-pair depth; however, as Δz increases, so does the averaging distance represented by that single point. A long averaging distance might not be appropriate if the vertical component of flux changes along the distance Δz. If the purpose of an experiment is to observe changes in vertical flux with depth, or if it is critical that a flux estimate represent a certain depth with precision, then the goal should be to have the smallest sensor spacing (window) possible.

On the other hand, the amplitude attenuation and time lag between two sensors must be large enough to be resolved by the resolution and response time of the sensors (Hatch et al., 2006). If the sensors are too close together, then the amplitude ratio can effectively approach one (or even surpass one due to measurement error), and the time lag can likewise approach zero. If the sensors are too far apart, then the tracer signal may be damped out below the sensor resolution at the lower depth, and the amplitude ratio can be effectively zero. If sensor spacing is two large, vertical resolution is also lost. Furthermore, the analytical heat transport models are most sensitive within defined ranges of A_r and Δt that vary depending on the magnitude and direction of flux and the thermal properties of the streambed sediment (Hatch et al., 2006; Keery

et al., 2007). For example, the Hatch amplitude method loses sensitivity as A_r approaches one or zero, near which small changes in the amplitude ratio can lead to large changes in the calculated flux due to a very low dA_r/dq. Selection of the best arrangement of sensors in a TP is therefore a particular problem, because the flux rates may not be known prior to equipment installation, or the flux rates may change with time or depth. Therefore, for the most flexible analysis Hatch et al. (2006) recommend the deployment of many sensors, so as to have the greatest number of pairs with the widest range of Δz values possible. However, the large number of potential pairs (1596 in the case of Cherry Creek) leads to a very large number of calculations.

As described in Methods, VFLUX has several features that allow the user to determine an ideal sensor spacing range. For any time step where the amplitude ratio is 1 or 0 (the sensor spacing is too small or two large), the program will not calculate either of the amplitude methods, and a null value is written. Likewise, for any time step where the time lag is 0 or negative, the program will not calculate either of the phase methods, and a null value is written. The application of the Hatch amplitude method in the VFLUX code also includes an optional sensitivity check, which will write a null flux value if the derivative dA_r/dq is below a certain value (0.001 d m^{-1}, by default), which indicates very poor sensitivity of the model due to non-ideal sensor spacing. The selection of an ideal window size for a specific application therefore involves achieving a sufficient percentage of non-null flux values while still maintaining a small-enough window to supply sufficient vertical resolution in space.

VFLUX displays bar charts of the percent of non-null flux calculations for different sensor pairs and sensor spacings (Figure 11). For each window size, a separate bar chart shows the percent of successful calculations made at each center-of-pair depth, and the number of bars and their spacing indicates the resolution in the vertical dimension (i.e., spatial resolution is

higher with smaller window sizes and a greater number of closely spaced depths). In relatively high-flux environments, such as the examples from Cherry and Ninemile Creeks, the ideal sensor spacing is larger than for low-flux environments; therefore, at small windows sizes, a large number of flux calculations are outside the model sensitivity range. However, because vertical flux changes with depth at both sites, we want high spatial resolution to resolve these changes, and larger window sizes may obscure or blur transitions in flux with depth.

At Ninemile Creek, amplitude ratios were generally high (ranging from 0.716 to 1.01 for all pairs and time steps), indicating that water flux was strong and in the downward direction. Because there are only 21 possible pairs in a TP of seven sensors, every possible window was used in VFLUX, i.e., 1, 2, 3, 4, 5, and 6 sensor spacings, which correspond to Δz values of 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30 m. At a Δz of 0.05 m, 17 percent of time steps had an $A_r \geq 1$ and the mean A_r was 0.965; at a Δz of 0.10 m, 9 percent of time steps had a $A_r \geq 1$ and the mean A_r was 0.944; by a Δz of 0.20 m, there were no time steps at which flux could not be calculated and the mean A_r was 0.903. The bar charts in Figure 11 break down these statistics for each window by depth. With a window of 1 sensor spacing (Δz of 0.05 m), Figure 11 shows that fewer than 35 percent of the flux calculations were within the model sensitivity range at the shallowest center-of-pair depth (0.025 m), where vertical flux was highest. A substantial percentage of calculations were also null at 0.075 m depth. With a window of 2 (Δz of 0.10 m), the shallowest depth of 0.05 m has an approximately 60 percent success rate, and with a window of 3 (Δz of 0.15 m), the shallowest depth is now up to 90 percent. We made the subjective decision that a Δz of 0.10 m best balanced spatial resolution and model sensitivity, because more than half of flux values were calculated within the model sensitivity range at even the shallowest depth (with

the highest vertical flux rate), and five different center-of-pair depths gave enough vertical resolution to describe changing flux patterns with depth.

As stated in Methods, the Cherry Creek DTS data were processed with windows of 8, 9, 10, 11, and 12 sensor spacings, which correspond to Δz values of 0.110, 0.124, 0.138, 0.152, and 0.166 m. These spacings were chosen because the amplitude ratios observed for this rod were generally high and lag times were short, meaning that the water flux at the TP location was strongly downwelling, and large sensor spacings were required to detect a change in amplitude. The pairs separated by the smaller Δz values of 0.110 and 0.124 m had many time steps at which flux could not be computed because the amplitude ratio was close to or greater than one. The values above one are due to slight errors in measurement or digital filtering. At Δz of 0.110 m, 21 percent of time steps for pairs with a center-of-pair depth less than 0.20 m had an $A_r \geq 1$, and the mean A_r among these time steps was 0.963; at a Δz of 0.138 m, 18 percent of time steps had an $A_r \geq 1$ and the mean A_r was 0.954; and at a Δz of 0.166 m, only 12 percent of time steps shallower than 0.20 m had an $A_r \geq 1$ and the mean A_r was 0.943. We made the subjective decision that a Δz of 0.138 m gave the best balance between vertical spatial resolution and model sensitivity for the entire TP, although it would also be possible to use a combination of multiple Δz values at different times and/or depths to maximize the sensitivity of the results in both space and flux rate (see Briggs et al., in revision.).

These examples show how the deployment of many sensors in a TP, coupled with the computing flexibility of a program like VFLUX, can help identify the best sensor spacing for a particular experiment and field site. In the cases of Cherry Creek and Ninemile Creek, high amplitude ratios due to high downward flux rates made the choice of sensor spacing even more important.

3.6 Flux results and a comparison of methods

The results of flux calculations at Ninemile and Cherry Creeks show that vertical flux rates at both sites are indeed strongly downwelling (approximately $5x10^{-5}$ m s^{-1} and $2x10^{-5}$ m s^{-1}, respectively, at the shallowest depths). A box plot of 2-hour flux estimates over the entire collection period (excluding the first and last 48 hours) versus the center-of-pair depth from Ninemile Creek is shown in Figure 12a. A similar plot from Cherry Creek is shown in Figure 12b. These two figures only include flux estimates calculated with the Hatch amplitude ratio method (Equation 6). As can be seen in the box plots, the distribution of flux estimates for each sensor pair is not normally distributed but has a positive skew with large positive outliers. The outliers are a result of decreasing sensitivity of the amplitude ratio model as the amplitude ratios become very close to 1, and could have been removed by employing the Hatch amplitude method derivative check (see the discussion above). In Figure 12, we have chosen to keep all outliers in order to illustrate the range of potential results, and we therefore use the median to represent the central tendency of a population of flux estimates.

At both sites, the vertical flux rates at the shallowest depths are the largest (Figure 12), with a median rate over time of $5.1x10^{-5}$ ($+1.4x10^{-5}$ or $-7.2x10^{-6}$ m s^{-1}) at 0.025 m depth at Ninemile Creek and $1.7x10^{-5}$ ($+5.5x10^{-6}$ or $-1.9x10^{-6}$ m s^{-1}) at 0.097 m depth at Cherry Creek (the plus-or-minus errors span the interquartile range). As the depth increases away from the streambed interface, the flux rate decreases towards zero. We interpret these spatial patterns to be due to relatively shallow, curved hyporheic flow paths that are not purely vertical but have a horizontal component that increases with increasing depth (also see Briggs et al., in revision). The profile at Ninemile Creek was installed at the head of a riffle in a pool-riffle sequence, where we would expect the initiation of a shallow, curved hyporheic flow cell. The profile at

Cherry Creek was installed just upstream of a beaver dam, which creates a step in hydraulic heat and likely initiates similar shallow hyporheic flow that becomes horizontal with depth. The flux rates near the streambed interface are therefore likely to be the best estimate of exchange between the hyporheic zone and the stream itself. Compared to previous methods in which the sensor spacing increased with increasing depth (Vogt et al., 2010), the use of a sliding window in the VFLUX method makes it possible to identify with greater precision changes in flux with depth, because it is easy to use the smallest sensor spacing—with the greatest spatial precision—that is still sensitive to the heat-transport model in use (Briggs et al., in revision).

Figure 13 shows calculated vertical flux rates through time for a single pair of sensors at 0.15 and 0.25 m depth (center-of pair depth is 0.20 m) at Ninemile Creek, calculated with the Hatch amplitude method. Also plotted are two sets of flux rates that were calculated on a daily basis by selecting the maximum and minimum temperature in each 24-hour period (similar to the methods of Fanelli and Lautz, 2008; Hatch et al., 2006; Lautz, 2010). The daily flux calculations were made on both the raw, unfiltered time series and on the diurnal signal after filtering it with DHR. Although the true flux values in the field are not known independently from our heat transport estimates, it is still useful to compare the estimates made from both the sample-by-sample and daily methods using both filtered and unfiltered thermal data, since all of these techniques have been used in the published literature (e.g. Fanelli and Lautz, 2008; Hatch et al., 2006; Keery et al., 2007). The daily flux calculations for the filtered time series agree fairly well with the results from VFLUX, except in the last five days, where they are slightly lower than the flux rates from VFLUX. The median flux value over the whole collection period is not significantly different between the two methods (Mann-Whitney rank test, $p=0.718$). However, the daily flux calculations made with the unfiltered time series do not agree well with the results

from VFLUX in Figure 13. The unfiltered flux estimates are more variable in time than both of the filtered flux estimates, and the median flux value is significantly different between the two methods at the 5-percent significance level (p=0.029).

3.7 Sensitivity and uncertainty analysis for thermal properties

We employed the VFLUX programs "vfluxsens" and "vfluxmc" to conduct basic sensitivity analyses on sediment and thermal properties and to estimate confidence intervals for flux estimates through time for the data from Ninemile Creek. The results of the sensitivity analysis for a Δz of 0.1 m at a center-of-pair depth of 0.2 m are presented in Table 2. At this magnitude of downward flux (1.03×10^{-5} m s^{-1} calculated with the base parameters) estimated flux values are most sensitive to thermal conductivity, closely followed by porosity, which determines the total heat capacity of the saturated matrix. Flux is least sensitive to the heat capacity of sediment and water, because these values vary little in nature (Lapham, 1989). Although only results from one pair of sensors are presented, these general patterns hold true for all of the flux results from Ninemile Creek. As the magnitude of downward flux increases, flux becomes more sensitive to porosity and thermal conductivity. A similar sensitivity analysis was performed for Cherry Creek (not shown), where flux magnitudes were smaller than at Ninemile. At flux rates above approximately 5×10^{-6} m s^{-1}, the Cherry Creek analysis shows a similar pattern to the Ninemile analysis, with increasing model sensitivity to porosity and thermal conductivity with increasing downward flux. However, at flux rates below 5×10^{-6} m s^{-1}, the model is more sensitive to thermal conductivity with decreasing flux rates, suggesting that model sensitivity to conductivity is at a minimum near 5×10^{-6} m s^{-1} (for the specific conditions at Cherry Creek). At Ninemile Creek, flux magnitudes were not small enough to observe this effect.

One question that arises when comparing different flux estimates from one site to another, or even from one depth to another, is whether differences in flux estimates are due to truly different magnitudes of water flux, or simply to different sediment thermal properties. For example, in order to say with greater certainty that vertical flux does indeed decrease with depth at Ninemile Creek (Figure 12a), we can test whether the apparent differences in flux could be due to changes in sediment texture alone. Select results from the Monte Carlo analysis of data from Ninemile Creek are shown in Figure 14. Flux estimates are plotted through time for two sensor pairs with center-of-pair depths at 0.15 and 0.20 m (Δz of 0.10 m), with upper and lower uncertainty limits (±2 standard deviations). The plot shows that uncertainty is greater at greater downward flux rates (at the shallower depth), but there is no overlap between the confidence intervals from the two depths. Results therefore suggest that the different flux rates measured at 0.15 and 0.20 m are statistically different. Error or uncertainty analyses such as the one presented here are an important component of physical flux estimation, and confidence intervals should generally be given along with reported flux values. The VFLUX software package and the Monte Carlo analysis program demonstrated here make the calculation of confidence intervals automatic and straightforward.

4. Conclusion

Time series of streambed temperatures have been successfully used by many researchers to estimate vertical water flux rates in shallow streambeds (e.g. Fanelli and Lautz, 2008; Hatch et al., 2006; Keery et al., 2007; Vogt et al., 2010), and these studies have each made advances in heat-transport modeling, signal processing, and experimental deployment. We have introduced a step-by-step workflow for analyzing temperature time series and calculating vertical water flux, which incorporates the advancements made in previous work and adds automation and new

techniques for processing large amounts of data from high-resolution sensor profiles. The method has been written into a computer program called VFLUX, which automates the entire process of calculating vertical flux rates from raw temperature time series. VFLUX is the first published computer program for calculating water flux from temperature time series to completely automate data pre-processing, including robust filtering with Dynamic Harmonic Regression (DHR), and the first method to use a sliding window to identify changing flux rates with depth at high spatial resolution. VFLUX also includes functions for data and result visualization, new tools for evaluating model sensitivity and ideal sensor spacing, and functions for sensitivity analysis and confidence interval calculation.

We have shown using two field examples that each step of the method is important for calculating accurate flux rates from real-world temperature data, which are typically characterized by several irregular signals that vary in time with different amounts of noise. The automated method processes time series consistently for every sensor in a profile, and it can calculate flux between every possible sensor pair separated by every possible sensor spacing, which is particularly useful when the ideal sensor spacing for a specific field site is unknown. An automated program such as VFLUX is of great assistance when deploying a large number of sensors, as in a high-resolution DTS profile, due to the very large number of potential calculations. We have shown how VFLUX can help the practitioner quantitatively evaluate the ideal spacing between sensor pairs, and how the program can be used in situations where many repeated calculations are desired, such as studies of model sensitivity to thermal parameters and Monte Carlo error estimation.

We have also shown how this method, which provides high spatial and temporal resolution, is both more accurate and reliable than previous methods at capturing changing flux

rates through time and changing vertical flux with depth. Our analysis of field data from Ninemile Creek, NY, and Cherry Creek, WY demonstrate how the VFLUX method can be used in situations where shallow hyporheic flow paths change direction from vertical near the surface to horizontal with depth, such as in shallow, curved hyporheic flow paths. For example, high-resolution DTS data from Cherry Creek processed with this method show vertical flux decreasing from approximately 1.7×10^{-5} m s^{-1} very near the sediment-water interface to near zero by approximately 0.5 m depth. A Monte Carlo uncertainty analysis at Ninemile Creek shows that the observed decrease in flux with depth cannot be due to changing sediment thermal properties alone. Unlike other methods used in the literature, VFLUX also supplies flux results at high temporal resolution, and therefore could also be used to investigate flux rates that change on hourly time scales in estuary environments or dam-controlled water bodies. We believe that the release of this method and the availability of the VFLUX program will make heat transport modeling a more approachable and widely-used technique for calculating vertical water flux rates using temperature.

Acknowledgements

Field data were collected in collaboration with Timothy Daniluk and Danielle Hare. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-0750965, and is partially based upon work supported by the National Science Foundation under Grant Nos. EAR-0911612 and EAR-0901480 and the Canadian Foundation for Innovation.

Tables

Table 1. Values of the physical parameters used to calculate flux rates at Ninemile Creek, NY and Cherry Creek, WY. Values listed in table are also the base or mean values used in the sensitivity and Monte Carlo analyses for Ninemile Creek. The standard deviation used in the Monte Carlo analysis is shown in parentheses. Thermal properties were estimated based on guidelines in Lapham (1989) and values in Fetter (2001) and CRC (2011).

Property	Symbol	Units	Value	
			Ninemile Creek	Cherry Creek
Total porosity	n	dimensionless	0.20 (0.04)[a]	0.35
Baseline thermal conductivity	λ_o	J s^{-1} m^{-1} $^oC^{-1}$	2.26 (0.31) [b]	1.30
Thermal dispersivity	β	m	0.001 (0) [c]	0.001
Volumetric heat capacity of sediment	C_s	J m^{-3} $^oC^{-1}$	$2.09x10^6$ ($3.1x10^4$) [d]	$2.09x10^6$
Volumetric heat capacity of water	C_w	J m^{-3} $^oC^{-1}$	$4.18x10^6$ ($2.1x10^4$) [e]	$4.18x10^6$

[a] Low and high n values for sensitivity analysis are 0.18 and 0.34, respectively.
[b] Low and high λ_o values for sensitivity analysis are 1.21 and 2.47 J s^{-1} m^{-1} $^oC^{-1}$, respectively.
[c] Low and high β values for sensitivity analysis are 0 and 0.1 m, respectively.
[d] Low and high C_s values for sensitivity analysis are $2.03x10^6$ and $2.15x10^6$ J m^{-3} $^oC^{-1}$, respectively.
[e] Low and high C_w values for sensitivity analysis are $4.14x10^6$ and $4.23x10^6$ J m^{-3} $^oC^{-1}$, respectively.

Table 2. Results of the sensitivity analysis for Ninemile Creek data (Δz of 0.1 m, center-of-pair depth of 0.2 m).

Property varied	Median flux value through time ($m\ s^{-1}$)	
	with low parameter value	with high parameter value
[none, using base values from Table 1]	1.03×10^{-5}	
Total porosity	1.01×10^{-5}	1.11×10^{-5}
Baseline thermal conductivity	0.86×10^{-5}	1.05×10^{-5}
Thermal dispersivity	1.02×10^{-5}	1.07×10^{-5}
Volumetric heat capacity of sediment	1.02×10^{-5}	1.04×10^{-5}
Volumetric heat capacity of water	1.034×10^{-5}	1.018×10^{-5}

Figures

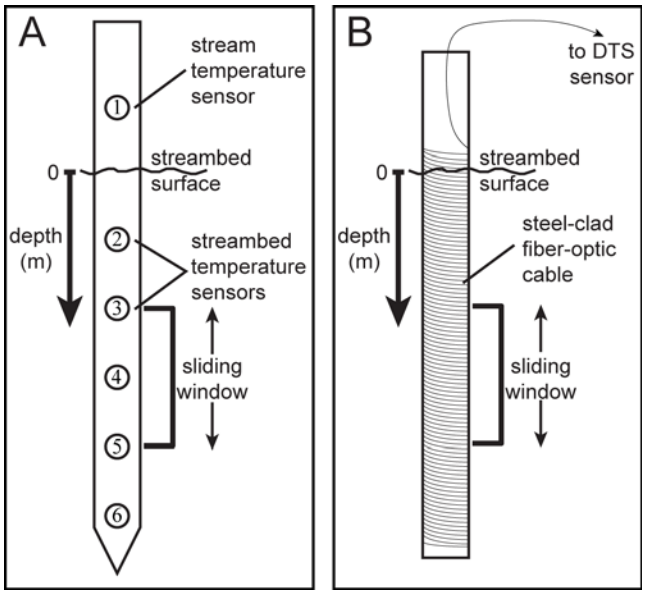

Figure 1. (a) A diagram of a typical vertical temperature profile (TP), showing five iButton Thermochron temperature sensors installed in the streambed and one in the water column, which gives the temperature of the water at the streambed interface. The sensors are not necessarily evenly-spaced. The sliding window (in this case, two sensor-spacings long) identifies sensor pairs between which flux will be calculated. (b) A diagram of a high-resolution DTS temperature profile, showing a coil of fiber-optic cable wrapped around a threaded PVC rod. Using a distributed temperature sensing device, the temperature of short segments (0.024 m) of the rod length can be measured, which are equivalent to temperature sensors. The sliding window again identifies sensor pairs for flux calculations.

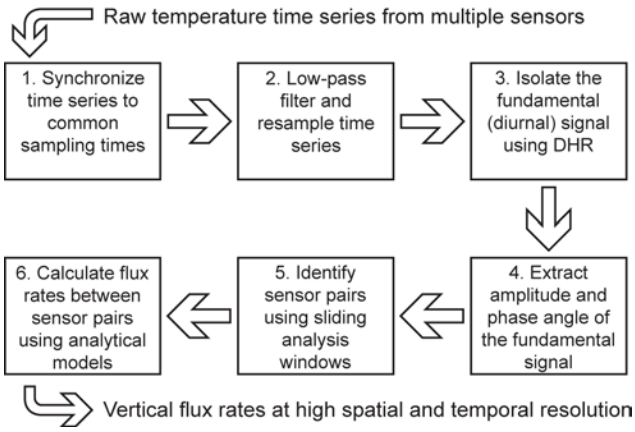

Figure 2. A flowchart of the six major steps in the VFLUX method.

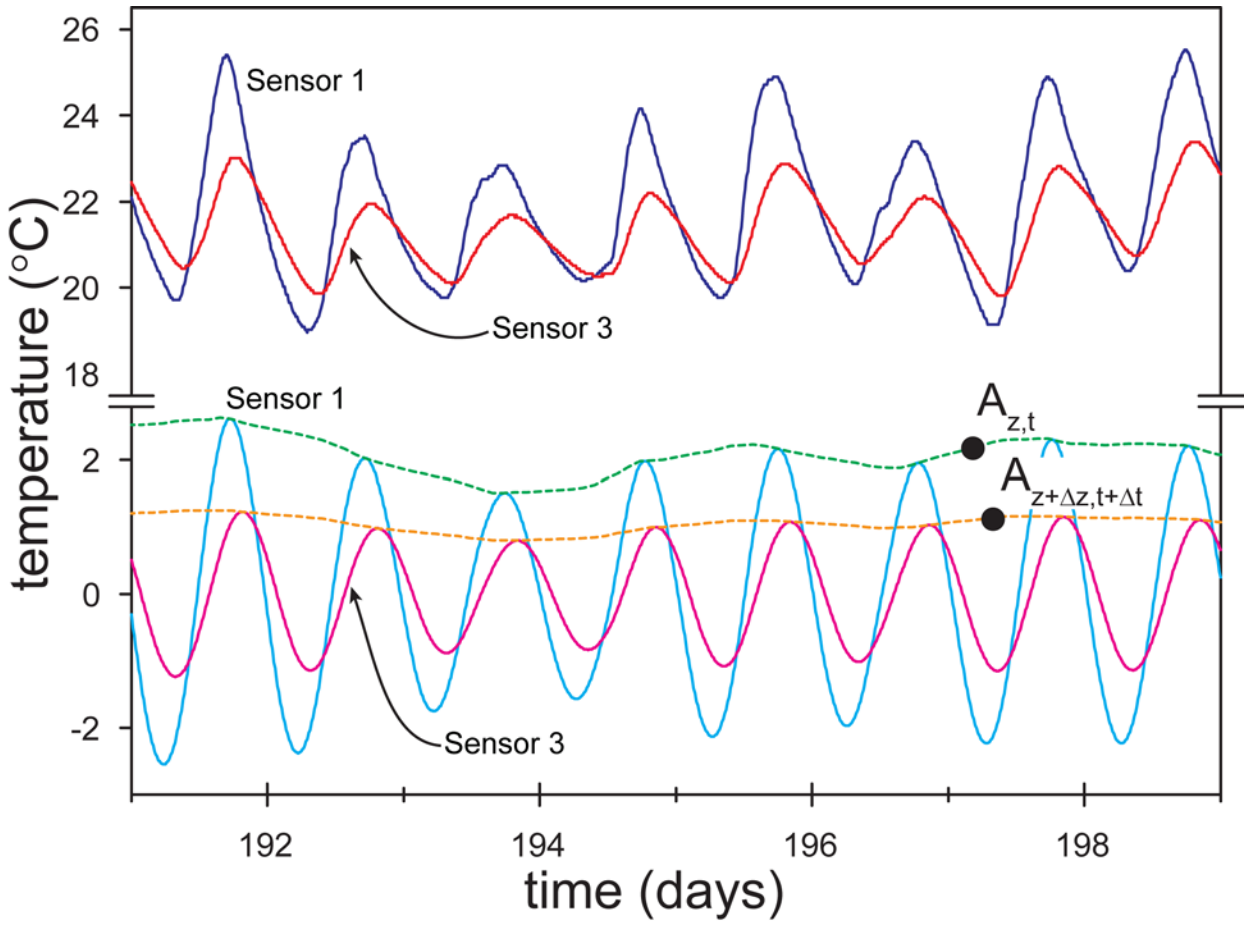

Figure 3. Two discrete time series from sensors 1 and 3 in Figure 1a. The raw temperature time series are shown in the upper part of the plot. In the lower part are shown the diurnal frequency components of the two time series, extracted using Dynamic Harmonic Regression (DHR). The dashed lines are the amplitudes of the diurnal components, extracted using DHR. The pair of points illustrates the time lag and amplitude attenuation between equivalent parts of the temperature signal as it progresses through different depths.

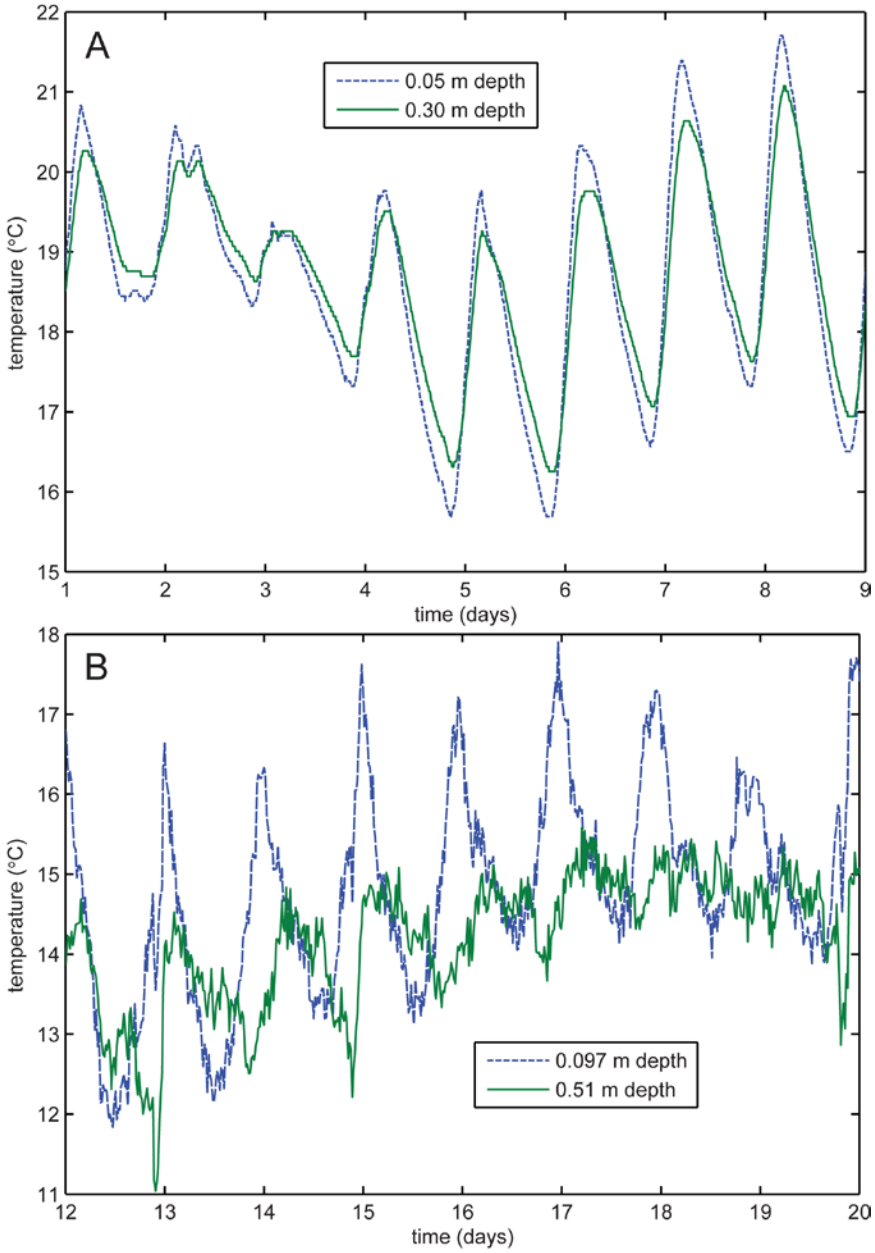

Figure 4. Examples of raw temperature time series data from (a) Ninemile Creek at 0.05 m and 0.30 m depth and (b) Cherry Creek at 0.097 m and 0.51 m depth. Note in both plots the amplitude attenuation and time lag between the shallower and deeper sensors, the asymmetric saw-toothed pattern of daily increases and decreases in temperature, the changing trend through time, and the amount of noise present in each time series.

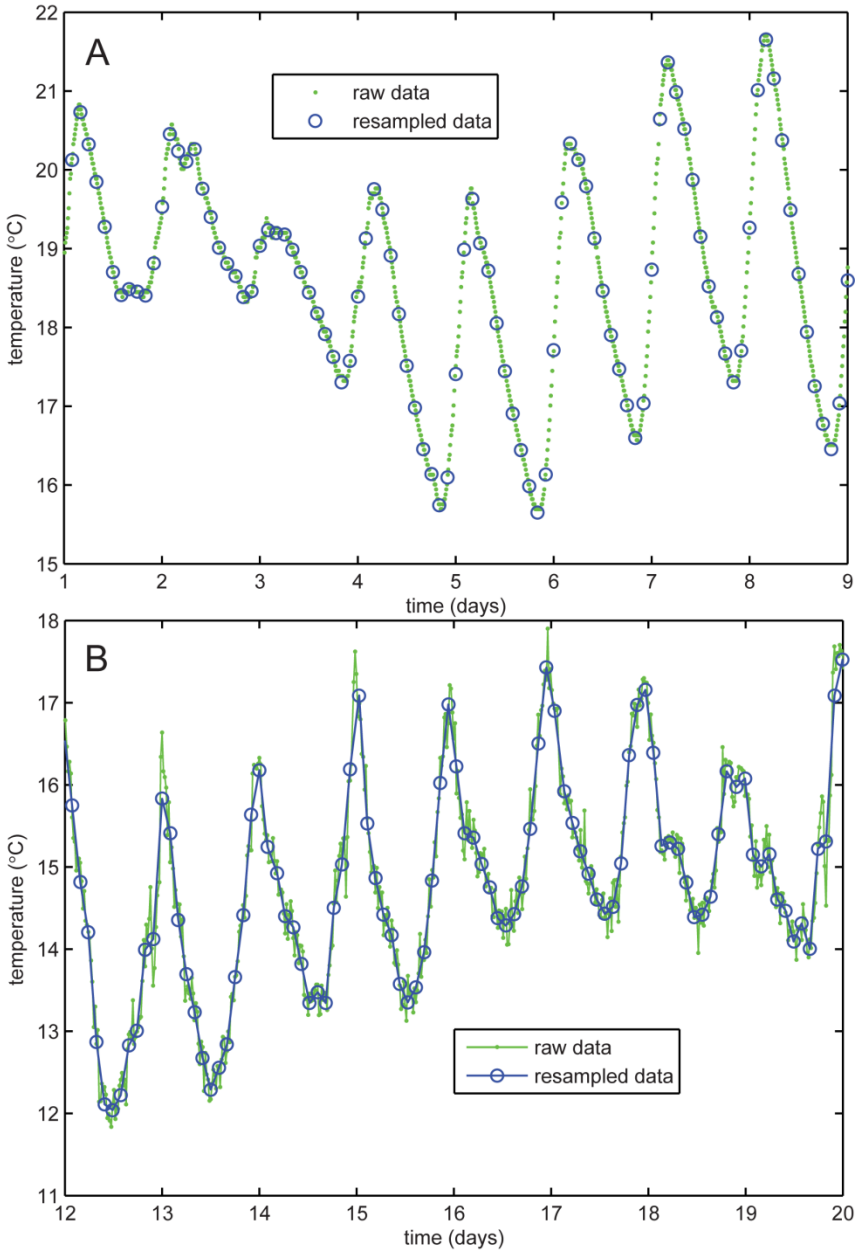

Figure 5. Examples of temperature time series data before and after being low-pass filtered and resampled. (a) From Ninemile Creek at 0.05 m depth. (b) From Cherry Creek at 0.097 m depth. The low-pass filter has visibly reduced the high-frequency noise in the Cherry Creek data.

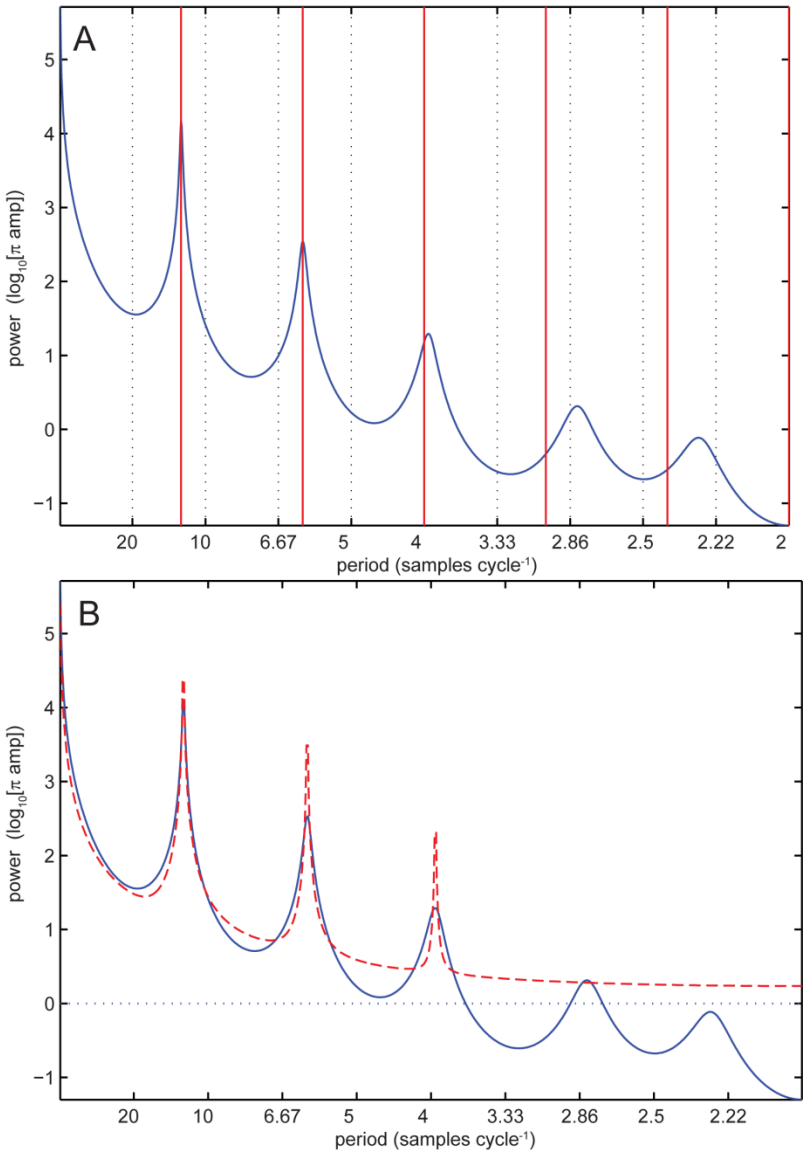

Figure 6. (a) The AR(12) spectrum (auto-regression spectrum of order 12) of the 0.05 m time series from Ninemile Creek, showing peaks near the fundamental frequency (12 samples per cycle) and the first and second harmonics (6 and 4 samples per cycle). The slope of the spectrum up and left towards the zero frequency (infinite period) is the trend component of the DHR model. The vertical lines are indices showing the expected periods of the fundamental signal and its harmonics (12, 6, 4, 3 and 2.4 samples per cycle). (b) The optimized DHR model for the AR spectrum in (a) (dashed line), showing a good model fit to the data (solid line) along the trend, fundamental, and first two harmonic peaks.

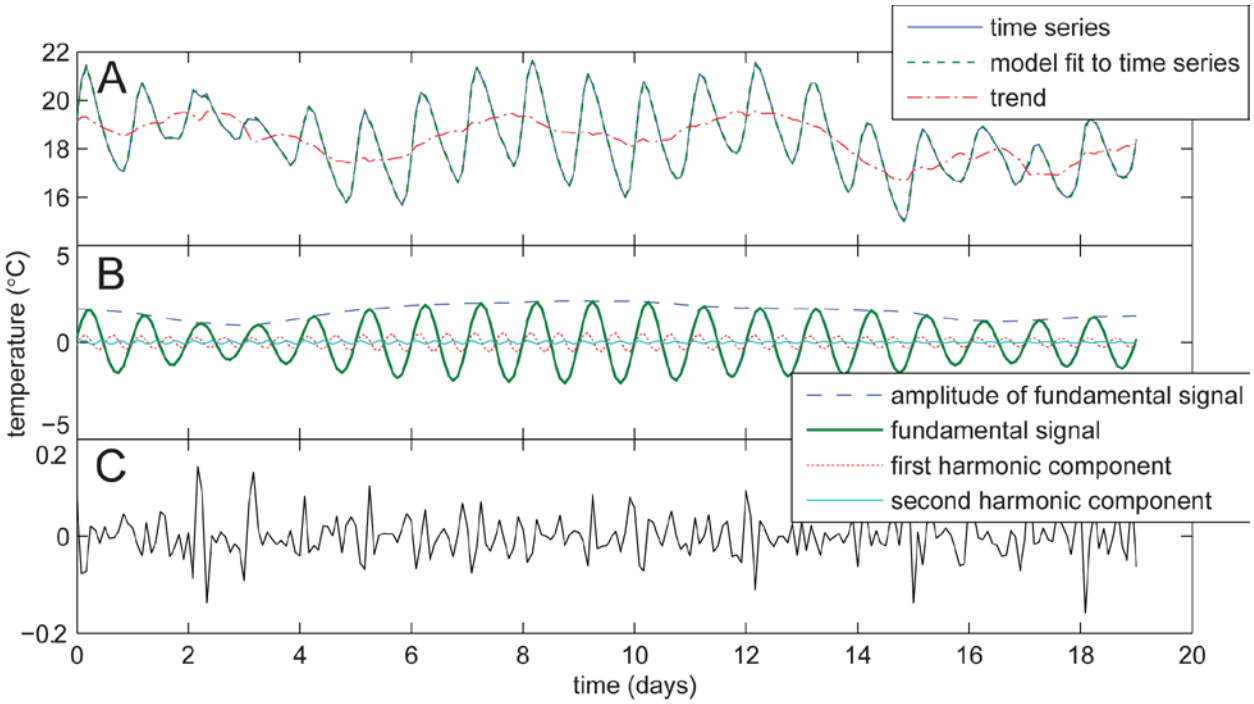

Figure 7. Results of DHR filtering of the 0.05 m time series from Ninemile Creek. (a) The original resampled time series, the DHR model fit to the time series in the time domain, and the DHR-identified trend. (b) The fundamental (diurnal) frequency and its amplitude, and the first two harmonics. (c) The residual, or the difference between the original data and the model fit.

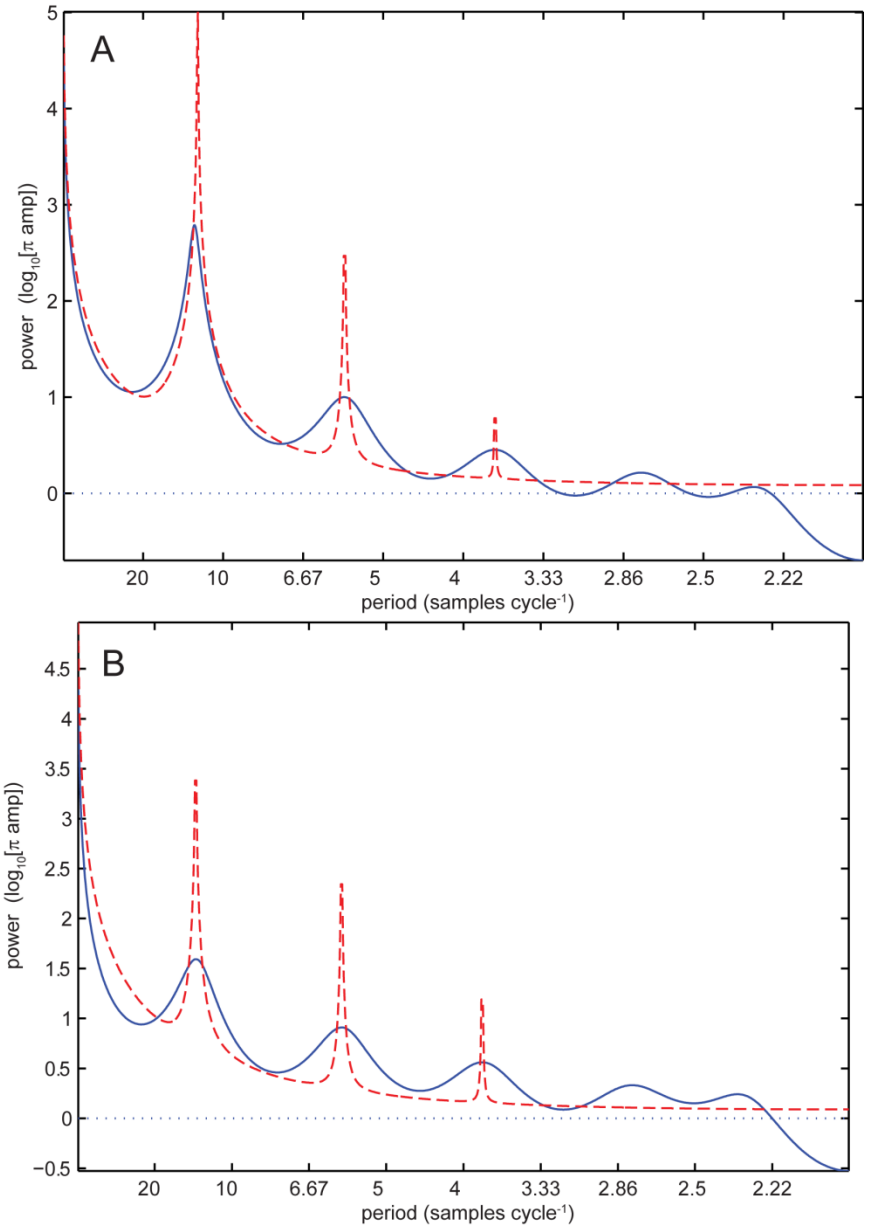

Figure 8. (a) The AR(12) spectrum of the 0.097 m time series from Cherry Creek (solid line), with the optimized DHR model fit (dashed line). (b) The AR(12) spectrum of the 0.51 m time series from Cherry Creek (solid line), with the optimized DHR model fit (dashed line).

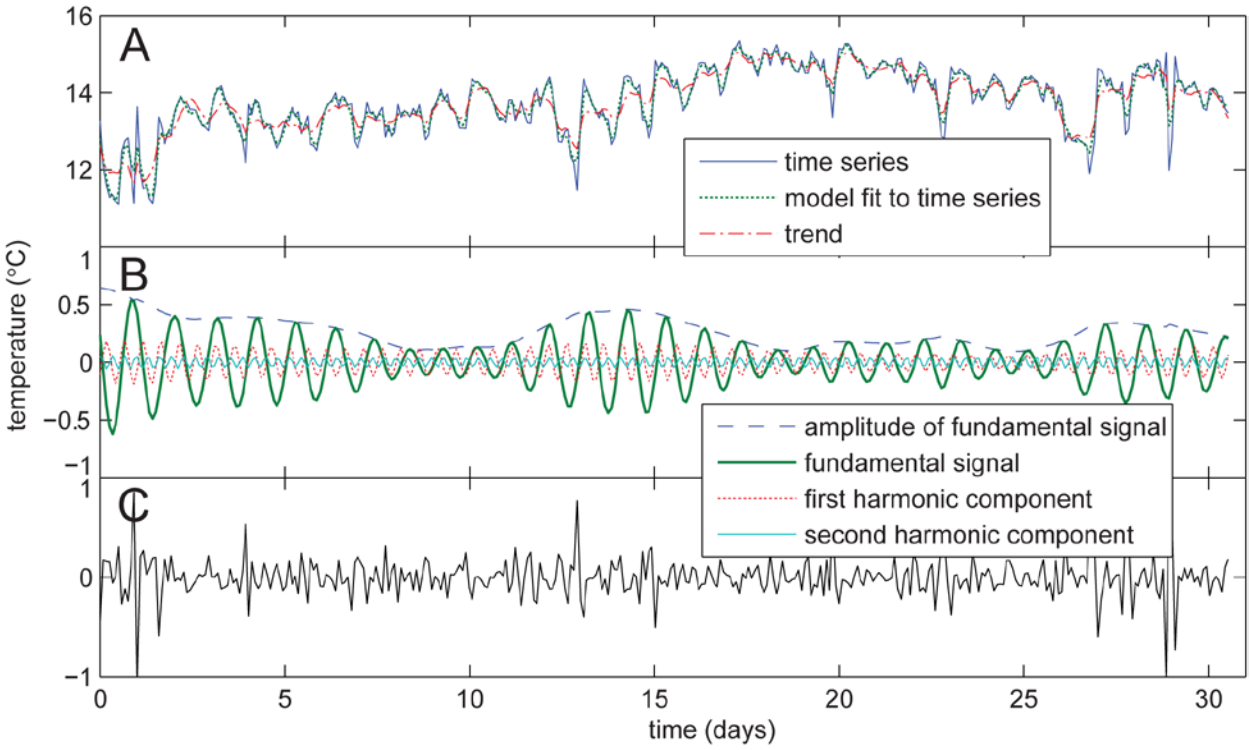

Figure 9. Results of DHR filtering of the 0.51 m time series from Cherry Creek. The fundamental component and its amplitude are identified, but the model fit to the raw data is not ideal, with some large spikes in the residual. Note that the average amplitude of the diurnal signal is approximately 0.27 ºC, only slightly greater than the ±0.2 ºC precision of the DTS equipment. (a) The original resampled time series, the DHR model fit to the time series in the time domain, and the DHR-identified trend. (b) The fundamental (diurnal) frequency and its amplitude, and the first two harmonics. (c) The residual, or the difference between the original data and the model fit.

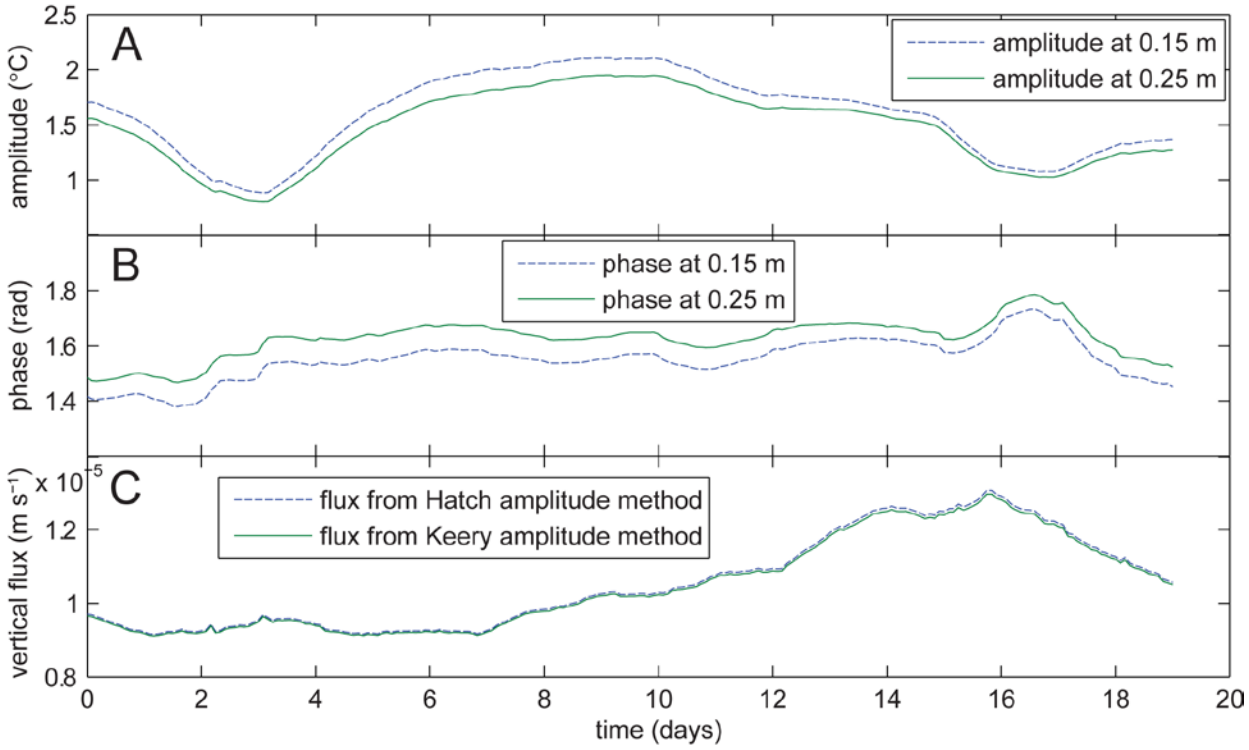

Figure 10. (a) A plot of the amplitude of the diurnal signal through time from sensors at Ninemile Creek at 0.15 and 0.25 m depth. (b) A plot of the phase angle of the diurnal signal from the same two sensors. (c) Flux calculated using the Hatch and Keery amplitude methods for the pair of sensors in (a) at a center-of-pair depth of 0.20 m.

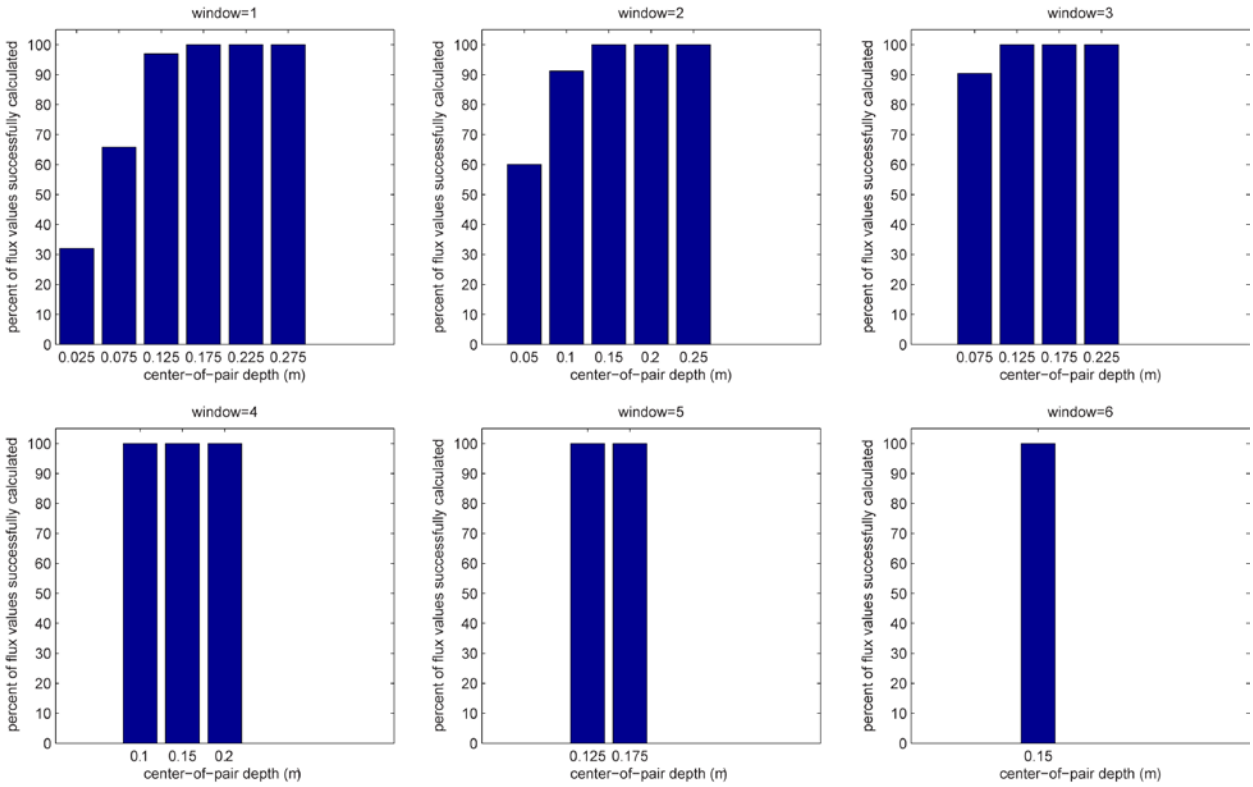

Figure 11. An example of the sensor spacing diagnostic charts produced by the visualization routine of VFLUX. This set of charts was produced using the Hatch amplitude method results from the Ninemile Creek data. Each bar chart for a specific window size (sensor spacing) shows the percent of flux values through time that were within the sensitivity range of the model at each center-of-pair depth.

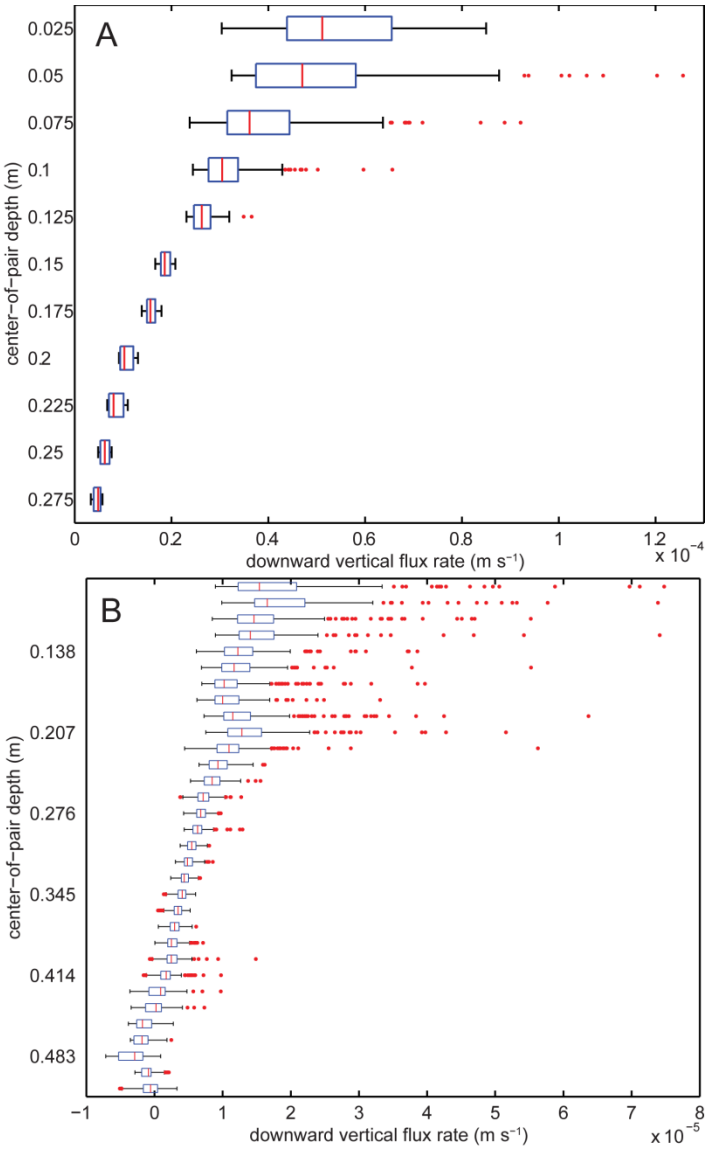

Figure 12. Box plots of 2-hour flux estimates over the entire collection period (excluding the first and last 48 hours), showing the decrease in vertical flux with depth at both study locations. The depth value of each box is the center-of-pair depth. The boxes represents the interquartile range and the vertical lines the median value. The whiskers extend to the most extreme data value that is beyond the boxes by less than 1.5 times the interquartile range, and any outliers beyond the whiskers are plotted individually as dots. (a) At Ninemile Creek, only showing sensor pairs where Δz is 0.05 or 0.10 m. (b) At Cherry Creek, only showing sensor pairs with a Δz of 0.138 m and with a center-of-pair depth of 0.51 m or less.

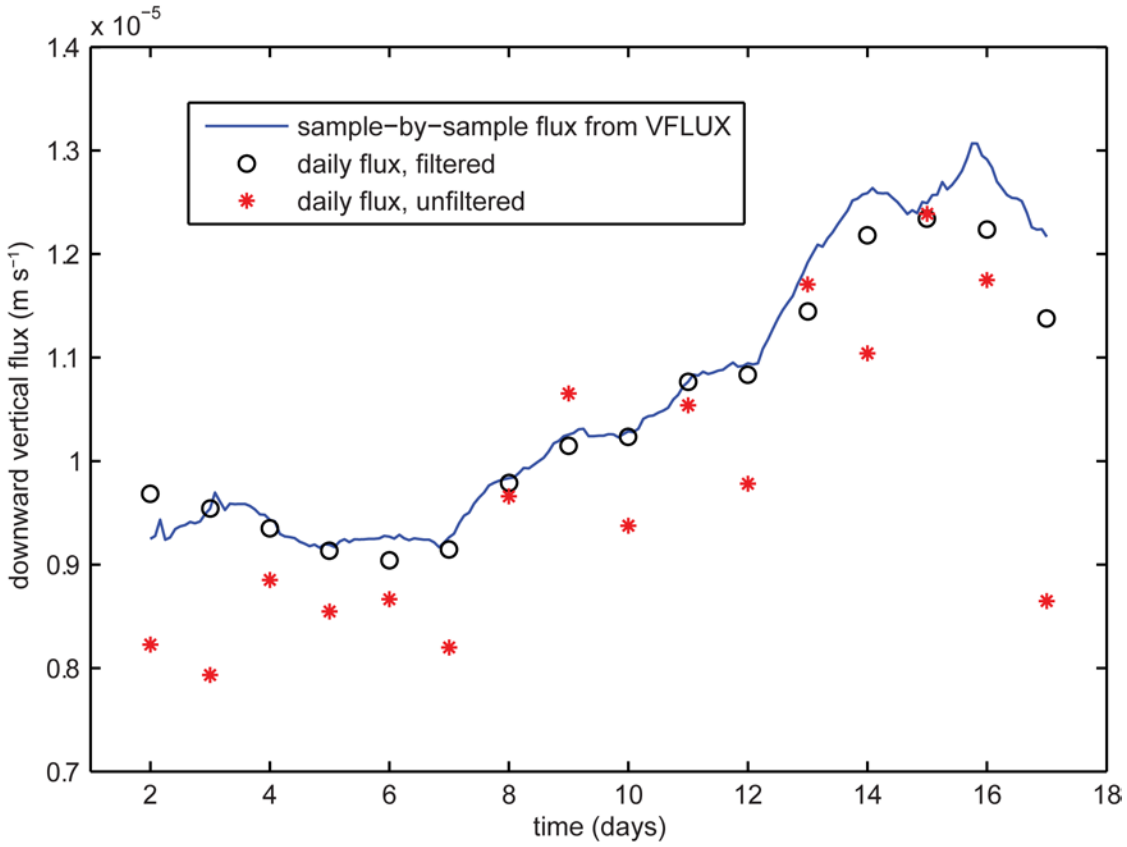

Figure 13. Vertical flux rates through time calculated for the pair of sensors at 0.15 and 0.25 m depth at Ninemile Creek. Flux rates are calculated with the VFLUX program, and also using the daily maxima and minima of the raw temperature time series and the time series after filtering with DHR.

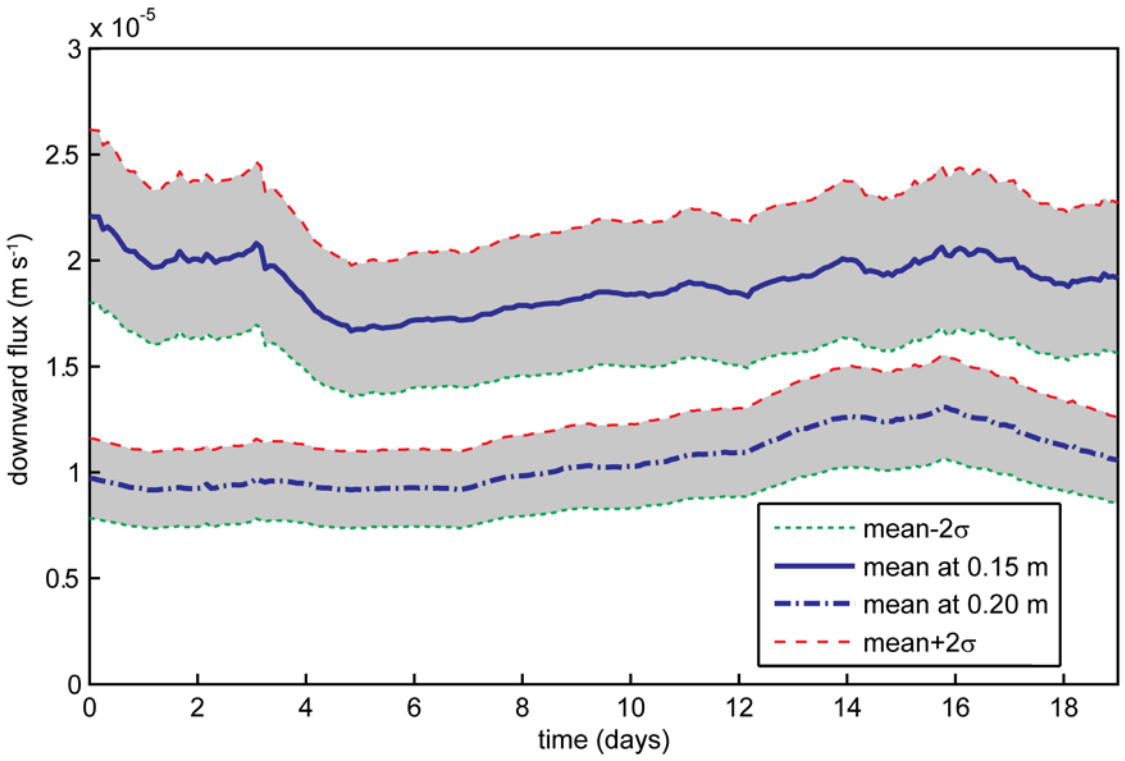

Figure 14. Results from the Monte Carlo analysis of flux uncertainty for two sensor pairs with center-of-pair depths at 0.15 and 0.20 m (Δz of 0.10 m). The upper and lower uncertainty limits represent the mean flux estimate ±2 standard deviations.

Chapter 2.

Spatial patterns of hyporheic exchange and biogeochemical cycling around cross-vane restoration structures: Implications for stream restoration design

Chapter 2 has been published as:

Gordon, R.P., L.K. Lautz, and T.L. Daniluk, 2013, Spatial patterns of hyporheic exchange and biogeochemical cycling around cross-vane restoration structures: Implications for stream restoration design, *Water Resources Research*, 49: 1-16, doi: 10.1002/wrcr.20185.

Abstract

Natural channel design restoration projects in streams often include the construction of cross-vanes, which are stone, dam-like structures that span the active channel. Vertical hyporheic exchange flux (HEF) and redox-sensitive solutes were measured in the streambed around 4 cross-vanes with different morphologies. Observed patterns of HEF and redox conditions are not dominated by a single, downstream-directed hyporheic flow cell beneath cross-vanes. Instead, spatial patterns of moderate (<0.4 md^{-1}) upwelling and downwelling are distributed in smaller cells around pool and riffle bed forms upstream and downstream of structures. Patterns of biogeochemical cycling are controlled by dissolved oxygen concentrations and resulting redox conditions, and are also oriented around secondary bedforms. Strong downwelling into the hyporheic zone (0.5-3.5 md^{-1}) was observed immediately upstream of structures, but was limited to an area 1-2 m from the cross-vane; these hyporheic flow paths likely rejoin the stream at the base of cross-vanes after residence times too short to alter nitrate concentrations or accumulate reaction products. Total hyporheic exchange volumes are ~0.4% of stream discharge in restored reaches of 45-55 m. Results show that shallow hyporheic flow and associated biogeochemical cycling near cross-vanes is primarily controlled by secondary bed forms created or augmented by the cross-vane, rather than by the cross-vane itself. This study suggests that cross-vane restoration structures benefit the stream ecosystem by creating heterogeneous patches of varying HEF and redox conditions in the hyporheic zone, rather than by processing large amounts of nutrients to alter in-stream water chemistry.

1. Introduction

The hyporheic zone is the area of the streambed and banks where water from the active channel mixes with sediment pore water and returns to the channel [*Triska et al.*, 1989; *White*, 1993; *Harvey and Wagner*, 2000; *Gooseff*, 2010]. Channel bed morphologies, such as steps, pools, and riffles, intensify the vertical flux of water through the hyporheic zone by a variety of hydraulic mechanisms [*Elliott and Brooks*, 1997; *Hester and Doyle*, 2008; *Buffington and Tonina*, 2009; *Tonina and Buffington*, 2009; *Endreny et al.*, 2011a]. Flow paths that redirect stream water through the hyporheic zone introduce dissolved oxygen, organic material, and nutrients to the streambed and increase the time that these solutes are in contact with sediments and organisms [*Findlay*, 1995; *Brunke and Gonser*, 1997; *Boulton et al.*, 1998]. In the streambed, enhanced geochemical and biological processes alter the water chemistry and create distinct gradients of solutes that are sensitive to oxidation-reduction (redox) conditions [*Champ et al.*, 1979; *Triska et al.*, 1989; *Jones et al.*, 1995; *Baker et al.*, 2000; *Morrice et al.*, 2000; *Lautz and Fanelli*, 2008]. Hyporheic exchange can influence surface water quality by chemically processing nutrients and pollutants in the streambed and by reintroducing chemically-altered water back to the stream [*McKnight et al.*, 2004]. Hyporheic exchange also maintains the health of aquatic species in the hyporheic zone, such as subsurface macroinvertebrate communities and fish embryos, by introducing stream water into benthic sediments and creating heterogeneous geochemical habitats in the streambed [*Brunke and Gonser*, 1997; *Boulton et al.*, 1998].

Because hyporheic exchange is an important component of stream ecology and is strongly controlled by stream morphology, and because the restoration of degraded streams creates dramatic changes in stream bed morphology, it is important that more studies focus on hyporheic exchange in restored streams. *Boulton* [2007], *Boulton et al.* [2010], and *Hester and*

Gooseff [2010; 2011] have emphasized the importance of the hyporheic zone in ecological restoration and called for a more complete approach to river restoration that includes hyporheic processes. However, post-construction assessment of restoration projects is not widespread [*Bernhardt et al.*, 2005; *Bernhardt et al.*, 2007; *Nagle*, 2007], and few studies have investigated the impact of in-stream restoration structures on hyporheic exchange fluxes or water chemistry, with some notable exceptions [*Kasahara and Hill*, 2006b; a; *Fanelli and Lautz*, 2008; *Kasahara and Hill*, 2008; *Lautz and Fanelli*, 2008].

Several studies have used numerical models to investigate hyporheic flow beneath instream obstructions analogous to restoration structures [*Kasahara and Wondzell*, 2003; *Gooseff et al.*, 2006; *Lautz and Siegel*, 2006; *Hester and Doyle*, 2008; *Crispell and Endreny*, 2009; *Endreny et al.*, 2011a; b; *Sawyer et al.*, 2011]. These studies typically show that a drop in water-surface elevation across a step or weir produces a single, deep hyporheic flow cell that connects downwelling areas upstream of the structure with upwelling areas downstream of the structure. In discussing the hyporheic effects of restoration structures, these modeling studies and reviews that reference them [e.g., *Hester and Gooseff*, 2010; 2011], concentrate on the hydraulics of the structure itself and emphasize that the type and size of the structure and the head drop across it, along with hydraulic conductivity, are the most important aspects of restoration structures for promoting ecological restoration of the hyporheic zone.

In the present field study, we measured patterns of hyporheic exchange and redox chemistry in restored stream reaches containing a common type of structure called a cross-vane. Cross-vanes are static, in-stream stone structures that span the active channel and are intended to reduce stream bank erosion and establish grade control [*Rosgen*, 2001; *NRCS*, 2007]. They are one of the most common structures called for in Natural Channel Design (NCD), used by federal

and local agencies to restore stability to degraded streams [*Doll et al.*, 2003; *Malakoff*, 2004; *Hey*, 2006; *NRCS*, 2007; *Lave*, 2009]. By directing water energy into the middle of the channel and creating a hydraulic step in the water surface profile, cross-vanes maintain a deep plunge-pool downstream of the structure and contribute to the formation of pool-riffle sequences up- and downstream in the restored reach [*Rosgen*, 2001]. Shallow pool-riffle sequences that surround cross-vanes are referred to as secondary bedforms in this study, and are considered an integral component of the cross-vane because they are purposefully created by the restoration process and maintained by the structure hydraulics, and often are absent in degraded streams pre-restoration. Secondary bedforms created or maintained by restoration structures have not received much attention in existing literature on the hyporheic effects of stream restoration [with the exception of scour topography around logs in *Sawyer et al.*, 2011].

The goal of this study is to characterize the patterns and magnitudes of hyporheic exchange flux (HEF) around restoration structures in complex field settings in order to understand the conditions they create in the hyporheic zone and the potential effects this may have on stream ecosystem health. By understanding the effects of restoration structures as constructed in real streams, we can better design stream restoration to achieve specific ecological goals. For example, if restoration structures create strong hyporheic flow cells with large volumetric HEF rates relative to stream flow and moderate retention times, they may significantly contribute to whole-stream nutrient processing and remediation [*Boulton et al.*, 1998]. This study extends the scope of those before it [*Fanelli and Lautz*, 2008; *Lautz and Fanelli*, 2008] by focusing on longer reaches of streams restored using NCD principles that contain specially engineered restoration structures, and by employing consistent methods across multiple systems with varying stream sizes and cross-vane morphologies.

In this study, we use hydrologic, thermal, and hydrochemical data collected around 4 cross-vanes in order to characterize the effects of restoration structures on patterns of HEF and biogeochemistry in the streambed hyporheic zone. We hypothesized that (1) cross-vanes promote hyporheic exchange between the stream and streambed, both by creating a step in the water surface profile and by maintaining secondary pool and riffle bedforms that drive exchange; and (2) downstream-directed flow cells beneath cross-vanes promote rapid HEF and associated biogeochemical cycling (relative to the streambed far from cross-vanes) and create large oxic zones in the hyporheic zone upstream and anoxic zones downstream of the structures. The second hypothesis was not supported by observations, as we found oxic downwelling zones were small in area (relative to structure size), and strong anoxic upwelling zones were not observed. Instead, secondary bedforms drive HEF and control redox conditions over the majority of the streambed area in each restored reach and, as a result, the majority of biogeochemical cycling in restored reaches is likely due to exchange around secondary bedforms.

2. Methods

2.1 Study site descriptions

Four cross-vane restoration structures were studied during the summer and fall of 2009 and 2010 (Fig. 1; Table 1). The 4 cross-vanes were constructed as part of restoration projects on 3 degraded streams in lowland, agricultural catchments of Central New York. Two cross-vanes (denoted A and B) were selected on a reach of Nine Mile Creek (contributing area to study reach: 177 km^2) in Marcellus, NY. The studied cross-vanes in Marcellus are 2 structures in a series of repeating cross-vanes and boulder J-hook structures [*Rosgen*, 2001]; Ninemile A is 100 m downstream of Ninemile B, with 3 J-hooks positioned irregularly between them around a meander bend. One cross-vane was selected on the West Branch of Owego Creek (contributing

area: 74.4 km^2) and another on Boyer Creek (19.3 km^2), a tributary to the West Branch, in Caroline, NY. The studied reach at Owego Creek was immediately downstream from a J-hook structure, and the cross-vane at Boyer Creek was one in a series of repeating cross-vanes. Data were collected at Ninemile A in both 2009 and 2010 (referred to as the NMA'09 and NMA'10 datasets), at Ninemile B (NMB) in 2009 and at Boyer Creek (Boyer) and W. Branch Owego Creek (Owego) in 2010. Secondary pool-riffle bedforms were identified upstream and downstream of the cross-vane at Ninemile A and downstream from the cross-vane at Owego and Boyer Creeks. The sediment texture was very similar among all 4 sites, and consisted of gravel and cobbles, with some silt and sand, and a few boulders in the plunge pool at Ninemile A. Backfill material placed upstream of the cross-vanes after construction was all loosely packed coarse gravel and cobbles, with some observed macropores, and therefore likely has a higher hydraulic conductivity (K) than the rest of the reaches, although K was not measured.

The stream reaches and specific cross-vanes were selected to represent a range of different morphologies (step heights, plunge-pool depths, and surrounding pool-riffle bedforms), and different elapsed time periods since construction (Table 1). Other factors that were considered in the selection process were stream size, substrate type, structure symmetry, the presence of repeating structures, and ease of access. Cross-vanes that were deemed failed by their designers were excluded, including those that had been covered by aggrading sediment or laterally bypassed by channel erosion and those in which structural elements had moved since construction. The choice to study only structures that met their stated design criteria was consistent with our goal to evaluate cross-vane design, rather than cross-vane construction success, and to focus on improving the design of future restoration structures.

2.2 Field and laboratory methods

At each study reach, the bed topography, cross-vane topography, and water surface elevation were surveyed and mapped (Fig. 2), and baseflow discharge was measured with a velocity meter (2010 only). Thermal and chemical data at each site were collected using temperature sensors and mini-piezometers installed in the streambed.

Temperature time series were recorded in the streambed to estimate vertical hyporheic exchange using heat tracing [*Hatch et al.*, 2006; *Keery et al.*, 2007]. Vertical temperature profiles (TPs) were constructed by embedding 7 iButton temperature sensors and loggers (DS1922L, Maxim, Inc., 0.0625 ºC resolution, 0.5 ºC accuracy, 130 s response time) [*Hubbart et al.*, 2005; *Johnson et al.*, 2005] into 2.5-cm diameter steel rods at depths of 5, 10, 15, 20, 25, and 30 cm in the streambed and at 5 cm above the streambed. Silicone sealant was used to waterproof the sensors and insulate them from the rod. TP rods were installed in the streambed along the reach thalweg and around the cross-vane and secondary bedforms (Fig. 2). Temperature time series were recorded at a 10-min sampling rate for the entire study period at each site (study period ranged from 14 to 29 days).

At one time during each study period in 2010, temperature in the streambed was mapped at high spatial resolution using a hand-held temperature probe (Traceable Digital Thermometer, Control Company, 0.001 ºC resolution, 0.05 ºC accuracy) [following *Conant*, 2004; *Lautz and Ribaudo*, 2012]. Temperature was measured at 70-90 locations (Fig. 2) in the shallow streambed by directly inserting the probe to a depth of 7 cm [similar to *Lautz and Ribaudo*, 2012]. Temperature mapping of the entire stream bed was completed in less than 2 h, and was conducted in the mid-afternoon when stream temperature was at a maximum and changing most

slowly. During bed mapping, stream temperature varied by less than 0.45 ºC at Owego, 1.33 ºC at NMA'10, and 0.45 ºC at Boyer.

Mini-piezometers (1.3-cm dia.) were installed in the streambed in a grid at each site (40-53 piezometers per site, Fig. 2), with a 5-cm screened interval centered at 20 cm depth (2009 sites) or 15 cm (2010 sites). Samples of pore water were drawn from the piezometers in a single day during each study period. Dissolved oxygen (DO) in pore water was measured in the field (at Owego and NMA'10). At the 2010 sites, stream water and shallow groundwater samples from riparian piezometers were also collected on the same day. Water samples were filtered using 0.7 μm glass fiber filters and analyzed for the ions chloride, nitrate, sulfate, sodium, ammonium, potassium, magnesium, and calcium by ion chromatography (Dionex ICS-2000), and for total dissolved iron, manganese, phosphorus, silica, and strontium using inductively-coupled plasma optical emission spectrometry (Perkin-Elmer OPTIMA 3300DV, 2010 samples only).

2.3 Heat transport modeling

The one-dimensional heat transport model of *Hatch et al.* [2006] was used to calculate vertical seepage flux (*q*, units of distance per time) between pairs of temperature sensors in the TPs [see Equation 6a of *Hatch et al.*, 2006]. Values of sediment thermal and textural properties used for all study sites are given in Table 2. Temperature time series recorded in TPs in the field were analyzed using the VFLUX program [*Gordon et al.*, 2012], which calculates vertical flux through time between all sensor pairs using the diel temperature signal and the *Hatch et al.* [2006] amplitude method. Vertical flux estimates were calculated every 2 h for each pair of sensors, and 48 h of results were discarded from the beginning and end of each series [*Gordon et al.*, 2012]. A Monte Carlo error analysis (1000 realizations) in VFLUX was used to estimate potential error due to uncertainty in sediment thermal and textural parameters. Parameters were

assumed to be normally distributed with means and standard deviations given in Table 2. Parameter uncertainty has been exaggerated here, and is more constrained by sediment textural observations than represented in the error analysis. The analysis does not take into consideration other potential sources of error, such as errors in sensor spacing (Δz) or temperature measurement, or error from violations of model assumptions [e.g. *Lautz*, 2010]. More details of the heat transport modeling, signal processing methods, and error analysis can be found in *Hatch et al.* [2006], *Keery et al.* [2007], *Lautz* [2010], *Gordon et al.* [2012], and *Lautz* [2012].

2.3 Data analysis methods

The mapped bed temperature measurements made with the hand-held probe were used along with estimated vertical flux rates from the TPs in order to develop an empirical relationship between vertical HEF and shallow streambed temperature at a fixed depth and constrained time. This method, developed by *Conant* [2004] and modified by *Lautz and Ribaudo* [2012], relates the strength of the temperature gradient in the shallow streambed to vertical flux [*Stonestrom and Constantz*, 2003; *Anderson*, 2005]. First, bed temperatures at TP locations were estimated by spline interpolation of the 70-90 measurements made by hand during bed temperature mapping. The interpolated temperatures at each of the TPs were then related by non-linear regression to the mean flux value (over the time period that bed temperatures were mapped) for the shallowest pair of sensors in each TP. In order to test the validity of the interpolated temperatures at TP locations, a second regression was also performed using the actual bed temperatures measured at the TP 5-cm sensor, and the regressions were compared. The first regression model was then used to infer the flux rate at each of the 70-90 locations where bed temperature was measured, creating a dense grid of flux estimates over the streambed.

Geochemical data from streambed pore water samples were analyzed for bivariate relationships and spatial patterns. DO concentrations, chemical species sensitive to biogeochemical cycling or changing redox conditions, and those correlated with these species were incorporated into a square-root-transformed principle component analysis (PCA) [*Lautz and Fanelli*, 2008; *Mencio and Mas-Pla*, 2008; *Woocay and Walton*, 2008], or for samples with missing values, probabilistic PCA [*Roweis*, 1998; *Tipping and Bishop*, 1999; *Porta et al.*, 2005].

3. Results

3.1 Vertical flux magnitudes and patterns

The diel temperature signal was observed at all functioning TP sensors and it was therefore possible to calculate vertical water flux (q, positive downward) for all functioning sensor pairs. No data were recovered from some sensors at the 2009 sites, which malfunctioned due to excessive vibration during installation. TPs with more than one sensor malfunction, but at which flux measurements were still possible, include TP01, 06, and 07 at NMA'09, and TP05 at NMB. No flux measurements were possible at TP06-10 at NMB. A moderate difference in amplitude was detectable between sensors spaced 5 cm apart, meaning that a Δz of 5 cm was within the sensitivity of the model [*Hatch et al.*, 2006; *Keery et al.*, 2007; *Gordon et al.*, 2012]. In this study, therefore, we use only flux measurements made with sensor pairs separated by 5 or 10 cm, because of the higher spatial precision afforded by closely spaced sensors. Monte Carlo analysis was used to define 95% (2σ) confidence intervals around each flux estimate due to parameter uncertainty. Flux values between 0.086 m d^{-1} and -0.119 m d^{-1} have confidence intervals that include zero, and are therefore considered not significant.

Spatial patterns and magnitudes of flux over the streambed showed similar patterns among the study sites, except at Boyer Creek. TPs were qualitatively grouped by location into 4

categories based on streambed morphology: (1) upstream of riffles and in the glides of shallow pools, (2) downstream of riffles and in the runs of shallow pools, (3) directly upstream of cross vanes, and (4) in the plunge pools downstream of cross-vanes (Fig. 3). In general, vertical flux was of relatively small magnitude over much of the streambed in categories 1, 2, and 4. At these locations, the widest range of vertical flux rates (median fluxes of 0.50 to -0.31 m d^{-1}) was observed at the shallowest depths in the streambed (between the stream and 5-cm sensors), which is the best measure of the magnitude of water exchange across the stream-streambed interface. Upstream of riffles (category 1), shallow exchange flux was downwelling (positive), with the exception of one outlier (NMA'10 TP10, installed just downstream of a boulder) (Fig. 3A). Downstream of riffles (category 2), exchange flux was upwelling (negative) at all TPs (Fig. 3B). In plunge pools downstream from the cross-vane drop (category 4), shallow exchange flux was primarily upwelling (as strong as -0.31 m d^{-1} at Owego TP10 and NMA'10 TP08), with 3 exceptions: Owego TP07, NMA'09 TP07, and NMA'10 TP07 (Fig. 3D). All 3 exceptions occurred at TPs directly in the path of a narrow turbulent jet of water that reached into the plunge pool from the cross-vane, just downstream from the point where most water spilled over the drop. At TP locations that were less than 1 m upstream from cross-vanes (category 3), strong downwelling was observed at all depths down to 30 cm (Fig. 3C). Measured median flux rates ranged from 0.15 m d^{-1} to over 3 m d^{-1} at a range of depths, but this strong downwelling flux was limited to the area of the streambed immediately upstream from cross-vanes, and was not observed as close as 2.4 m upstream from the cross-vane at NMA'10 TP05, which instead followed the shallow upwelling pattern characteristic of locations downstream from riffles.

Patterns of hyporheic flux observed at Boyer Creek were unlike those seen at the other sites (Fig. 4), which may be related to loss of water from the stream channel during the study

period, since dry summer weather caused the apparently ephemeral stream to cease flowing after July 3. Flux results show that Boyer Creek was losing water to the subsurface across all TP locations throughout the study period (median fluxes of 0 to 2.7 m d^{-1}). Generally higher downwelling rates were measured upstream of the riffles and cross-vane than downstream of these bedforms, but no clear patterns based on morphologic category are discernable (Fig. 4).

3.2 Streambed temperature and flux relationships

Temperature measured in the streambed directly using the hand-held probe ranged 24.00-27.43 ºC at Owego, 24.61-27.70 ºC at NMA'10, and 23.32-24.5 ºC at Boyer. Bed temperatures estimated at TP locations using spline interpolation ranged 24.17-26.39 ºC at Owego and 25.09-26.82 ºC at NMA'10. These temperatures compare favorably to those measured directly at the TP locations by the 5-cm sensors at the time of mapping (Fig. 5).

Regressions of flux rates to interpolated bed temperature at TP locations yielded strong relationships (quadratic r^2=0.71, p=0.002 at Owego, cubic r^2=0.76, p=0.003 at NMA'10), showing that higher bed temperatures were predictive of stronger downwelling flux, as expected (Fig. 5). Relationships between flux and measured bed temperature at the 5-cm TP sensors was even stronger (r^2=0.99, p<0.001 at Owego, r^2=0.97, p<0.001 at NMA'10), and the regression equations of the two relationships are very similar for each site (Fig. 5).

Flux rates estimated at the hand-held probe locations using the empirical models from Fig. 5 (using interpolated bed temperature) ranged from -0.26 to 1.40 m d^{-1} at Owego and -1.00 to 7.03 m d^{-1} at NMA'10; however, flux rates at the high end of this range are improbably large due to extrapolation of the non-linear regression beyond its steep upper end. We therefore assigned any temperature at NMA'10 greater than 26.82 ºC (corresponding to a flux of 2.08 m d^{-}

^{1}) a flux rate of 2.08 m d^{-1} in Fig. 6. Similarly, temperatures at Owego greater than 26.39 ºC were assigned a flux rate of 0.72 m d^{-1}.

Maps of exchange flux over the streambed surface (Fig. 6) were used to estimate the instantaneous volume of stream water that was exchanging with the hyporheic zone over the study reaches at Owego and NMA'10, at the time of streambed temperature mapping. Only calculated flux values outside the estimated uncertainty limits around zero (0.086 and -0.119 m d^{-1}) were considered as significantly different from zero. Out of 647 m^2 bed area at NMA'10, 29% was downwelling, 44% was upwelling, and the remainder was not significantly different from zero. The total flux of stream water into the bed was 140 m^3 d^{-1}, or 0.4% of total stream discharge, and the total flux of water out of the bed into the stream was 99 m^3 d^{-1}, or 0.3% of total stream discharge. Out of 321 m^2 at Owego, 35% was downwelling and 20% was upwelling. Total flux of water into the bed was 30 m^3 d^{-1}, or 0.4% of stream discharge, and the flux of water out of the bed was 11 m^3 d^{-1}, or 0.2% of stream discharge.

3.3 Streambed geochemistry

Relationships between solutes were similar among the study sites, with the exception of Boyer Creek. Solutes with low variability in streambed pore water include Cl^-, Na^+, Mg^{2+}, Ca^{2+}, and Sr (Table 3). Concentrations of these generally conservative solutes were not significantly different (α=0.05) between pore water and stream water, while riparian groundwater was significantly different from streambed pore water, with few exceptions. Solutes with relatively high variability included DO, NO_3^-, NH_4^+, Fe, Mn, and total dissolved phosphorus (TDP) (Fig. 3). Species known to be redox sensitive (DO, NO_3^-, NH_4^+, Fe, Mn, and SO_4^{2-}) showed significant correlations to one another, except for Fe. In general, DO was highest in stream water, and NO_3^- values higher than stream water were only observed in 6 piezometers at Owego and nearly all

piezometers at Boyer. Pore water NH_4^+ concentrations were equal to or greater than the very low stream values, and stream values of SO_4^{2-} were at or near the high end of the pore water range.

Example plots of redox sensitive solutes are shown in Fig. 7. DO appears to be a direct or indirect control on many of the measured solutes. NO_3^- decreases with DO in the streambed, with NO_3^- differences from stream water greatest at the lowest DO concentrations; however, NO_3^- begins to decrease from stream water concentrations with DO as high as 10 mg l^{-1} (Fig. 7A). SO_4^{2-} also decreases with decreasing DO, but in a more threshold fashion, not decreasing significantly below the stream concentration unless DO is at or below ~4 mg l^{-1} (Fig. 7B). Fig. 7C (NO_3^- versus SO_4^{2-}) shows several processes occurring in the hyporheic zone. Several piezometers appear on a linear mixing line between steam water and shallow groundwater, while most of the others appear on the plot in an "L" shape pattern, in which NO_3^- appears to be consumed while SO_4^{2-} holds steady near the stream value of 14.5 mg l^{-1}. Only when NO_3^- is near zero does SO_4^{2-} decrease from the stream value to below groundwater concentration and towards zero (Fig. 7C). Similar patterns appear in plots of NO_3^- versus Mn and NH_4^+, except that both are produced in the hyporheic zone as NO_3^- is consumed (Figs. 7D and 7E). Two other dissolved species, while not generally recognized as redox sensitive [*Champ et al.*, 1979; *Baker et al.*, 2000] appear to have relationships with the redox-sensitive solutes: K^+ and Si both have positive relationships with NH_4^+ ($r= 0.76$, $p<0.001$ and $r=0.53$, $p<0.001$, respectively; Figs. 7F and 7G) and negative relationships with DO ($r=-0.46$, $p=0.001$ and $r=-0.44$, $p=0.002$, respectively; not shown). Manganese, NH_4^+, K^+, and Si concentrations are all greater in groundwater than in stream water, but are much higher in hyporheic pore water than in groundwater (Figs. 7D-G).

Geochemical patterns at Boyer were different than the other sites. Streambed pore water was very stream-like, with conservative ion concentrations in a tight range near stream values

(Table 3). Redox-sensitive solutes were also much less variable and closer to stream concentrations. Streambed NO_3^- and SO_4^{2-} concentrations were not statistically different from the stream (p=0.062 and 0.101), and there was only very slight K^+, NH_4^+ and Mn production.

While the relationships among DO and the redox-sensitive solutes are non-linear in some cases (see Fig. 7), they do exhibit consistent positive and negative relationships that co-vary, and can be made more linear by applying the square root transform. We applied PCA separately to each site, using the species NO_3^-, NH_4^+, SO_4^{2-}, and K^+, plus DO, Mn, and Si for the 2010 sites. PCA results (Table 4) show that the first principle component (PC1) explained the majority of total variation in each data set. PC1, which we interpret as a linear "redox index" has strong, significant correlations with all of the individual species, with positive loadings and correlations for DO, NO_3^-, and SO_4^{2-}, and negative loadings and correlations for NH_4^+, K^+, Mn, and Si (Tables 4 and 5), except K^+ at NMB. Positive values of the redox index therefore indicate oxidizing conditions, and negative values indicate reducing conditions. PC2 and PC3 did not have consistent loadings or correlations with solutes among different sites.

Spatial patterns of redox-driven biogeochemical cycling in the streambed are summarized by the redox index in Figs. 8 and 9. Patterns were similar between the study sites, except for Boyer Creek. In general, oxidizing zones (high NO_3^-, low NH_4^+, etc.) are found in the glide directly upstream of cross-vanes, where downwelling water flux was strongest, and in glides upstream of secondary riffles. Reducing zones (low NO_3^-, high NH_4^+, etc.) are found in secondary pools upstream of the cross-vanes and in runs downstream of secondary riffles. Plunge pools do not show primarily reducing or oxidizing conditions, and are more variable among sites.

In order to summarize patterns of flux and biogeochemical cycling and their relations to bed topography and water surface elevation, longitudinal profiles of interpolated flux rates and

the redox index were made down the thalweg of each site (Fig. 9). Downwelling zones, primarily upstream of the cross-vane and at riffle heads, are associated with oxidizing conditions, and upwelling zones, primarily in pools downstream from secondary riffles, are associated with reducing conditions. Pearson correlations between flux and the redox index are somewhat low but significant, with coefficients of 0.37 ($p<0.001$) at Owego and 0.40 ($p<0.001$) at NMA'10. Correlations between flux at TP locations and interpolated redox index are higher in the 2009 sites, but are necessarily limited by the small number of data points (Fig. 9). Pearson coefficients are 0.80 ($p=0.006$, $n=10$) at NMA'09 and 0.89 ($p=0.04$, $n=5$) at NMB, however, the correlation at NMB is controlled by a single point (TP04), and is therefore not valid with the data available.

4. Discussion

4.1 Patterns of hyporheic exchange around cross-vanes

Heat tracing methods and error analysis were successfully used to measure vertical HEF and clearly differentiate between upwelling and downwelling zones. Vertical patterns of flux with depth (Fig. 3) showed that the greatest variability in flux across the streambed occurs at the shallowest depths, between the stream and 5 or 10 cm. These flux rates across the streambed interface are important for ecosystem health because they represent the magnitude of the exchange of water and dissolved constituents between the channel and hyporheic zone. With increasing depth, vertical flux rates showed a tendency to decrease towards zero, and we attribute this pattern to penetration of the TP rods through shallow, curved hyporheic flow paths that become more horizontal with depth [similar to *Briggs et al.*, 2012]. This pattern was most evident upstream and downstream of secondary riffles (Figs. 3A and B), where shallow hyporheic flow paths are expected to start and end. This vertical pattern is not an artifact of the

one-dimensional heat tracing method, which can measure the vertical component of oblique flow even under non-ideal conditions [*Lautz,* 2010; 2012].

Empirical relationships between bed temperature and vertical flux rates were well-characterized (Fig. 5). The similarity between the regression curves when using interpolated bed temperature and temperature measured directly at the TP indicates that interpolation yields a satisfactory approximation of temperature and simply adds randomly distributed error around the true value (Fig. 5). The shape of the regression curves in Fig. 5 is also to be expected from heat transport theory [*Stonestrom and Constantz,* 2003; *Anderson*, 2005]. Because the shallow streambed can never be warmer than the stream or colder than deep groundwater, there will be two asymptotic limits to the streambed temperature, at the groundwater and stream water values [*Conant*, 2004; *Lautz and Ribaudo,* 2012]. The empirical relationship in Fig. 5B, approximated with a cubic function, shows stream temperature approaching both limits, especially where strong downwelling flux caused the shallow bed temperature to approach the stream temperature of ~27 ºC. Fig. 5A only appears to cover the convex-up part of the relationship, probably because the range of flux was narrower at Owego than at NMA'10, and was therefore fit with a quadratic.

Spatial patterns of upwelling and downwelling flux around cross-vane steps, plunge pools, and secondary riffles and pools were different and more complex than typically represented in modeling studies of in-stream structures [e.g. *Lautz and Siegel*, 2006; *Hester and Doyle*, 2008]. Results of hyporheic flux measurements at Ninemile and Owego Creeks show that these cross-vanes create a small zone of intense downward flux (~0.5-3.5 m d^{-1}) in the glide area immediately upstream from the structure drop. However, this zone does not extend beyond ~2 m upstream from any of the cross-vanes, and vertical flux does not exceed 0.4 m d^{-1} elsewhere in the studied reaches. It is likely that these zones of strong downwelling are partially created or

facilitated by observed areas of coarse, loose backfill material with high conductivity that were placed just upstream of the structures during cross-vane construction. Furthermore, these strongly downwelling areas may not be connected to deep (and presumably long-residence time) hyporheic flow paths that extend beneath the cross-vanes for two reasons: (1) Strong upwelling flux was not measured downstream of cross-vane structures as predicted by models [e.g. *Sawyer et al.*, 2011], instead flux patterns in the plunge pools are both up- and downwelling and moderate in magnitude; and (2) water was qualitatively observed exiting the streambed through macropores and gaps in the downstream face of the loosely-constructed stone structures, presumably limiting the amount of water being forced into deeper and longer-residence time flow paths. However, it is possible that flux measurements missed small areas of intense upwelling, or that flow paths are longer and deeper than expected and upwell in dispersed zones farther downstream than the plunge pools. These results show that the studied cross-vanes create only limited zones of strong downwelling, and are unlikely to dramatically increase total hyporheic exchange along restored stream reaches, even in sections of repeated structures placed 25 to 150 m apart like at the Boyer, Owego, and Ninemile Creek restoration projects.

In plunge pools, moderate upwelling was measured, as well as limited zones of downwelling (Figs. 3 and 6). One TP in each plunge pool at NMA'10, NMA'09, and Owego was located in jets of high turbulence just downstream of where most water poured over the cross-vane structure. These 3 TPs have distinctive patterns of vertical flux with depth (Fig. 3D), where flux is downwelling (at up to 1.3 m d^{-1}) at the shallowest depths, but decreases to near zero by 7.5 to 12.5 cm depth. Shallow downwelling flux in a generally upwelling zone is likely caused by the inertial force of water impinging on the streambed and/or turbulent momentum just

downstream of the cross-vane step and has been predicted in dynamic modeling studies [*Endreny et al.*, 2011a; *Sawyer et al.*, 2011], but has not been observed in field settings before now.

Patches of both upward and downward low-magnitude exchange were observed around secondary pool and riffle structures, with downwelling evident upstream and at the heads of riffles (glides), especially where the plunge pool transitions into the first riffle downstream of the cross-vane, and upwelling downstream of riffles (runs) and in small pools (Fig. 6). The one exception, at NMA'10 TP10, was observed just downstream from a partially-buried boulder, which are known to create localized exchange zones [e.g., *White et al.*, 1987]. Combined with the patterns of flux with depth that suggest shallow, curved hyporheic flow paths, these results show that patterns of upwelling and downwelling over the majority of the streambed area consist of shallow hyporheic flow cells organized around secondary pool and riffle bedforms, similar to patterns observed around microforms in simulations of *Kaser et al.* [2012].

Boyer Creek was a distinct outlier in both flux and chemical patterns. The stream was strongly losing water to the subsurface over its entire bed area. Furthermore, the studied stream reach, apparently ephemeral, was completely dry by the end of the study (see Section 3.1). These results show that, in this case, an ephemeral stream in the last stages of channel flow loses water at a rate higher than potential hyporheic exchange caused by pressure differences on the bed around cross-vanes and secondary bedforms. As such, water lost to the streambed does not return to the stream and the hyporheic zone is non-existent.

Vertical flux maps (Fig. 6) and total hyporheic exchange estimates show that the volume of water downwelling and upwelling at both NMA'10 and Owego were of the same order of magnitude. Total volumes of water exchanged between the stream and hyporheic zone in the studied reaches was small compared to total stream discharge in the channel (total downwelling

volume was 0.4% of total stream discharge at both NMA'10 and Owego, in stream reaches of 55 and 45 m, respectively). These figures are much smaller than estimates of HEF relative to total stream discharge in some smaller streams [*Kasahara and Wondzell*, 2003], and smaller than values for other in-stream structures [*Hester and Doyle*, 2008; *Crispell and Endreny*, 2009] cited in hyporheic restoration reviews [*Hester and Gooseff*, 2011]. Much of these differences between systems may be due to different hydraulic conductivity (K), which varies by orders of magnitude between sites and can be a dominant control on HEF volume [*Hester and Gooseff*, 2011]. Alternatively, our measurements may underestimate total hyporheic exchange by failing to measure small areas of intense upwelling between our sampling locations or dispersed upwelling downstream from our instrumented reach. Total hyporheic exchange may also be larger than reported here if there are substantial amounts of shallow exchange around individual sand and gravel particles, which may not be observed with our shallowest (5-cm deep) temperature sensors. Because discharge through the hyporheic zone is so small, we expect the processing of solutes in these restored hyporheic zones to have a small effect on total stream solute loads.

4.2 Patterns of biogeochemical cycling around cross-vanes

Patterns of solutes in the streambed were used to further understand the dynamics of water exchange and the connections between HEF and biogeochemical cycling around stream restoration structures. Despite the small magnitudes of HEF, we found that the hyporheic zone still contains significant proportions of stream water that undergoes chemical change due to biogeochemical reactions in streambed sediments. Biogeochemical cycling in the streambed at Ninemile Creek and Owego Creek was evident based on co-varying changes in redox-sensitive solutes, as would be expected based on commonly-accepted redox sequences: aerobic respiration followed by NO_3^- reduction or denitrification, Mn(IV) reduction to Mn(II), Fe(III) reduction to

Fe(II), SO_4^{2-} reduction, and finally methanogenesis [*Champ et al.*, 1979; *Baker et al.*, 2000]. Evidence for most of these processes was observed in the hyporheic zone in this study, including the consumption of DO along with apparent NO_3^- reduction and NH_4^+ production from organic matter mineralization in anoxic flow paths [*Triska et al.*, 1989; *Jones et al.*, 1995; *Valett et al.*, 1996; *Pretty et al.*, 2006; *Lautz and Fanelli*, 2008; *Zarnetske et al.*, 2011], soluble Mn production [*Baker et al.*, 2000; *Lautz and Fanelli*, 2008], and SO_4^{2-} reduction [*Valett et al.*, 1996; *Morrice et al.*, 2000]. Because these latter processes are initiated in the hyporheic zone at different threshold levels of redox potential, SO_4^{2-} and Mn have non-linear relationships to DO and other redox-sensitive solutes like NO_3^- (Figs. 7B-D).

Evidence of NO_3^- production, possibly due to nitrification [*Duff and Triska*, 2000] was observed at several locations at Owego and many at Boyer. The presence of NO_3^- production at Boyer is assumed to be due to the consistent, moderately-strong downwelling of stream water into the hyporheic zone, which allows for contact time between organic matter, DO, and organisms that may be adequate to promote nitrification. At Owego, a zone of NO_3^- production appeared upstream of a large riffle where moderate downwelling was observed. No NO_3^- production was observed in strong downwelling zones near cross-vanes, perhaps because retention times were too short for reaction products to accumulate [*Lautz and Fanelli*, 2008]. As a result, the hyporheic zones at the studied sites are not considered to be significant sources of NO_3^- to stream water, although NO_3^- may be produced and contributed to groundwater at Boyer.

The processes associated with observed relationships of K^+ and Si to reduced pore water species and upwelling flux are not as well documented in previous literature. High Si concentrations in reducing zones were also found by *Lautz and Fanelli* [2008], where the authors attributed it to the dissolution of buried siliceous particles. It may be that dissolved Si is found at

the upwelling ends of longer subsurface flow paths due to longer contact time with particulate Si, and not directly due to redox conditions. Potassium is a nutrient found in leaf litter that is rapidly released during decomposition by leaching and mineralization [*Blair*, 1988], and may simply be released during the decomposition and respiration of organic matter in hyporheic flow paths. However, K^+ release from clay minerals may also be related to redox processes. Potassium participates in cation exchange [*Barshad*, 1954] and may be displaced by NH_4^+ ions as they accumulate in anoxic pore water, causing K^+ desorption [*Evangelou et al.*, 1994]. The addition of NH_4^+ has also been found to increase the likelihood of K^+ to participate in cation exchange with Ca^{2+} and Mg^{2+} [*Lumbanraja and Evangelou*, 1990]. On the other hand, the presence of SO_4^{2-} ions causes a stronger binding of K^+ to clays [*Evangelou et al.*, 1994], so a reduction in SO_4^{2-} concentration may also cause K^+ desorption. Finally, it has been shown that the reduction of Fe(III) to Fe(II) in clay minerals causes a decrease in K^+ sorption and an enhancement in the replaceability of K^+ by Mg^{2+} [*Barshad and Kishk*, 1968]. All of these processes have the potential to create elevated K^+ concentrations along reducing hyporheic flow paths, such as were observed in upwelling zones. Although no previously published work to our knowledge has observed K^+ co-varying with other redox-sensitive solutes in the hyporheic zone, the observation of this pattern at Owego and in two successive years at Ninemile A suggests it is not anomalous.

As a result of the consistent downwelling flux across the bed at Boyer, pore water chemistry was uniformly stream-like for most of the measured solutes, and redox chemistry was mostly oxidizing. These geochemical results support the physical flux results, and show that stream water was infiltrating into the bed throughout the reach.

The co-variation of the solutes DO, NO_3^-, NH_4^+, Mn, SO_4^{2-}, K^+, and Si was successfully summarized with a single principle component or “redox index” at all Ninemile and Owego

sites, suggesting that they are similarly governed by microbial respiration under varying redox conditions [a method also used by *Lautz and Fanelli*, 2008; *Briggs et al.*, in revision]. Spatial patterns of the redox index (Figs. 8 and 9) and the correlations between flux and redox generally show that oxidizing conditions are prevalent in downwelling zones and reducing conditions in upwelling zones.

The spatial patterns of the redox index help corroborate physical flux measurements and elucidate the biogeochemical effects of HEF in restored stream beds. At the Ninemile sites, where the water surface drop over the cross-vane was >22 cm, the streambed was most oxidizing in a narrow zone directly upstream of the cross-vane, and also oxidizing where the plunge pool transitions into the first riffle. At Owego, where the drop was only 10 cm, the streambed redox chemistry showed less variation around the cross-vane, and the most oxidizing conditions were upstream of the secondary riffle at the bottom of the reach (Fig. 9). Owego, where there was a small standing wave downstream of the cross-vane, also had a stronger oxidizing zone in the plunge pool than at NMA'10, where turbulent downwelling was observed. At all sites except Boyer and NMB, the streambed was most reducing in pools upstream from the cross-vane, and not in the plunge pool downstream of the cross-vane, where momentum from the cross-vane drop and turbulence in the water column created variable flux and redox conditions (Fig. 9).

The only site that was sampled in two consecutive years showed remarkable repeatability in time. Both flux (Figs. 3 and 9) and chemical patterns (Figs. 8-9) at the Ninemile A site are visually similar between the 2009 and 2010 summer field seasons. Furthermore, the redox index down the thalweg profile (Fig. 9) is strongly correlated between the two data sets ($r=0.91$, $p<0.001$) over the distance that they overlap in space. This correlation shows that our methods are repeatable and, more importantly, chemical conditions in the hyporheic zone may be stable in

time, at least at similar flow conditions. It appears that benthic niches specifically defined by their chemistry may persist for multiple seasons in the same areas of the streambed.

4.3 Implications for nutrient processing and ecosystem health

The consistent and predictable spatial patterns found around cross-vanes in this study suggest that these restoration structures do promote hyporheic exchange around steps in the water surface profile and secondary pool and riffle bedforms. However, evidence for strong downstream-directed flow cells beneath cross-vanes was not observed in measurements of flux or chemistry. While the strongest flux rates were measured upstream of cross-vanes, these flow cells do not appear to be connected to long flow paths that strongly upwell reduced solutes downstream of cross-vanes. Instead, cross-vanes and secondary bedforms create heterogeneous patches of low-magnitude hyporheic exchange and redox-driven biogeochemical cycling in the shallow hyporheic zone over the majority of the streambed area. Along with the low volumetric rate of HEF measured over these reaches, these results suggest that cross-vanes in the studied systems do not process significant amounts of nutrients or pollutants for the whole-stream system, and that the influence of cross-vanes on ecosystem health in these streams is primarily realized in benthic habitat heterogeneity, rather than in-stream geochemical remediation.

The role of the hyporheic zone in stream ecosystems has recently been reexamined as it has become clear that the volumetric rate of hyporheic exchange flows are often a small fraction of total stream discharge [e.g. *Wondzell*, 2011], as was found in this study. The importance of the streambed hyporheic zone to stream channel chemistry, and the ability of the hyporheic zone to mitigate pollutants such as nitrate transported in channel water, is related to the mass transport of solutes through the hyporheic zone [*Findlay*, 1995], as well as reaction rates [*Harvey and Fuller*, 1998]; however, the relative volume of the hyporheic zone varies widely and is often quite small,

especially in systems with low K [*Wondzell*, 2011]. On the other hand, heterogeneous patches of upwelling and downwelling water with strong gradients in redox conditions, solute concentrations, and temperature provide diverse benthic habitat for hyporheic organisms, even if overall HEF is small [*Brunke and Gonser*, 1997; *Boulton et al.*, 1998; *Wondzell*, 2011].

The view of the hyporheic zone as a chemical processing zone for the stream and downstream ecosystems can be thought of as the stream-centric perspective, and the view of the hyporheic zone as a patchwork of heterogeneous benthic niches can be thought of as the benthic-centric perspective. Clearly these viewpoints are not mutually exclusive, and both are important in hyporheic restoration [*Boulton*, 2007; *Hester and Gooseff*, 2010]. However, the results of this study suggest that the benthic-centric perspective is more applicable to the hyporheic zone in the restored stream reaches studied here, which are representative of many agricultural, lowland streams in Central New York in slope, sinuosity, bed material, and history of degradation and restoration, although differences in K could alter these conclusions for other streams. The studied cross-vanes do not dramatically promote the potential processing of large amounts of pollutants in the hyporheic zone, and instead appear to contribute to stream health by creating heterogeneous patches of hyporheic exchange and biogeochemical cycling in the hyporheic zone. Secondary bedforms, which are created and maintained by cross-vane hydraulics, are mainly responsible for the observed patchiness and heterogeneity of HEF and redox chemistry across the streambed area in restored reaches, similar to microforms in models of *Kaser et al.* [2012].

5. Conclusions

We found that the 4 studied cross-vanes do promote HEF around themselves and around secondary bedforms maintained by structure hydraulics. However, we did not find the expected pattern of strong downwelling flux and oxidizing conditions upstream of the cross-vane and

strong upwelling flux and reducing conditions downstream of the cross-vane. Instead, we found a more limited zone of rapid downwelling and stream-like geochemistry 1-2 m upstream of cross-vanes, and variable flux and redox conditions downstream. In the downstream plunge pools, we observed shallow downwelling in areas affected by turbulence and momentum from the cross-vane drop, which has been modeled before but never measured to our knowledge. We also found that aspects of practical cross-vane construction, such as backfill and gaps, affect their ability to drive HEF, but are not mentioned in literature [e.g. *Doll et al.*, 2003]. In reaches of 45-55 m, total measured HEF volumes were ~0.4% of stream discharge.

Our results suggest that flux immediately around these cross-vanes performs less biogeochemical processing for the whole-stream ecosystem than anticipated, and has limited impact on benthic ecology because a small percentage of total bed area is affected. The low total reach-scale HEF volume also shows that restoration of these reaches did not lead to high rates of exchange, and the impact on total biogeochemical cycling and stream solute concentrations is not great. Although this is likely due to local conditions such as low K [*Hester and Gooseff*, 2011], our sites are fairly representative of agricultural, lowland streams in Central New York.

Secondary pool and riffle bedforms upstream and downstream of the cross-vanes did, however, create heterogeneous patches of flux and biogeochemical reactions in the streambed hyporheic zone. Although exchange flux magnitudes were generally low, shallow streambed pore water was mostly stream-derived, and distinct gradients of biogeochemical redox conditions were observed. Heterogeneous patches of upwelling and downwelling water, temperature, dissolved oxygen, and redox-sensitive solute concentrations in these reaches provide a diverse benthic environment in the hyporheic zone. We expect that cross-vanes that successfully

promote the creation of secondary pool-riffle morphology are more effective at hyporheic restoration than cross-vanes with large steps alone.

These conclusions show that hyporheic restoration in streams such as ours should focus more on creating small secondary bedforms and less on the size of large structures and the hydraulics immediately around them, unless it is to maintain or augment smaller bedforms. The expense of building large structures spaced 25-150 m apart might not be realized in significantly increased reach-averaged HEF and nutrient processing in many streams. We suggest that a better use of resources might be to concentrate on smaller, more closely spaced drivers of bedform heterogeneity, such as large woody debris, if restoration of the hyporheic ecosystem is the goal.

Acknowledgements

Kathleen McGrath, Jason Hamidi, and Martin Briggs assisted with field work. The authors would like to thank the Associate Editor, Audrey Sawyer, Erich Hester, and one anonymous reviewer for comments that improved an earlier manuscript. This work was supported by the National Science Foundation under Grant Nos. EAR-0911612 and EAR-0901480, a National Science Foundation Graduate Research Fellowship under Grant No. DGE-0750965, and the Syracuse Center of Excellence.

Tables

Table 1. Dimensions and other details of study reaches and cross-vanes.

Site name	Year of restoration	Dates of data collection	Vane width[a] (m)	US vane height[b] (m)	DS vane height[c] (m)	Water surface drop[d] (m)	Pool depth[e] (m)	Surveyed reach length (m)	Base-flow discharge (m^3 s^{-1})	Watershed area (km^2)
Ninemile A	2004	8/27-9/17, 2009 7/30-8/12, 2010	17	0.30	0.61	0.26	0.90	39 (2009) 55 (2010)	0.40 (2010)	177
Ninemile B	2004	9/19-10/17, 2009	14	0.30	0.44	0.22	0.91	34	-	177
Boyer	2006	6/22-7/6, 2010	8	0.00	0.29	0.21	0.53	69	0.03	19.3
Owego	2006	7/9-7/26, 2010	10	0.22	0.75	0.10	1.30	45	0.079	74.4

[a] approximate width of stone structure measured from bank-to-bank across plunge pool.
[b] difference between cross-vane top elevation and bed elevation immediately upstream of cross-vane, at point where water spills over.
[c] difference between cross-vane top elevation and bed elevation immediately downstream of cross-vane, at point where water spills over.
[d] difference between water surface elevation immediately upstream and downstream of cross-vane, at point where water spills over.
[e] difference between cross-vane top elevation (at point where water spills over) and lowest elevation of plunge pool.

Table 2. Mean and standard deviation (in parentheses) values of sediment thermal and textural properties used for modeling heat transport at all sites. Properties were estimated based on sediment observations at field sites and guidelines in *Lapham* [1989] and *Fetter* [2001].

Property	Symbol	Units	Value
Total porosity	n	dimensionless	0.20 (0.04)
Baseline thermal conductivity	λ_o	J s^{-1} m^{-1} $^oC^{-1}$	2.26 (0.31)
Thermal dispersivity	β	m	0.001 (0)
Volumetric heat capacity of sediment	C_s	J m^{-3} $^oC^{-1}$	2.09×10^6 (3.1×10^4)
Volumetric heat capacity of water	C_w	J m^{-3} $^oC^{-1}$	4.18×10^6 (2.1×10^4)

Table 3. Mean concentrations and standard deviations (in parentheses) of dissolved solutes for streambed pore water, stream water, and shallow riparian groundwater at the study sites.

	n	Cl mg l^{-1}	NO_3^- mg l^{-1}	SO_4^{2-} mg l^{-1}	Na^+ mg l^{-1}	NH_4^+ mg l^{-1}	K^+ mg l^{-1}	Mg^{2+} mg l^{-1}	Ca^{2+} mg l^{-1}	Fe μg l^{-1}	Mn μg l^{-1}	P mg l^{-1}	Si mg l^{-1}	Sr mg l^{-1}	DO mg l^{-1}
NMB pore water	38	39.3 (3.52)	0.33 (0.6)	8.3 (4.19)	21.9 (2.37)	0.71 (0.39)	2.44 (0.64)	16.9 (3.51)	37.8 (5.75)						
NMA'09 pore water	35	41.4 (6.18)	0.26 (0.36)	8.9 (4.13)	24.1 (2.59)	0.87 (0.8)	2.31 (0.26)	15.6 (2.81)	31.8 (4.98)						
NMA'10 pore water	48	48.8 (5.54)	0.46 (0.58)	10.4 (4.31)	28.7 (3.51)	0.68 (0.63)	2.55 (0.57)	17.8 (1.35)	37.1 (5.4)	8.2 (18.3)	129 (177)	32 (23)	2.76 (1.09)	119.5 (15.3)	2.9 (2.4)
NMA'10 stream	4	46.4 (0.47)	1.98 (0.04)	14.5 (0.16)	26.9 (0.22)	0.01 (0.01)	1.55 (0.08)	18.2 (0.11)	29.7 (5.16)	11.1 (7)	2 (1)	26 (4)	1.51 (0.05)	105.6 (12.2)	10.7
NMA'10 groundwater	1	39.7	0.13	5.1	24.8	0.22	1.66	9.9	44.6	29.8	123	147	1.86	526.5	
Owego pore water	52	25.3 (1.87)	0.41 (0.39)	5.6 (1.95)	13.4 (0.82)	0.57 (0.61)	1.59 (0.35)	6.2 (0.86)	29.6 (3.39)	32.5 (46.7)	648 (1013)	33 (31)	1.26 (0.76)	55.4 (9.4)	3.2 (2.8)
Owego stream	4	25.9 (0.04)	1.02 (0.1)	6.9 (0.02)	13 (0.02)	0.02 (0)	1.02 (0.01)	6.3 (0.01)	26.6 (0.86)	9.8 (4)	3 (1)	11 (3)	1.07 (0.26)	49.2 (1.4)	11.8
Owego groundwater	1	28.3	0.01	5.4	15.9	0.01	1.98	5.7	29.3	5.8	12	13	0.67	46.9	
Boyer pore water	40	18.8 (0.13)	0.68 (0.13)	6 (0.11)	10.7 (0.07)	0.01 (0.01)	1.27 (0.03)	3.7 (0.14)	12.3 (1.48)	1.4 (1.3)	4 (13)	23 (4)	0.97 (0.31)	26.3 (2.5)	
Boyer stream	4	18.9 (0.16)	0.55 (0.01)	6.1 (0.03)	10.7 (0.06)	0.01 (0)	1.23 (0.02)	3.8 (0.01)	12.9 (1.03)	1 (0.6)	0.4 (0.1)	18 (6)	1.33 (0.02)	29.4 (1.3)	

Table 4. Results of the principle component analysis (PCA) at each study site. Shown are the Eigenvector loadings, Eigenvalue, and percent of explained variance for the first 3 principle components.

Principle Component	Eigenvalue	explained variance (%)	Eigenvector loadings						
			NO_3^-	SO_4^{2-}	NH_4^+	K^+	Mn	Si	DO
Ninemile A 2009:									
PC1	2.78	69.6	0.45	0.53	-0.54	-0.47			
PC2	0.55	13.8	0.81	-0.41	0.40	-0.15			
PC3	0.51	12.7	0.37	0.35	-0.09	0.86			
Ninemile A 2010:									
PC1	4.21	60.5	0.41	0.38	-0.40	-0.41	-0.34	-0.41	0.28
PC2	0.99	14.3	0.41	-0.29	0.37	-0.08	0.22	0.16	0.73
PC3	0.74	10.6	0.17	0.26	-0.35	-0.32	0.71	0.38	-0.15
Ninemile B:									
PC1	1.89	47.2	0.63	0.52	-0.56	-0.12			
PC2	1.40	35.1	-0.06	-0.49	-0.37	-0.78			
PC3	0.43	10.8	0.71	-0.22	0.64	-0.22			
Owego:									
PC1	3.89	57.4	0.44	0.41	-0.38	-0.45	-0.38	-0.31	0.22
PC2	1.23	18.2	0.32	-0.32	0.17	-0.25	0.06	0.55	0.64
PC3	0.80	11.8	0.13	0.37	0.61	-0.01	0.47	-0.42	0.27

Table 5. Pearson correlation coefficients between individual square-root-transformed solutes and the first principle component (the redox score). All p-values for tabulated correlations are less than 0.001.

Site	NO3-	SO42-	NH4+	K+	Mn	Si	DO
Ninemile A 2009	0.75	0.88	-0.91	-0.78			
Ninemile A 2010	0.85	0.78	-0.82	-0.84	-0.69	-0.84	0.59
Ninemile B	0.87	0.73	-0.77	n.s.[a]			
Owego	0.88	0.81	-0.69	-0.87	-0.79	-0.62	0.57

[a] n.s. = correlation not significant; p=0.34.

Figures

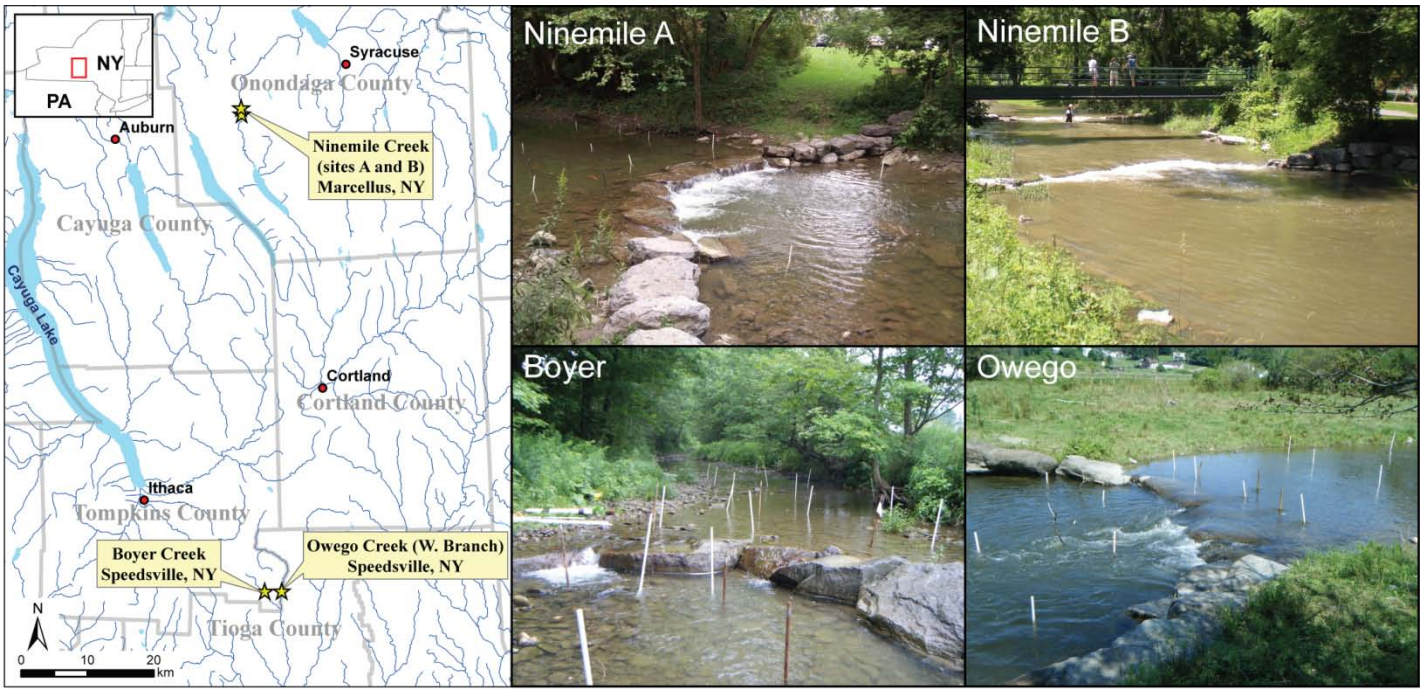

Figure 1. A regional map and photographs of cross-vanes in Central New York State that were selected for this study.

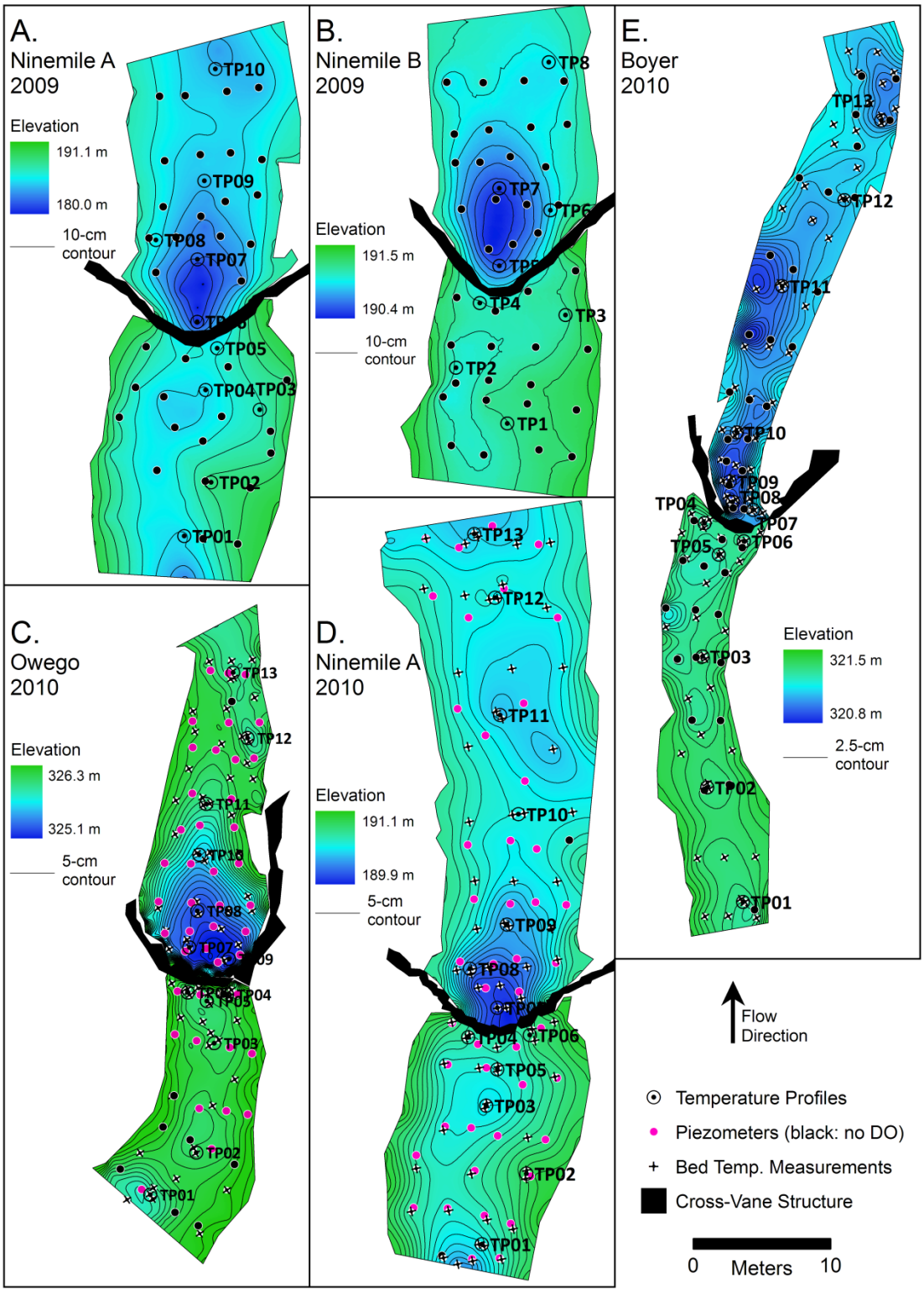

Figure 2. Surveyed bed topography, cross-vane structures, and locations of installed instrumentation in study reaches. Piezometers shown with black circles instead of pink are those at which DO was not measured.

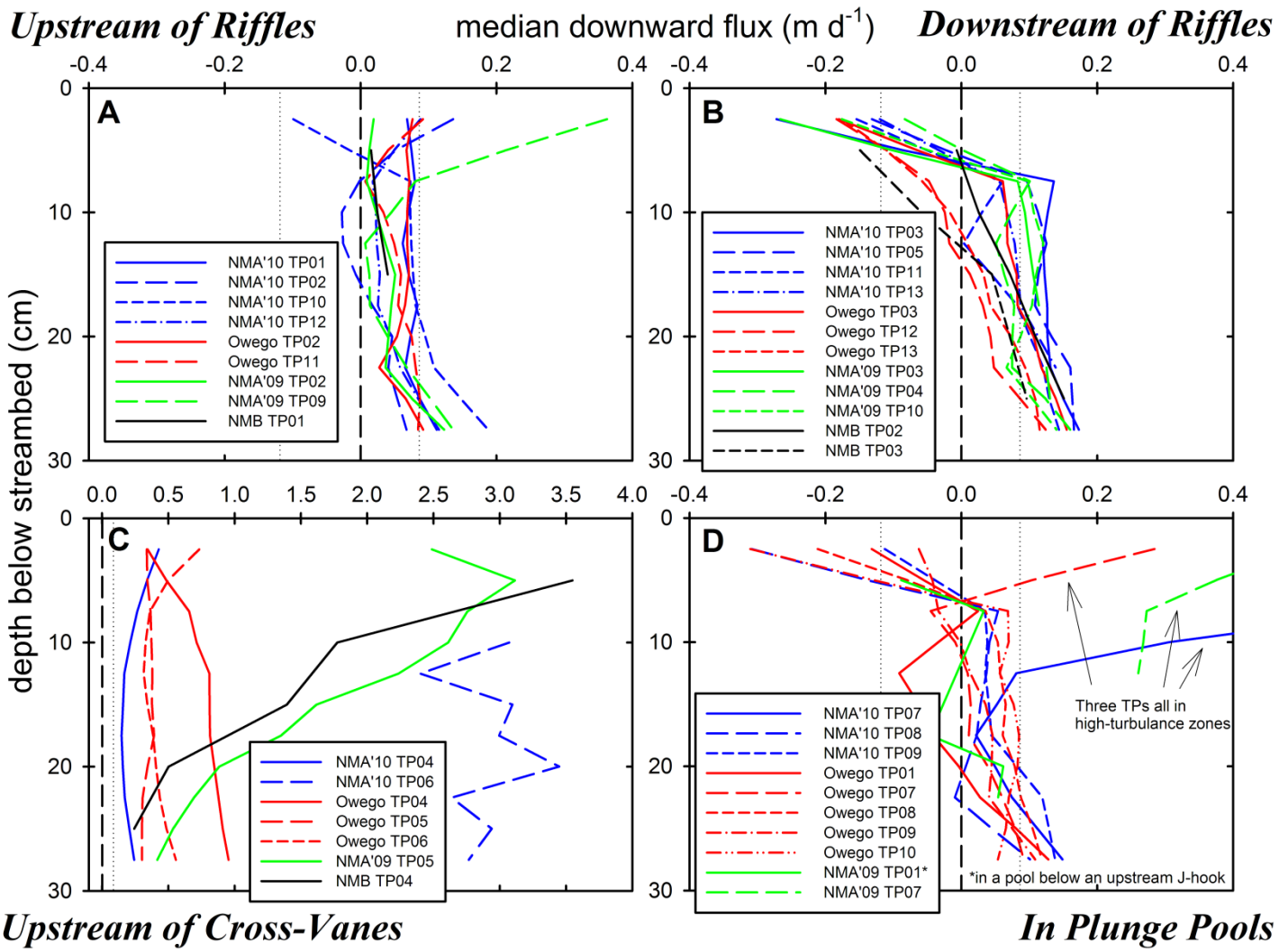

Figure 3. Median flux versus depth profiles for TP locations, divided into four categories based on morphology: (A) upstream of riffles and in the glides of shallow pools, (B) downstream of riffles and in the runs of shallow pools, (C) directly upstream of cross vanes, and (D) in the plunge pools downstream of cross-vanes. Flux values are the median flux estimate throughout the study period. Depth values are the midpoints between sensor pairs (with Δz values of 5 or 10 cm). Confidence intervals for each TP are not shown for clarity; instead, only the width of uncertainty around zero is shown by dotted lines. Included are TPs from all sites except Boyer; NMA'09 TP08 is not shown because of its location near the water's edge (median flux for the shallowest sensor pair at TP08 was -0.16 m d^{-1}).

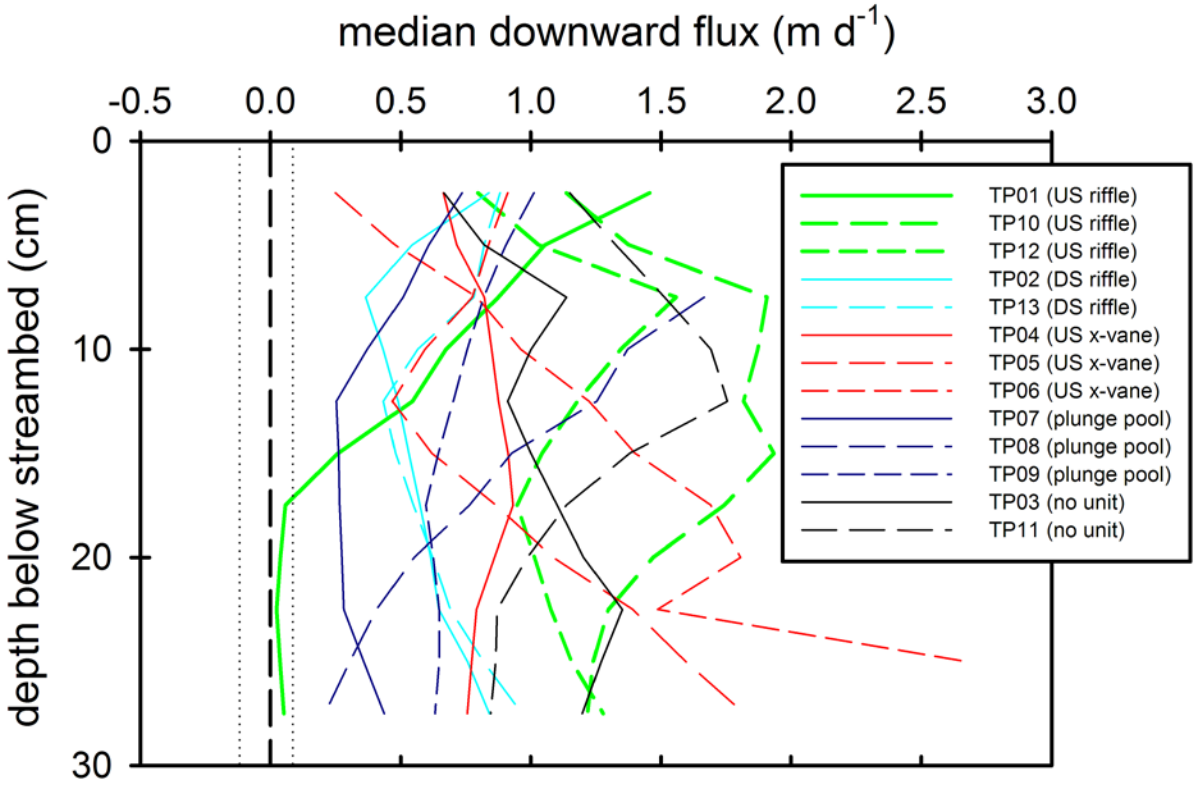

Figure 4. Median flux versus depth profiles for TP locations at Boyer, similar to Figure 3. Legend indicates the morphological category assigned to each TP: upstream of riffle, downstream of riffle, upstream of cross-vane, in plunge pool, or no distinct unit identified.

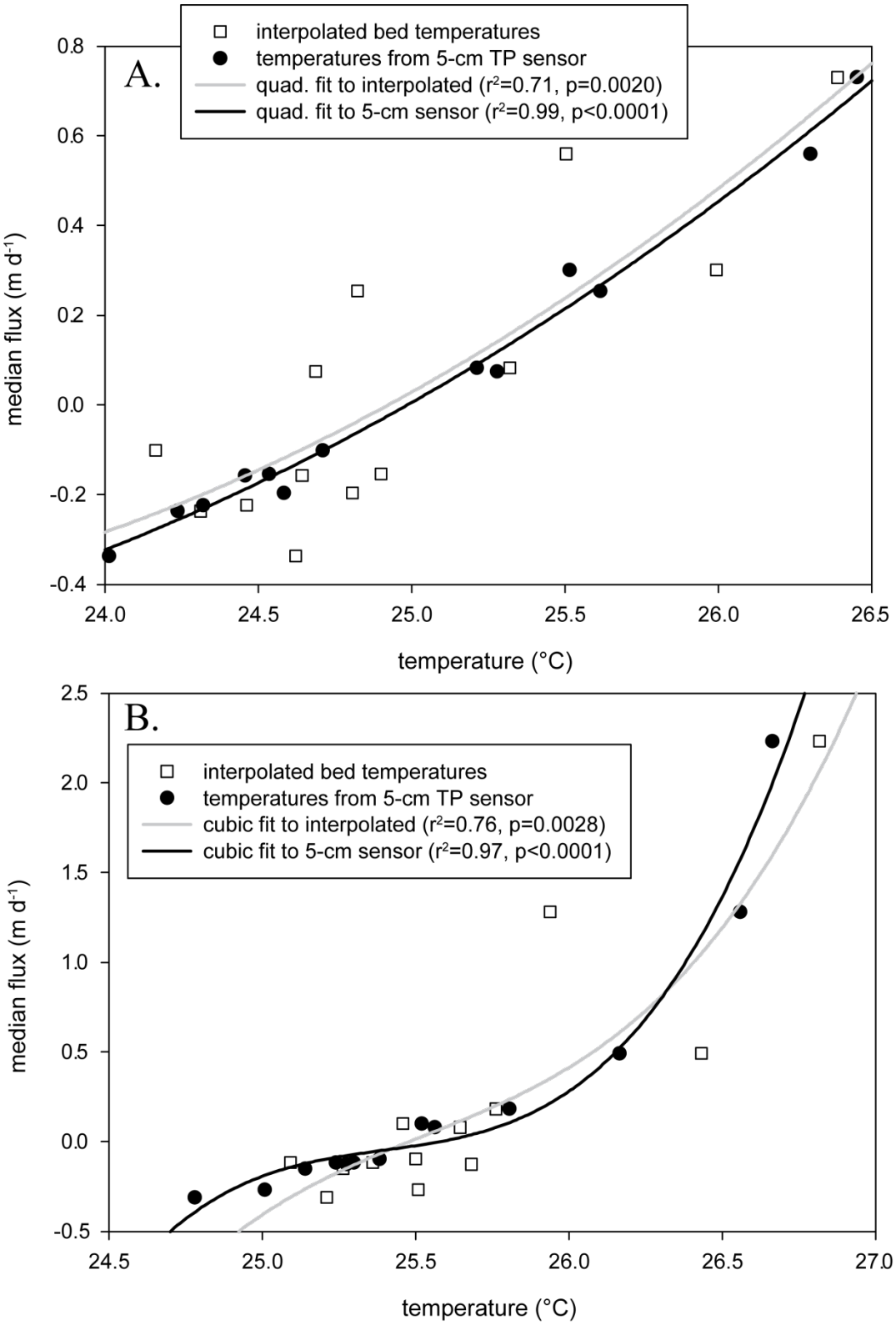

Figure 5. Median flux values at TP locations regressed against bed temperature measured at a single point in time for (A) Owego and (B) NMA'10. See discussion in text, Sections 3.2 and 4.1.

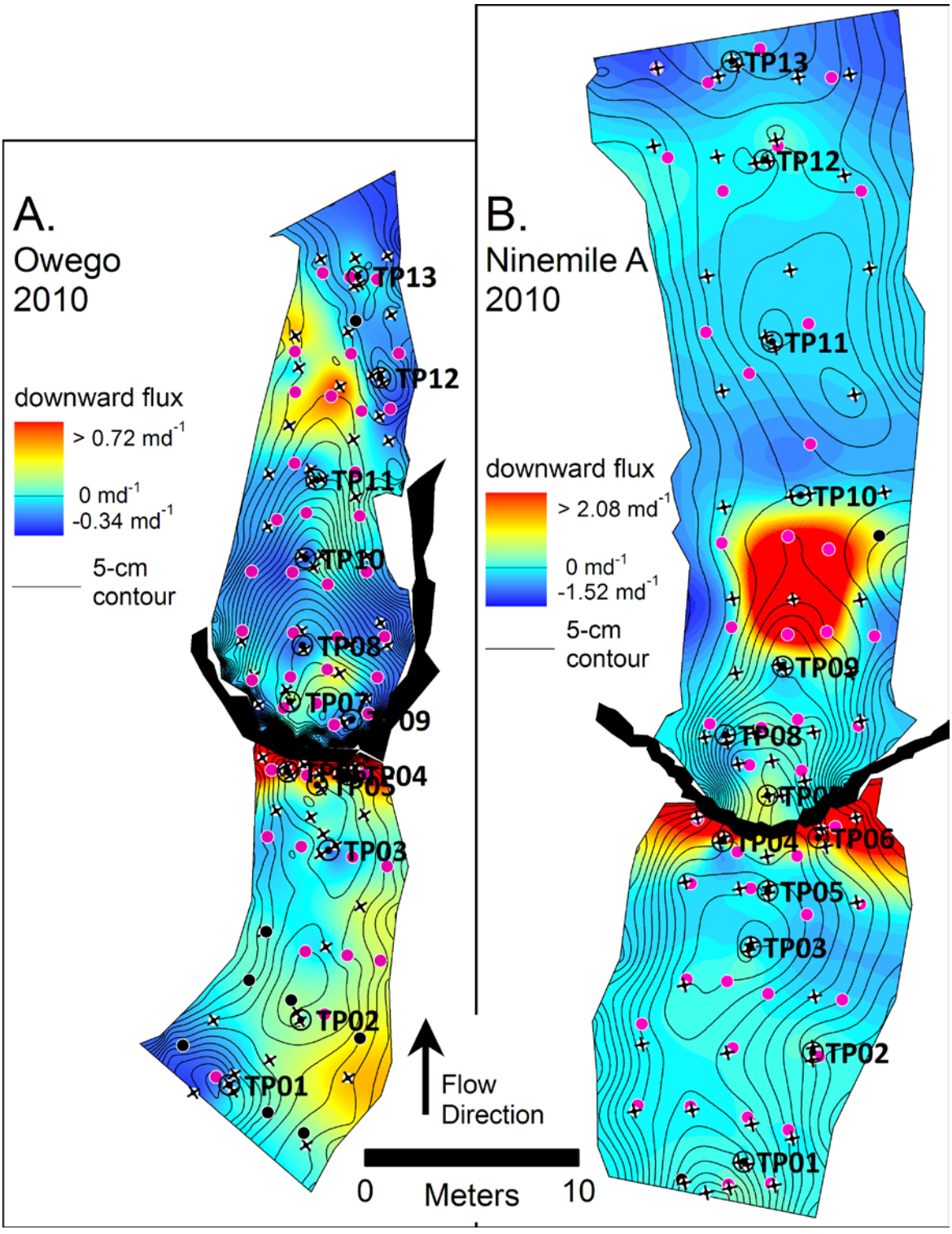

Figure 6. Spatial patterns of vertical water exchange flux (units of m d^{-1}, positive downward) inferred from measured bed temperatures using the regression equations from Figure 5. Interpolation between bed temperature measurement locations was performed using the spline method. The color cyan represents zero flux. Symbols as described in Figure 2.

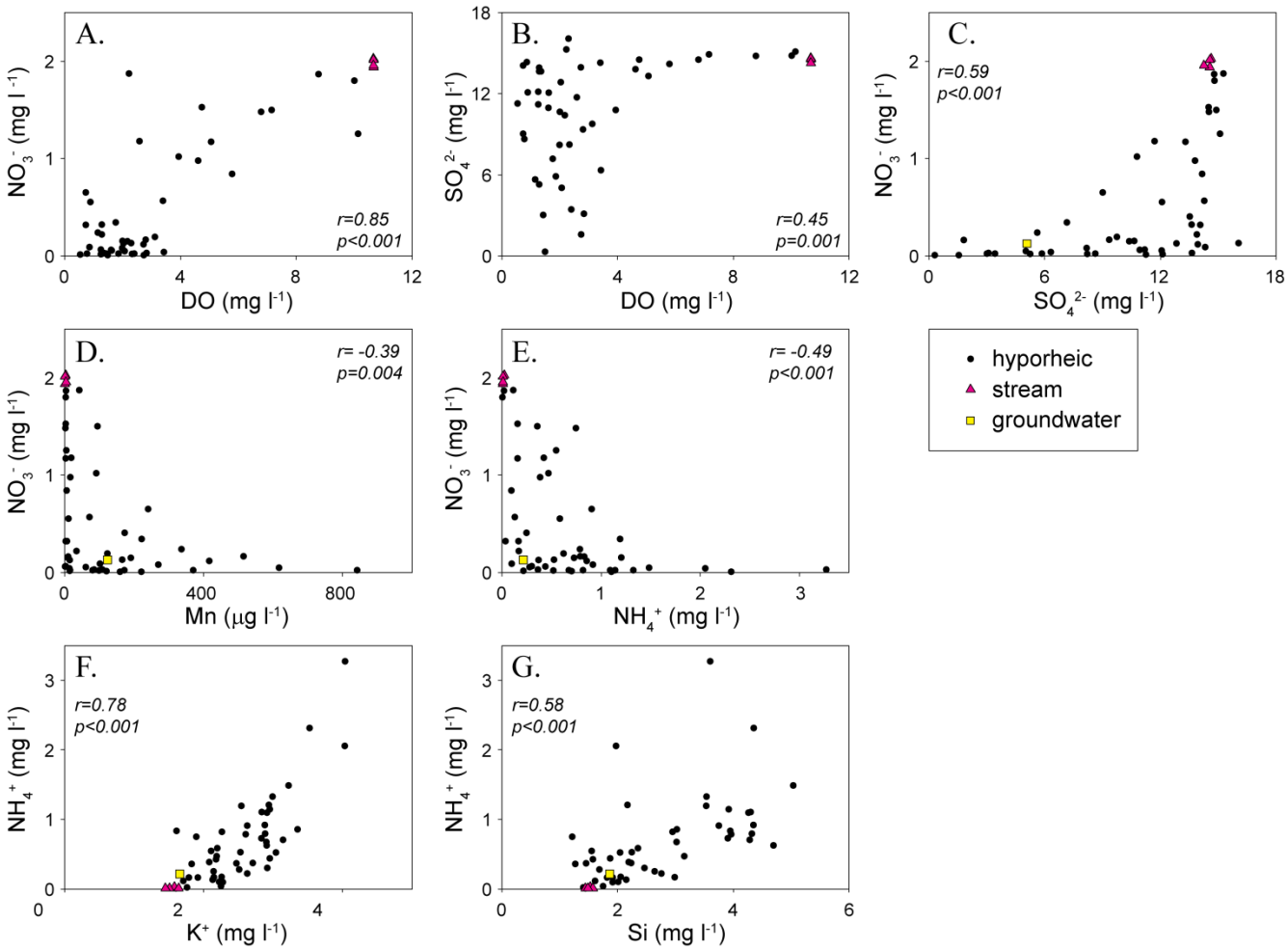

Figure 7. Bivariate plots of redox-sensitive solute concentrations in the stream, shallow groundwater, and streambed at Ninemile A (NMA'10). Patterns are typical of all sites except for Boyer.

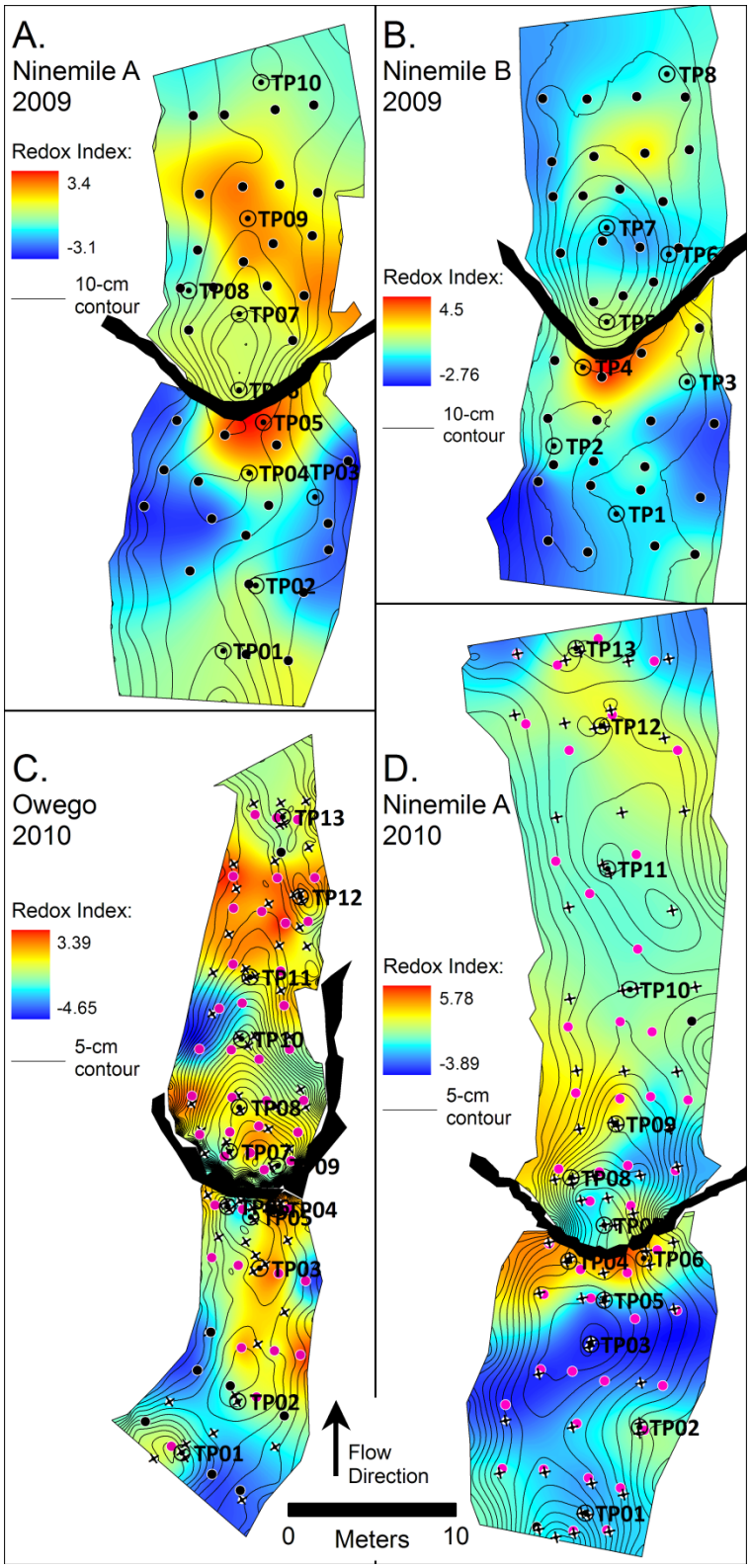

Figure 8. Spatial patterns of the redox index (first principle component) in the streambed. Positive values (warm colors) indicate oxidizing conditions, and negative values (cold colors) indicate reducing conditions. Interpolation between piezometer locations was performed using the spline method. Symbols as described in Figure 2.

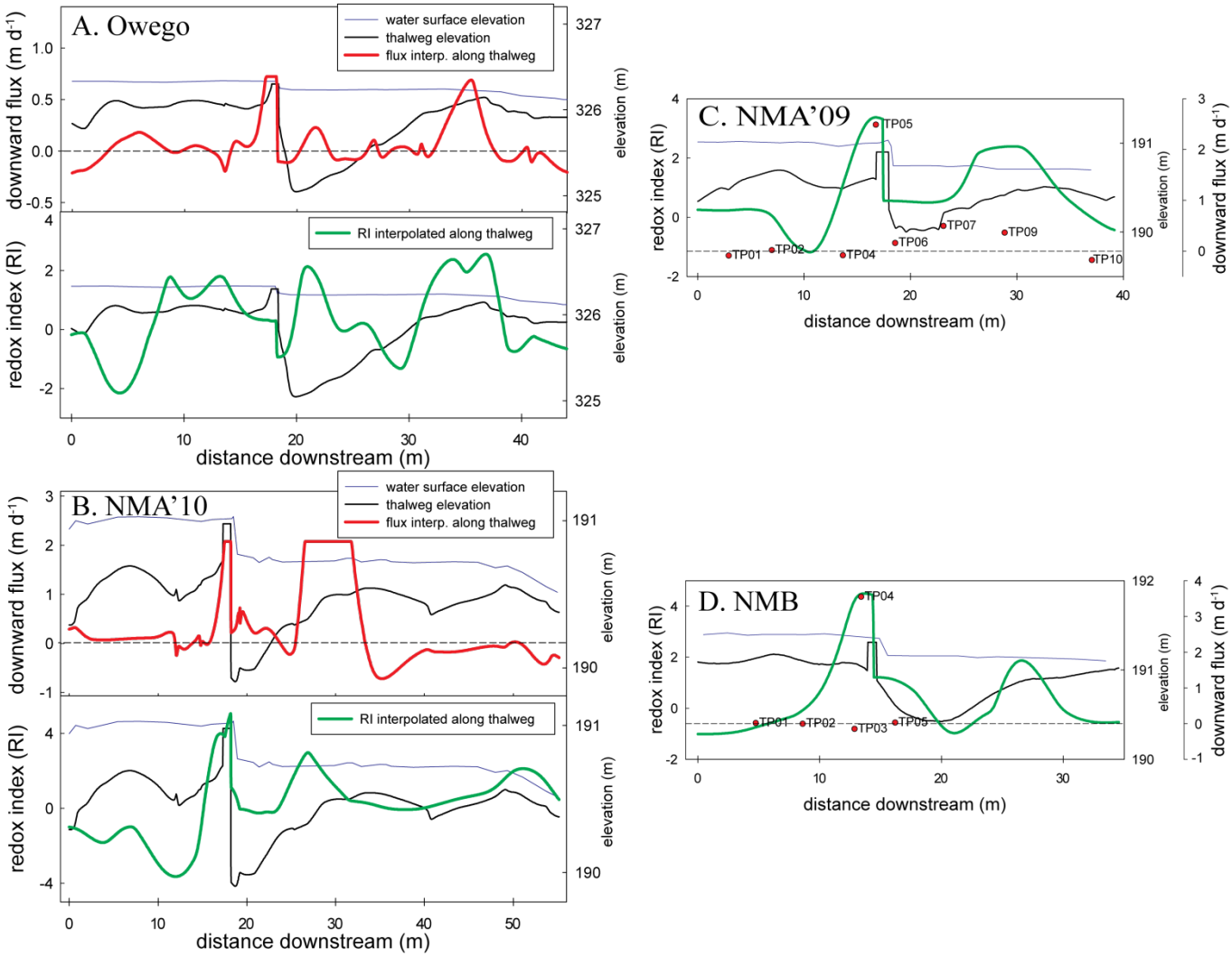

Figure 9. Downward vertical flux rates and interpolated values of the redox index down the thalwegs at (a) Owego, (b) Ninemile A 2010, (c) Ninemile A 2009, and (d) Ninemile B. For 2010 sites, flux is reported as the value of the interpolated flux surface derived from the empirical relationship between flux and bed temperature measurements (Figure 6). The maximum reported flux values at the 2010 sites are limited to 2.08 m d^{-1} (NMA'10) and 0.72 m d^{-1} (Owego), but these should only be interpreted as lower bounds on the true flux magnitude in these portions of the profiles (see text in Section 3.2). For the 2009 sites, flux is reported as calculated from the shallowest pair of temperature sensors at TP locations.

Chapter 3.

Sources and pathways of stream generation in high-elevation meadows of a tropical glacierized mountain range

Chapter 3 is prepared as a manuscript in the format of *Hydrological Processes*, to be submitted in 2014 as:

Gordon, R.P., L.K. Lautz, B.G. Mark, J.M. McKenzie, and D. Chavez, Sources and pathways of stream generation in high-elevation meadows of a tropical glacierized mountain range.

1. Introduction

Alpine (highland or high-elevation) watersheds are one of the most important sources of fresh water in the world, and in semi-arid regions, mountain drainages supply a vastly disproportionate share of the water used by populations downstream (Barnett *et al.*, 2005; Viviroli *et al.*, 2007). More than 80% of the water supply in the semi-arid tropics and subtropics originates in mountain watersheds, affecting over half of humanity (Messerli, 2001). Alpine watersheds are dominated by snow and ice melt processes, which are highly impacted by global climate change (Barnett *et al.*, 2005). Fossil water resources preserved in glaciers and snowpack are being permanently lost due to world-wide melt and recession (Barry and Seimon, 2000; Barnett *et al.*, 2005; Barry, 2006; Kaser *et al.*, 2006; Nogues-Bravo *et al.*, 2007; Vuille *et al.*, 2008; Milner *et al.*, 2009). Both the timing and the total quantity of water discharge from glacierized mountain areas will be affected by continuing climate change; specifically, dry season discharge will be reduced as water supplied by melting ice and snow is minimized, and both seasonal and inter-annual runoff will become more variable as precipitation becomes the dominant control on streamflow (Mark and Seltzer, 2003; Barnett *et al.*, 2005; Milner *et al.*, 2009). The effects of glacier-loss on seasonal runoff variation will be felt most acutely in the tropics, because mid-latitude glaciers amplify seasonal runoff variation while tropical glaciers smooth the variation, an effect that decreases significantly as glaciers recede (Kaser and Osmaston, 2002; Kaser *et al.*, 2003; Mark and Seltzer, 2003).

As seasonal discharge patterns shift and the buffering capacity of meltwater released from ice and snow is lost, the storage capacity of alpine landforms has an increasing influence on seasonal and inter-annual streamflow (Barnett *et al.*, 2005). Although groundwater storage in glacierized valleys is usually assumed negligible in watershed modeling studies (e.g. Kaser *et al.*,

2003; Caballero *et al.*, 2004), common proglacial features such as meadows, moraines, talus slopes, and fans may store substantial groundwater that is released to streams at different rates in different seasons, with consequences for hydrologic modeling and water resource management (Clow *et al.*, 2003). Understanding sources of streamflow and the storage capacity of natural reservoirs in proglacial watersheds is important for alpine hydrology, and has potential impacts on the way hydrologists model alpine processes and predict future impacts of deglaciation on water resources; however, very little research has explored the subsurface processes contributing to streamflow in glacierized tropical catchments.

Little is known about the role of groundwater in alpine watersheds (Hood *et al.*, 2006). Using hydrograph separation or end-member mixing analyses, several studies (Sueker *et al.*, 2000; Huth *et al.*, 2004; Liu *et al.*, 2004; Brown *et al.*, 2006) have estimated a wide range for the percent contribution of alpine groundwater to streamflow at the catchment scale. At a more detailed scale, a series of studies (Hood *et al.*, 2006; Roy and Hayashi, 2009; McClymont *et al.*, 2010; Langston *et al.*, 2011; McClymont *et al.*, 2011; Muir *et al.*, 2011) have used a variety of geophysical and more traditional hydrologic techniques to investigate groundwater storage and interactions among a number of alpine geomorphologic units including proglacial lakes, moraines, talus slopes, springs, and a small meadow, all in a single watershed in the Canadian Rocky Mountains. McClymont et al. (2010) estimated the storage and residence time of a talus slope-meadow complex, suggesting future studies should investigate different types of alpine meadow complexes and their impacts on watershed runoff and water quality.

No hydrologic research has been published on the unique meadows, talus slopes, or moraines that are likely storage reservoirs in tropical glacierized catchments (Mark *et al.*, 2005; Mark and McKenzie, 2007; Baraer *et al.*, 2009). As a result, proglacial hydrologic models

frequently oversimplify the contribution of groundwater to watershed budgets, either ignoring it altogether or treating it like a lumped black box, calibrated empirically (Hood *et al.*, 2006). Interactions between groundwater and streamflow have likewise been neglected or decoupled in simulations of climate change effects on alpine hydrology (Huntington and Niswonger, 2012). Studies that have estimated the relative contributions of groundwater and glacial melt to total stream discharge have done so using mass-balance and geochemical end-member mixing approaches at the basin and sub-basin scale (e.g. Wagnon *et al.*, 1998; Mark *et al.*, 2005; Baraer *et al.*, 2009). This type of study is very useful in large, poorly instrumented mountain regions, but necessarily ignores potential recharge of proglacial aquifers by glacial meltwater and transformations in stream chemistry due to hydrologic turnover, potentially underestimating the importance of glacial recession. We expect that interactions between groundwater in geomorphic reservoirs and surface water in proglacial streams affect the quantity, timing, and chemistry of downstream discharge. Although complex interactions between surface water and groundwater have been recognized for some time (e.g. Winter *et al.*, 1998), the interactions between groundwater and mountain streams at the reach scale (10 m to 1 km) is a recent subject of interest (e.g. Covino and McGlynn, 2007; Bencala *et al.*, 2011; Covino *et al.*, 2011). Understanding these groundwater-surface water interactions is important for making accurate estimates of glacial meltwater contributions to the total hydrologic budget, and for our ability to make predictions about future surface water resources.

In this paper, we report the results of an investigation of streamflow generation, groundwater-surface water interactions (GWSWI), and water sources in two proglacial high-elevation meadow complexes in the tropical Andes of Peru (Figure 1). The objective of this research is to determine the impacts of high-elevation meadows and other proglacial geomorphic

features on surface water quantity and quality in glacierized tropical mountain catchments, with a focus on streamflow generation and GWSWI between streams and proglacial subsurface aquifers during the dry season (austral winter). We studied the GWSWI between proglacial stream channels and shallow aquifers in wet meadows and adjacent moraine, talus, and fan features by measuring exchange rates between groundwater and stream channels and by investigating the isotopic and geochemical contributions, mixing, and transformations of source waters.

2. Site Description

The Cordillera Blanca of Peru (Figure 1) is the most glacierized tropical mountain range in the world. It is also the largest and highest mountain range in Peru, with a peak at Huascarán (6,768 masl). In the 1970s, 723 km^2 of glacier area in the Cordillera accounted for 40% of Peru's glacial volume (Ames *et al.*, 1989), but glaciers have since been retreating rapidly (Mark *et al.*, 2010). The glacierized catchments on the western side of the range supply two-thirds of the dry-season discharge of the Santa River (Rio Santa) (Mark *et al.*, 2005), which flows to the Pacific Ocean. The Rio Santa powers multiple hydroelectric plants, which supply electricity to cities, villages, polymetalic mining operations, and heavy industry in the watershed. Water from the Rio Santa is also used for small-holder agriculture in the upper basin, and to irrigate commercial agriculture operations on the arid coast. The upper Rio Santa basin adjacent to the Cordillera Blanca is known as the Callejón de Huaylas, with an area of 4,900 km^2 and a population of 267,000 (Mark *et al.*, 2010). The climate in the Callejón de Huaylas is semi-arid with a highly seasonal precipitation regime; 80% of precipitation falls in the austral summer, between October and May, and totals 800 to 1200 mm yr^{-1} (Bury *et al.*, 2011). The average annual temperature, dependent on elevation, is between 0 and 9 degrees C, with annual differences in monthly mean temperature much less than the daily variation (Kaser *et al.*, 1990).

Glacial meltwater provides approximately half of the dry-season stream discharge in some Cordillera Blanca valleys (Mark *et al.*, 2005), and meltwater buffers streamflow throughout the range, making it less susceptible to inter-annual variations in precipitation (Mark and Seltzer, 2003; Barnett *et al.*, 2005; Milner *et al.*, 2009). The remainder of dry-season streamflow is supplied by groundwater. In the future, when glacier loss has reduced the influence of meltwater on streams of the Cordillera Blanca, groundwater discharge will be the sole dry-season source of stream water for irrigation, municipalities, and hydropower generation in the Rio Santa watershed, putting livelihoods and economies at risk (Bury *et al.*, 2011; Baraer *et al.*, 2012). A better understanding of the dynamics of alpine groundwater, including sources, pathways, and storage, is therefore vitally important for future planning.

Most major glaciers on the western side of the Cordillera Blanca are situated above low-gradient meadows or wetlands locally called “pampas” or “bofedales” (Figures 1b-c), and glacial meltwater typically flows via a stream network through a complex landscape of lakes, moraines, fans, and talus slopes around these meadows (Figure 2). Kaser et al. (2003), in a study of glacial water storage in the Cordillera, noted that watersheds with “extensive, swampy plains and plateaus” appear to have additional storage capacity that complicates the effects of glaciers on discharge. Mark and McKenzie (2007) suggested that groundwater flow in the Cordillera Blanca valleys is derived from stored wet-season precipitation, and proposed that proglacial wet meadows (pampas) are likely groundwater reservoirs but require further study. The meadows are generally understood to form by paludification in low-gradient valley bottoms or from lakes dammed by moraines or rockfalls (Earle *et al.*, 2003; Squeo *et al.*, 2006), and are ecologically characterized by a community of grass, sedge, and wetland plants in hydric soils. We therefore expect the meadows to be underlain by layers of peat and lacustrine deposits over buried talus,

till, and bedrock.

For this study, we performed investigations at two meadow complexes in two glaciated valleys on the western side of the Cordillera Blanca, which were chosen due to their contrasts in stream discharge, meadow extent, adjacent geomorphic features, bedrock geology, and position relative to glacier termini. The first site included a meadow and proglacial stream system at the head of the Damanda branch of the Llanganuco valley. The meadow system (9.017ºS, 77.604ºW, 4370 masl) and adjacent geomorphologic features cover approximately 0.5 km^2 and include terminal moraine deposits, an alluvial/debris fan, rock falls, several groundwater springs at different geomorphic positions, and over 2 km of stream channel in tributaries draining two proglacial lakes (Figures 1b and 2a). This meadow complex was chosen because of its small size and low stream discharge, the presence of terminal moraines, talus slopes, and a fan that all terminate at the meadow margins, and the fact that the groundwater system is relatively contained in a hanging glacial cirque with a single, measurable outflow over bedrock. Modeling studies predict that the Llanganuco catchment could experience over a 60% decline in dry season discharge if glaciers melt completely (Baraer *et al.*, 2012). Elevation of the watershed studied at Llanganuco ranged between 4340 m at the outflow and 6108 m at the peak of Chakrarahu, and is 12.5 km^2 in area, 1.5% of which is meadow, and 43% glacier (estimated from Quickbird-2 satellite imagery, 2003). The lowest glacier terminus is at approximately 4750 m elevation. Discharge measured at the outflow was 115 L/s during our study. The majority of bedrock in the Llanganuco valley is granodiorite/tonalite of the Miocene Cordillera Blanca batholith (Wilson *et al.*, 1995); however, although not mapped as such due to ice cover in the Twentieth Century (Wilson *et al.*, 1995), parts of the valley above Laguna Broggi appear to be composed of the Upper Jurassic Chicama formation, a unit of marine black shales high in pyrite and other sulfide

minerals that is more commonly found on the eastern side of the Blanca (Love *et al.*, 2004).

The second site was a 1.5-km long reach of the Quilcay stream in the middle of the Quilcayhuanca valley, upstream from an old water divertment named Casa del Agua (9.465ºS, 77.379ºW, 3905 masl) (Figures 1c and 2b). This site was chosen because the valley is typical of the long, hanging valleys of the Blanca, where a series of wet meadows contained behind moraines in the upper valley are situated above larger and more continuous valley-fill grassland with very low slope (1.5º) but steep talus-mantled valley walls. The studied stream reach at its upper end meanders through a wet meadow at approximately 3970 m elevation, and then flows through a valley-crossing moraine, at which point it is incised in a narrow gully. After dropping ~50 m, it flows out of the moraine into a lower meadow and passes through the Casa del Agua flume. The studied watershed is 69 km^2 in area, with a highest point of 6309 masl at the peak of Chinchey; 3.0% of the watershed is meadow, and 26% glacier (estimated from ASTER satellite imagery, 2007). The lowest glacier terminus is at approximately 4700 m elevation. Discharge measured at Casa del Agua was 1180 L/s during our study. Bedrock of the Quilcayhuanca valley is again primarily Cordillera Blanca granodiorite, but with substantial Chicama formation found in the upper glacierized valley, above our study site (Love *et al.*, 2004).

Comparing the two sites, Llanganuco is characteristic of the proglacial geomorphology at the headwaters of westward-flowing Cordillera Blanca valleys, and Quilcayhaunca is characteristic of the extensive mid-valley meadow systems. By studying them together we can make conclusions about how these different components of the landscape function and influence water resources.

3. Methods

3.1 Tracer dilution gauging

Each studied stream was divided into reaches approximately 100-200 m long for dilution gauging and water balance experiments using water tracers, which were measured and mapped using handheld GPS and aerial photography. Because all the valleys identified for this study are not continuously gauged, tracer dilution gauging is the only practical method available to directly quantify stream discharge. By performing multiple dilution tracer experiments at different scales over several reaches throughout the study areas, we were able to calculate not only net changes in discharge over distance, but also gross losses and gains of stream water—which can occur concurrently—to and from underlying aquifers over stream reaches of interest. Tracer tests used the fluorescent dye Rhodamine WT (RWT) to label stream water and estimate changes in stream discharge and gross water gains and losses. RWT is a commonly used tracer for measuring water flow, and is considered conservative in most lotic waters (Stream Solute Workshop, 1990). Except where noted, the tracer tests were conducted with the "double-slug" (two instantaneous slug injections) channel water balance method, described in the following paragraph and Figure 3, after Harvey and Wagner (2000) and Payn et al. (2009).

First, short mixing lengths (tens of meters) are identified upstream of the top and bottom of each stream reach based on field assessment and criteria of Day (1977). A tracer mass or slug M_D is released one mixing length above the downstream end of the reach, and tracer concentration through time, $C_D(t)$, is measured with a calibrated fluorometer (C3, Turner Designs, Sunnyvale, CA) as the tracer pulse passes the downstream end of the reach. Using the dilution gauging method (Day, 1977; Kilpatrick and Cobb, 1985), discharge at the downstream end of the reach, Q_D, is calculated from the observed tracer concentration breakthrough curve:

$$Q_D = \frac{M_D}{\int_0^T C_D(t)dt} \quad (1)$$

where T is the time at which concentration returns to zero after the tracer pulse passes. Next, a second slug of tracer, M_U, is released one mixing length above the upstream end of the reach, and its concentration is measured using a second fluorometer at the upstream end, yielding $C_U(t)$. Q_U, discharge at the upstream end of the reach, is calculated from M_U and $C_U(t)$, similar to Q_D in Equation 1. In this paper, Q values are reported in units of L s^{-1}. Finally, the tracer mass M_U, after having traveled the entire reach, is observed passing the downstream fluorometer, yielding $C_{UD}(t)$. The change in discharge over the reach, or net gain, $\Delta Q = Q_D - Q_U$, is the net subsurface inflow to the stream or, if negative, the net loss or outflow of stream water to the subsurface. The recovered mass, M_{rec}, which is the mass of tracer observed at the downstream end of the reach, is calculated as

$$M_{rec} = Q_D \int_0^T C_{UD}(t)dt \quad (2)$$

The mass of tracer that is lost over the reach due to exchange of tracer-labeled water with the subsurface is $M_{loss} = M_{rec} - M_U$, which allows a range for the gross loss of labeled water to be calculated. The minimum gross loss (Case 1 in Figure 3, in which all losses occur before all gains) is given by $Q_{loss,min} = Q_U M_{loss} / M_U$. The maximum gross loss (Case 2 in Figure 3, in which all losses occur after all gains) is given by $Q_{loss,max} = Q_D M_{loss} / M_{rec}$. The true value of Q_{loss} must be between these two values. The range of gross gains of unlabeled water into the stream is then simply given by the net gain minus the range of losses, $Q_{gain,min} = \Delta Q - Q_{loss,min}$, and $Q_{gain,min} = \Delta Q - Q_{loss,min}$. In this paper, ΔQ, Q_{loss} and Q_{gain} values are variously reported as discharges (L s^{-1}), discharges as a fraction of Q (unitless), discharges normalized by reach length (L s^{-1} km^{-1}), or discharges as a fraction of Q normalized by reach length (km^{-1}); all of these unit conventions are useful for comparing reaches with varying discharges and/or lengths (Covino *et*

al., 2011). The double-slug method relies on several assumptions, including perfect mixing of the tracer into stream water over the mixing lengths, steady discharge through time, and complete mass recovery (no loss of tracer) over the short mixing reaches (Harvey and Wagner, 2000; Payn *et al.*, 2009). Schmedel et al. (2010) quantified uncertainties in the channel water balance method due to violations of these assumptions, as well as uncertainty in data collection and measurement, and calculated a mean error of approximately 8% in measurements of Q, ΔQ, and gross gains and losses, almost all of which was due to incomplete mixing of the tracer.

At Llanganuco, double-slug experiments were conducted over nine stream reaches in the meadow (labeled A-G and X-Y in Figure 4), and three additional single-slug experiments were made to measure the surface discharge from each of the two lakes and discharge off of the moraine onto the meadow. A single-slug experiment is a simple dilution gauging measurement, in which one tracer mass is released and observed one mixing reach downstream, yielding a single discharge measurement by Equation 1. Experiments were conducted on July 10-11 (reaches A-F) and July 24-25 (all others) of 2012, during the southern tropical dry season. At Quilcayhuanca, double-slug experiments were conducted over seven stream reaches (labeled 1-7 in Figure 2b) in the upper meadow, over the moraine, and in the lower Casa del Agua meadow, all on July 16-17, 2012. Due to data loss, $C_U(t)$ for Reach 7 is unavailable, so only Q_D and $Q_{loss,max}$ can be calculated for that most upstream reach.

3.2 Synoptic sampling and analyses

Synoptic sampling of glacial lake, channel, spring, and shallow well water was performed on July 16-17, 2011 and July 22-25, 2012 at Llanganuco (Figure 2a) and July 26, 2012 at Quilcayhuanca (Figure 2b), all during dry seasons. Synoptic sampling captures the variation in geochemistry among end-members and mixed waters along stream profiles. In 2011, the

Llanganuco main channel up to Laguna Broggi was sampled approximately every 200 m, tributaries to the channel were sampled near their confluence, and water was collected from springs at the base of the moraine, debris fan, and talus slopes, as well as artesian springs in the meadow itself. Additionally, channel and spring samples were collected from farther down the Demanda valley from the study area shown in Figure 2a. In 2012, the Llanganuco campaign was repeated in the shown study area only, with channels sampled every 100 m, shallow wells sampled, and the Laguna 69 tributary and lake sampled (Figure 2a). At Quilcayhuanca, the main channel, one spring at the base of the moraine, and four groundwater wells were sampled (Figure 2b).

Sampled waters were measured for pH and conductivity in the field using a hand-held multi-probe (Multi 340i, WTW, Weilheim, Germany) and samples were filtered in the field using 0.45 μm nylon filters. Samples were later analyzed for the ions fluoride, chloride, bromide, nitrate, sulfate, sodium, potassium, magnesium, and calcium by ion chromatography (Dionex ICS-2000, Thermo Fisher Scientific, Waltham, MA) and for stable isotopes of water ($\delta^{18}O$ and $\delta^{2}H$) using a cavity ring-down spectrometer (L2130-i Water Isotope Analyzer, Picarro, Santa Clara, CA). Bicarbonate and carbonate concentrations were determined by charge balance and pH. Concentrations of fluoride, chloride, nitrate, sulfate, sodium, potassium, magnesium, and calcium in surface water samples at Llanganuco were incorporated into a principle component analysis (PCA) (Hooper, 2003; Lautz and Fanelli, 2008; Mencio and Mas-Pla, 2008; Woocay and Walton, 2008; Gordon *et al.*, 2013). A PCA for Quilcayhuanca is not feasible, due to the number of samples being close to the number of variables.

4. Results

4.1 Stream discharge and water balances

The magnitude of stream discharges as well as the relative gains and losses over study reaches varied widely within sites and between the two sites (Tables 1-2, Figure 4). At Llanganuco, the highest discharges were found in the lower half of the meadow (Reaches A-C), where discharge measured by dilution gauging ranged from 108 to 115 L s^{-1}. Reaches A-C were also consistently gaining (positive ΔQ) from lateral subsurface inflow, with a total net gain of 7 L s^{-1}, or 6% of average discharge, over 485 m of stream length (14 L s^{-1} km^{-1} or 13% of average discharge, normalized by reach length). The reaches in the upper half of the meadow (Reaches D-G), in contrast to those in the lower, had much smaller discharges (4.1 to 9.3 L s^{-1}) and showed net losses of stream water to the subsurface (negative ΔQ). Over the consecutive Reaches D-F, the net loss was 2.5 L s^{-1} or 31% of average discharge over 678 m (3.8 L s^{-1} km^{-1} or 53% of average discharge, normalized). The longitudinal patterns can be seen clearly in Figure 5a, which shows ΔQ as negative in the upper meadow and positive in the lower meadow. The other two meadow reaches (X and Y) were on a tributary to the main stem that flowed off of the Laguna 69 debris fan, and joined the main stem between reaches C and D (Figure 4). Discharges in these reaches ranged from 16 to 21 L s^{-1}. Reach X was losing discharge at a normalized rate per km (71% of Q_{av}) similar to the nearby Reach D (80% of Q_{av}), but Reach Y was gaining a large amount of water per km (180% of Q_{av}), mostly from a collection of diffuse springs and rivulets that were observed flowing into the channel from a very wet area at the nearby debris fan-meadow interface.

The balance of gross gains and losses in the studied reaches at Llanganuco also showed contrasts between the upper and lower sections of the meadow. Four of the five reaches that

experienced net loss in the upper meadow also showed gross gains concurrent with gross losses, meaning that the reaches were exchanging water with the underlying shallow meadow aquifer in both directions (Reaches E-G and X, Table 1, Figure 5a-b). On the other hand, no gross losses of stream water were observed in reaches in the lower part of the meadow that experienced net gains (Reaches A-C) or Reach Y.

Contributions to the meadow streams from the adjacent geomorphic units was also highly variable at Llanganuco. We found that the majority of the stream water that exits the study area (115 L s^{-1} at the outlet) was issuing from a single spring situated at the base of the debris fan near the confluence of Reaches X and D, which we named the Big Spring (Figure 4). By measuring all other flows upstream and downstream of the Big Spring, its discharge was estimated at 82 L s^{-1}. Other measurable flows from the debris fan into meadow channels totaled 23 L s^{-1}, including spring and surface flow at the head of Reach Y and spring flow at the head of Reach G, but excluding the Big Spring. At the very top of the debris fan, channel outflow from Laguna 69 was measured as 22 L s^{-1}, all of which disappears into the loose, blocky subsurface material near the top of the fan. Flow from the Broggi moraine onto the meadow was through two spring-fed channels totaling 5.2 L s^{-1}, which join the main stem above Reach F. Channel outflow from Laguna Broggi onto the moraine was measured as 36 L s^{-1}, all of which is lost to the subsurface on the upper moraine.

Discharge measured in the stream at Quilcayhuanca ranged from 760 L s^{-1} in the middle of the moraine to 1500 L s^{-1} at the lowest point at Casa del Agua (Table 2). Delta-Q values varied depending on the reach position relative to meadows and the moraine (Table 2, Figure 5c), but are generally higher than the ΔQ values at Llanganuco as a fraction of Q, meaning there was more variability in streamflow. The most upstream Reach 6 was positioned at the bottom of

the upper meadow, before the stream enters the moraine, and was gaining at a normalized rate per km equal to 81% of Q_{av}. Reach 5, the first reach in the moraine, was losing 270% of Q_{av} per km, Reach 4, in the middle of the moraine, was gaining 34% of Q_{av} per km, Reach 3 at the bottom of the moraine, was gaining 86% of Q_{av} per km, Reach 2, the first meadow reach just downstream of the moraine margin, was gaining 270% of Q_{av} per km, and Reach 1, the most downstream reach, was losing 84% of Q_{av} per km. Normalized gross gains and losses were also larger than at Llanganuco, with the highest gross gains and losses in the meadow Reaches 1 and 6 and at the top of the moraine (Reach 5) and the lowest in the middle of the moraine Reach 4 (Figure 5). All reaches had concurrent gains and losses, but Reach 2 was the most lopsided, with a $Q_{loss,min}$ of -27 L s^{-1} and a $Q_{gain,min}$ of 580 L s^{-1}.

4.3 Geochemical and isotopic composition of waters

The chemical composition of surface and groundwater is dominated by calcium and sulfate and/or bicarbonate at both Llanganuco and Quilcayhuanca, and the general patterns look similar plotted on Piper diagrams (Figure 6). Upstream channel samples plot high in the calcium-sulfate corner, with more downstream samples tending towards groundwater in the calcium-bicarbonate corner. A more subtle shift towards greater sodium in downstream and spring samples is also apparent. For the 2012 Llanganuco surface waters, where enough samples were taken to make correlations (n=41), sulfate is strongly positively correlated with magnesium (Pearson $r=0.97$, $p<0.001$) and more weakly with calcium ($r=0.63$, $p<0.001$). Fluoride is positively correlated with both sodium ($r=0.53$, $p<0.001$) and chloride ($r=0.61$, $p<0.001$), and negatively correlated with sulfate ($r=-0.73$, $p<0.001$), but the correlation between sodium and sulfate is very weak ($r=-0.38$, $p=0.013$). Sodium and chloride are only weakly correlated ($r=0.46$, $p=0.002$). At Quilcayhuanca, which is dominated in the upper catchment by the sulfide-rich

Chicama formation, sulfate values were as high as 105.1 mg L^{-1} in the stream and pH was low (~3.8). At Llanganuco, the Broggi lake outlet is the most concentrated in sulfate (28.5 mg L^{-1}), but pH throughout the watershed is generally circumneautral to weakly alkaline (6.8-7.9 in stream water and up to 8.5 in springs). Laguna 69, on the other hand, is not as heavily sulfate dominated, with proportionally more bicarbonate than Laguna Broggi. The valley above Laguna Broggi is composed of Chicama formation, but none is exposed above Laguna 69, where the bedrock is granodioritic.

At Llanganuco, the concentration of sulfate drops and sodium rises in general as water moves downstream from both lakes, and some springs and wells are enriched in sodium. Plotting sodium versus sulfate is therefore useful to elucidate dominant relationships between groups of samples (Figure 7a). Samples fall into several groupings in Figure 7a. Laguna Broggi starts very high in sulfate, but the concentration drops quickly, moving in a horizontal line as water flows over and through the moraine. Laguna 69 is moderately high in sulfate and low in sodium, but the sodium concentration jumps up in the channel samples on the fan, and jumps up again in the springs and channels at the fan base. The upper meadow channel samples (Reaches D-F) sit in a cluster in the middle of the plot, and the Big Spring and lower main channel samples form another cluster between the upper channel and the base of the fan. Three springs sit apart; two artesian springs (S07 and S08) that were found in raised mounds in the middle of the wet meadow, and a single spring issuing from the base of the talus slope to the west of the meadow. Due to its position in the watershed and its very different chemistry, the talus spring likely has no glacial influence and is completely precipitation derived. At Quilcayhuanca, a similar plot shows that the four stream samples decrease in sulfate and increase in sodium towards the spring and well samples as they move downstream (Figure 7b).

In order to simplify the multidimensional geochemistry and to analyze mixing and other relationships between ions and samples, a principle component analysis (PCA) was performed with the samples at Llanganuco excluding groundwater wells. The first two principle components (PC1 and PC2) explain 78% of the total variance of the data set, and the third component explains only an additional 8%, so PC1 and PC2 were retained and plotted (Figure 7c). The first component, PC1, is positively correlated with fluoride ($r=0.85$, $p<0.001$), chloride ($r=0.84$, $p<0.001$), nitrate ($r=0.82$, $p<0.001$), and sodium ($r=0.53$, $p<0.001$), and negatively correlated with potassium ($r=-0.60$, $p<0.001$), calcium ($r=-0.64$, $p<0.001$), and especially sulfate ($r=-0.90$, $p<0.001$) and magnesium ($r=-0.95$, $p<0.001$). The second component, PC2, is positively correlated with fluoride ($r=0.279$, $p=0.077$), sodium ($r=0.66$, $p<0.001$), potassium ($r=0.70$, $p<0.001$), and calcium ($r=0.53$, $p<0.001$). In general, samples appear in similar groupings as in Figure 7a.

The stable isotopic composition of water is reported as $\delta^{18}O$ and $\delta^{2}H$, relative to Vienna Standard Mean Ocean Water (VSMOW) (Figure 8). The distribution of stable isotopic composition was relatively narrow, with $\delta^{18}O$ ranging from -17.3 to -14.8‰ and $\delta^{2}H$ from -128 to -108‰. This is a small range relative to the expected seasonal variation; for example, monthly means over two years at Marcapomacocha, Peru, at a similar elevation (4477 m) to the Llanganuco meadow, range from -5.21‰ $\delta^{18}O$ and -24.4‰ $\delta^{2}H$ in August to -20.33‰ $\delta^{18}O$ and -149.7‰ $\delta^{2}H$ in March, with a precipitation-weighted mean of -16.25‰ $\delta^{18}O$ and -115.5‰ $\delta^{2}H$ (IAEA, 2013). All samples fall near or above the global meteoric water line (GMWL), with similar slope (linear regression slopes of 7.9 at Llanganuco and 7.6 at Quilcayhuanca) and deuterium excesses (range of 10.0-13.0), except for Laguna Broggi and its outlet down the moraine, which fall below the GMWL with deuterium excess values down to 7.3, indicating

evaporation effects in the lake or the series of lakes above it. At Llanganuco, groundwater from wells is the most enriched in heavy isotopes (mean of -15‰ $\delta^{18}O$, excluding GW3), while samples near the base of the moraine and in the upper meadow are the most depleted. The majority of samples from the main channel and Big Spring fall in the middle of the plot near -16.3‰ $\delta^{18}O$. Similarly at Quilcayhuanca, all the stream samples also cluster near -16.3‰ $\delta^{18}O$, with most groundwater wells more enriched.

5. Discussion

5.1 Groundwater-surface water interactions

Gains and losses of water were unequally distributed across the landscape during our study. The streams at both study sites were gaining and losing water at different rates in locations controlled by geomorphology, especially by the strongly varying slopes of geomorphic units (fan, moraines, and meadows), as we would expect from standard groundwater flow fields around breaks in topographic slope (Winter *et al.*, 1998). Figure 9 presents conceptual models illustrating subsurface flow paths and resulting GWSWI based on our results and interpretation.

At Llanganuco, the outlets of both glacial lakes were losing water to the subsurface at the top of the debris fan and moraine (Figure 4), where slopes become steep (average slopes are 12º and 7º for fan and moraine, respectively). Both channels run completely dry before they descend much of their respective units, which means that glacial lake water is recharging aquifers beneath these units. Similarly at Quilcayhuanca, the stream loses net water to the subsurface moraine over the upper reach of that unit (Reach 5, Figure 5c), which has an average slope of 5º. Near the bottom of these steep units, water is discharging from the subsurface and feeding streams, which was subjectively observed at Llanganuco in the springs and incipient channels that appear around the bases of the fan and moraine, and measured as a positive ΔQ at Quilcayhuanca in Reach 3,

the last reach in the moraine. This influence of water suggests that there exist flow cells of groundwater moving through these steep units and connecting recharge at the top with discharge at the bottom (Figure 9).

In the meadows, stream channels closest to the margins of the fan and moraine at Llanganuco (Reach Y and the springs and channels upstream of Reach F) and at the base of the moraine at Quilcayhuanca (Reach 2 and the spring in Figure 2) are strongly net gaining water; however, once a short distance away from the margins, reaches that flow through the upper portions of meadows (Reaches D-G and X at Llanganuco and Reach 1at Quilcayhuanca) are net losing to the underlying meadow. In the lower portions of meadows, the streams are net gaining again (Reaches A-C and Reach 6, at the downstream end of the upper meadow at Quilcayhuanca), suggesting that there are groundwater flow cells through the meadows, as well (Figure 9).

At Llanganuco especially, it is also interesting to compare the absolute fluxes of water as they vary over the catchment. Because the stream exits the study area over a bedrock outcrop and waterfall, we can be certain that we measured the total outflow from the watershed (neglecting potential fracture flow through bedrock) at 115 L s^{-1}. We found that the majority of this outflow (71%) was not flowing from the upper stream reaches through surface channels, but was issuing from a single source in the Big Spring, at the base of the Laguna 69 fan and relatively far down in the catchment, which makes the source of Big Spring water a topic of interest. Other fractional components of the total outlet discharge are the upper meadow channel (6%), the Laguna 69 tributary (17%), and lateral groundwater inflow to the lower meadow channel (6%). As a result of the dramatic inflow at the Big Spring, discharge in the upper meadow channels was a small fraction of discharge in the lower reaches. Also, if its origin is

assumed to be completely from debris fan groundwater, than the large spring discharge makes the debris fan a net contributor to stream flow during the dry season; the fan aquifer receives 22 L s^{-1} of surficial water from the glacial lake, but contributes 105 L s^{-1} to the meadow channels, mostly through springs at its base. The excess water produced could be from wet season precipitation stored in the fan aquifer, or from glacial lake water that seeps through the lakebed and into the fan. In contrast to the fan, the moraine was a net sink of surface water during our study, receiving 36 L s^{-1} of water from the Lagua Broggi outlet, but contributing only 5 L s^{-1} to the meadow streams, and demonstrating that different aquifer units of the proglacial landscape have varying potential as stores of groundwater and as sources of streamflow during the dry season. The wet meadow along Reaches A-C contributed a small absolute amount of discharge to the total stream outflow (7 L s^{-1}). These reaches passed through the wettest part of the meadow, adjacent to the talus slope on the west side of the study area, where several springs were located at the talus margin and in spongy vegetated mounds that rose above the rest of the meadow (the artesian springs), so we assume that most of this inflow came from groundwater on the western side. Although the absolute contribution to discharge measured over this approximately ½ km of channel was small, it must be remembered that this meadow is a small example high in the watershed, and many glacial valleys of the Cordillera Blanca contain tens of kilometers of wet meadows along their streams.

The study site at Quilcayhuanca was relatively lower in its valley than the headwaters studied at Llanganuco, with a considerably larger contributing area. Clearly absolute discharges were much larger at Quilcayhuanca as a result, but interestingly, net and gross gains and losses were also larger as a normalized fraction of total discharge (opposite to the general pattern found by Covino et al (2011) of decreasing relative gross loss with increasing discharge), even in

meadow Reaches 1,2, and 6. Gross gains and losses occurred concurrently in every measured reach, and were often up to 4 times average discharge per km. The net gain over 1.2 km of studied stream length was 200 L s^{-1}, which is a large proportion of the total discharge of approximately 1100 L s^{-1}; however, it is possible that this large net discharge is partially a product of error in the dilution gauging method (Schmedel *et al.*, 2010).

That there was more interaction with the subsurface at Quilcayhaunca than at Llanganuco is likely due to differences in subsurface hydraulic conductivity or heterogeneity, but these differences may be representative of longitudinal gradients in Cordillera Blanca valleys or differences between glacial deposits closer or farther from modern glacial termini. We therefore expect that groundwater will have a greater impact on stream chemistry through fractional hydrologic turnover (Covino *et al.*, 2011) as streams flow downwards through successive meadows and moraines in their valleys. We should also note, however, that even in the headwaters of the Llanganuco valley, almost all of the water exiting the catchment spent some time in the subsurface, and half of the total outflow originated as groundwater downstream of the glacial lakes. Clearly, groundwater in both proglacial geomorphic units and wet meadows has a considerable influence on streams, but the original sources of groundwater (precipitation or stream water lost to the subsurface) is not always obvious.

5.2 Water sources, mixing and transformation

High sulfate concentrations in water proximal to glaciers has been used as a tracer of glacial water in mixing models (Mark *et al.*, 2005; Baraer *et al.*, 2009) and been explained by the high concentration of sulfide minerals in bedrock and the high rates of weathering beneath glaciers and in freshly exposed till. In this study, however, it is difficult to use sulfate as an indicator of glacial contribution because of the differing influence of the Chicama formation in

the study basins, which leads, for example, to Laguna Broggi having higher sulfate concentrations than Laguna 69, despite being farther from glacier termini. We therefore used plots of solutes or transformed solutes to interpret relationships between different sampled waters in the complex terrain and channel configuration of the Llanganuco site. Although similar to the sodium versus sulfate plot (Figure 7a), the PCA analysis plot (Figure 7c) takes the total variation of all solutes into account, and performs better at partitioning causal relationships. (For example, water flowing down and through the moraine does not change PC1 values in Figure 7c, while it does change sulfate concentrations in Figure 7a.) PC1, which explains the most variation in the data, appears to separate the samples influenced by Laguna Broggi, with its high Chicama signature, from the other samples. We therefore interpret PC1 as an indicator of bedrock provenance, with negative values indicating more Chicama influence, and positive values indicating Blanca batholith influence. This allows PC2, orthogonal to PC1, to be relatively free of this influence. With its correlations to sodium, potassium, calcium, and fluoride, we interpret PC2 as an indicator of silicate weathering processes (Appelo, 1993), which would have a greater influence on groundwater in proportion to its contact time with silicate sediments.

Mixing relationships throughout the Llanganuco catchment are difficult to determine using Figures 7a or 7c, because there is a wide scatter of points instead of clear mixing lines, and because the end members of different mixing relationships are not clear. For example, the two glacial lakes have different compositions, the shallow meadow groundwater is not chemically consistent, and the Big Spring and catchment outflow are in the middle of the plot, where they could represent varying combinations of several potential end members. However, when focusing on the downstream samples, the geochemistry supports the composition of the outlet discharge estimated above using tracer dilution data (71% Big Spring, 17% fan, 6% upper

meadow, and 6% meadow groundwater); the outflow is close to the Big Spring, but shifted slightly towards samples at the base of the fan. The source of the Big Spring water itself is difficult to determine conclusively based on the geochemistry. It sits slightly apart from the other springs and channels at the base of the debris fan, shifted to the right in both plots towards the upper meadow samples, and may represent mixing of these two water types in the subsurface. Other clusters of points appear to be controlled not so much by mixing as by transformations in geochemistry concurrent with changes in geomorphology and GWSWI discussed above. Both lakes are set apart from the channels just downstream from them, which is especially evident in the increase in PC2 from Laguna 69 to the channels on the debris fan. PC2 jumps again between these channel samples and samples at the base of the fan. Water flowing down the moraine also increases PC2 quickly in a short distance. Substantial shifts in chemistry appear to be punctuated by changes in slope. Groundwater flowing through these units and discharging to the surface at their bases is significantly higher in silicate weathering products than channel and lake water recharging at the tops of the units. Other samples with high silicate weathering products include the well samples (not shown in Figure 7c), the artesian meadow springs, and the talus spring (especially high in sodium), suggesting that these waters have had extended contact with silicate minerals in the subsurface (Dethier, 1986). Because of the non-conservative geochemical behavior in waters sourced from the two lakes and in samples with heavy groundwater influence, it is not possible to estimate percentages of glacial lake water in the outlet stream or springs.

Stable water isotopes, however, do not react or change composition in the subsurface (Buttle, 1994), and the lack of local evaporation effects in the dataset (with the exception of Laguna Broggi) makes them conservative in this context. The main controls on isotopic composition in the Cordillera Blanca are season, the elevation of source precipitation, and

potentially the influence of glacial meltwater (Mark and Seltzer, 2003; Mark and McKenzie, 2007). All of the samples at both sites are in a narrow range close to the precipitation weighted mean for their approximate elevation, but variations within sites may be due to minor differences in the elevation of sources. At Llanganuco, the groundwater in most of the meadow wells was the most enriched (approximately -15‰ $\delta^{18}O$), suggesting that these samples represent precipitation that fell directly on the meadows, while samples from streams and springs have sources at higher elevation. GW3 is the outlier, and appears to be influenced by water from the upper meadow channel. Samples sourced from near Laguna 69 fall tightly together in the center of the plot, including the lake itself and the talus spring, which were set apart in the geochemistry plots by their high silicate weathering products, suggesting that all of these samples are sourced from a similar elevation, with little mixing with lower elevation groundwater during the descent from lake to meadow. It is also interesting that Laguna 69 is so similar in isotopic composition to the waters at the base of the fan, because it suggests that the observed changes in geochemistry are non-conservative transformations in the subsurface and not due to conservative mixing. The fact that the talus spring, with no likely glacial influence, is so similar to the glacially derived Laguna 69 samples in isotopic composition, suggests that stable isotopes are not useful for determining percent glacial melt at Llanganuco and instead represent the elevation of source precipitation. Assuming an average elevation effect of 3.25‰ $\delta^{18}O$ per km (Clark and Fritz, 1997), the apparent elevation difference between the Big Spring and groundwater wells is 390 m, putting the mean source at 4765 masl, approximately the same elevation as the glacier terminus. The artesian meadow springs have a composition intermediate between the Laguna 69 samples and the wells, likely a result of mixing of these sources in the deeper meadow aquifer. The most depleted samples, from the base of the Broggi moraine, have an apparent elevation difference of

710 m above the meadow wells, or 5085 masl, while Laguna Broggi itself appears to be enriched by evaporation from a similar source. Samples from the upper meadow channel are less depleted than the base of the moraine and are likely a mixture of that source and groundwater from the meadow. In the lower meadow, however, the channel does not appear to be influenced to an obvious degree by meadow precipitation, and channel flows are primarily sourced from higher-elevation water, which supports the conclusion from the tracer dilution results that meadow groundwater does not have a large influence on this stream.

At Quilcayhuanca, there was more evolution in water geochemistry observed along the channel than in the continuously sampled reaches at Llanganuco, likely due to the greater fractional hydrologic turnover between the channel and subsurface. Because of repeated gross gains and losses of stream water from and to the underlying aquifer, stream chemistry can be strongly influenced by groundwater chemistry and subsurface geochemical conditions over distance (Covino *et al.*, 2011). From upstream to downstream, the samples move linearly toward the groundwater and spring samples, which have less sulfite and more sodium (Figure 7b). It appears that at Quilcayhuanca, the groundwater has a more uniform composition among the shallow wells and the spring issuing from the base of the moraine, and it alters the channel chemistry in a predictable fashion. The groundwater samples are more isotopically enriched (with the exception of GW3) than the channel samples, which cluster tightly together at -16.3‰ $\delta^{18}O$ likely due to their higher elevation source. Because of the large amount of groundwater-surface water exchange observed at Quilcayhuanca, the isotopic and geochemical influence of the stream source waters has likely been diluted by groundwater through fractional hydrologic turnover, and the sources of streamflow are probably at a higher average elevation than the isotopic composition suggests.

5.3 Groundwater and surface water as resources

Our results show that in the dry season, groundwater is an important influence on and component of streamflow. In the highest headwaters of glacierized catchments, groundwater makes up a large percentage of incipient stream flow, although it is not clear in this study what percentage of this groundwater discharge is from stored precipitation versus recharging glacial melt. In these relatively small meadow systems, the contribution to streamflow by meadow precipitation is certainly low. However, in the mid-valley meadow systems, groundwater is more certainly derived from local precipitation sources, primarily recharged during the wet season. In these more robust valley aquifers, GWSWI are substantial and influence stream discharge rates and geochemistry. At the high groundwater inflow rates measured over the study reaches at Quilcayhuanca (200 L s^{-1} total net inflow over 1.2 km), groundwater contributions would be substantial over tens of km, representing an important dry season water resource.

This study has implications for calculating percent glacier melt in tributaries of the Rio Santa by conservative geochemical mixing models. Both the large-scale GWSWI and non-conservative transformations seen at Llanganuco and the fractional hydrologic turnover measured at Quilcayhuanca would tend to increase the influence of subsurface chemistry on stream water dissolved constituents, beyond the expected influence exerted by the volumetric contribution of groundwater to stream flow. Although we stress that these methods are still valid on a broad scale and are useful and necessary in large, poorly instrumented regions such as the Cordillera Blanca and Callejón de Huaylas, the dynamics observed in this study would cause some basic conservative mixing models to underestimate the contribution of glacial melt to stream discharge. This underestimation means that, in some locations, glacial resources might be even more important to water supply than was previously thought.

6. Conclusion

We found that groundwater is an important element in the proglacial hydrologic landscape, although the magnitude and type of impact varies from headwaters to mid-valley and between geomorphic units (moraines, fans, meadows, and talus). In general, gains and losses of water to and from stream channels are unequally distributed in the dry season, and are primarily controlled by geomorphic unit, slope, and valley position. Losses of stream water to the subsurface were observed at the upper ends of meadows and steeper geomorphic units, while gaining stream reaches were observed at the lower ends of meadows and where steep geomorphic units terminate into gently sloped meadows, leading to predictable and generalizable groundwater flow cells (Figure 9). Glacial lake water in stream channels was observed recharging aquifers in fan and moraine units, which represented a net loss of surface water to the moraine at Llanganuco, but which was more than made up by discharging groundwater from the Llanganuco debris fan and the moraine at Quilcayhuanca.

Changes in slope and geomorphic unit also punctuate changes in geochemistry, which shows evidence of influence by both conservative mixing and chemical weathering at the headwater site. Stable isotopes mainly indicate the elevation of source precipitation, which shows that meadow groundwater is a small component of the stream flow compared to higher lakes, glaciers, and precipitation in the headwater Llanganuco catchment. At the mid-valley site, stream water geochemistry trends towards the groundwater geochemical composition as water flows downstream, due to direct net inflow and hydrologic turnover.

The two study sites can be generalized to represent two points in a downstream continuum from headwaters to mid-valley in a typical Cordillera Blanca valley. The headwaters are dominated by water from glacial lakes, although much of it passes through the groundwater

system in steep proglacial geomorphic units, reemerging through springs at the meadow margins, rather than being transported directly in stream channels. Along these subsurface flow paths, the water acquires a geochemical signature from silicate weathering. The mid-valley systems, with their extensive stepped meadows and moraines, are dominated by net inflows from precipitation-derived groundwater and hydrologic turnover (gross gains and losses) with the underlying groundwater aquifer. In such a continuum, groundwater contributions to net stream discharge increase with distance downstream, while the influence of groundwater on stream chemistry changes more subtly, from an in-situ weathering influence high in the catchments to a fractional mixing influence as streams interact with groundwater in lower meadow systems. By the time they reach the Rio Santa at the edge of the Cordillera Blanca, streams are certainly heavily influenced by groundwater in both their total discharge and water chemistry.

There are several unknown aspects of groundwater in proglacial valleys that deserve further attention. For example, the total storage volumes and residence times of the groundwater aquifers in meadows and other units remains unknown; this information would help us understand how different aquifers might respond to reduced-meltwater scenarios, and how long the coupled groundwater-surface water system would take to adjust to different flow or precipitation regimes. A better understanding of the details of subsurface flow paths would also constrain the influence of heterogeneity in moving water through the units studied. Further work in this area should concentrate on subsurface exploration (drilling and non-invasive geophysical methods) to map the depth to bedrock and estimate porosity and heterogeneity, leading to estimates of total storage capacity, as well as modeling of residence times or dating of water.

Acknowledgements

The authors would like to thank AnneMarie Glose for assistance in the field, and Patrick Burns, Alex Eddy, and Michel Baraer for assistance with spatial and geochemical data. This work was supported by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-0750965, a CUAHSI Pathfinder Fellowship, a GSA Graduate Student Research Grant, and the National Science Foundation under Grant No. EAR-0901480.

Tables

Table 1. Llanganuco dilution gauging and water balance results.

Reach	Description	Distance up stream (m)		Length (m)	Q (L s^{-1})			ΔQ (L s^{-1})	Q_{loss} (L s^{-1})			Q_{gain} (L s^{-1})		
		upstream	downstream		upstream	downstream	ave		min	max	ave	min	max	ave
A	lower meadow, outlet	208	7	200	110	115	113	5.0	0.0	0.0	0.0	5.0	5.0	5.0
B	lower meadow	419	208	212	109	110	110	0.7	0.0	0.0	0.0	0.7	0.7	0.7
C	lower meadow	530	457	73	108	111	110	3.0	0.0	0.0	0.0	3.0	3.0	3.0
D	upper meadow	799	591	207	8.0	6.8	7.4	-1.2	-1.0	-0.9	-0.9	-0.3	-0.3	-0.3
E	upper meadow	971	799	172	8.9	8.0	8.5	-0.9	-1.6	-1.7	-1.6	0.7	0.9	0.8
F	upper meadow	1107	971	136	9.3	8.9	9.1	-0.4	-1.1	-1.2	-1.2	0.7	0.8	0.7
Moraine[a]	flow off Broggi moraine					1.5								
G	upper meadow	1371	1208	163	4.2	4.1	4.2	-0.1	-0.2	-0.2	-0.2	0.0	0.0	0.0
Broggi[a]	Laguna Broggi outflow					36								
X	debris fan trib	737	616	120	21	19	20	-1.7	-2.9	-3.1	-3.0	1.2	1.4	1.3
Y	debris fan trib	872	737	135	16	21	19	4.4	0.0	0.0	0.0	4.4	4.4	4.4
69[a]	Laguna 69 outflow					22								

[a]single-slug gauging sites

Table 2. Quilcayhuanca dilution gauging and water balance results.

Reach	Description	Distance up stream (m)		Length (m)	Q (L s^{-1})			ΔQ (L s^{-1})	Q_{loss} (L s^{-1})			Q_{gain} (L s^{-1})		
		upstream	downstream		upstream	downstream	ave		min	max	ave	min	max	ave
1	meadow, Casa del Agua	300	29	270	1500	1200	1300	-300	-1400	-18000	-9600	1100	17000	9300
2	top of meadow	472	300	172	930	1500	1200	550	-27	-45	-36	580	600	590
3	bottom of moraine	624	472	152	820	930	880	110	-320	-600	-460	430	710	570
4	moraine	840	624	216	760	820	790	58	-110	-140	-120	170	200	180
5	top of moraine	1003	881	122	1200	900	1100	-350	-820	-1700	-1300	460	1300	900
6	bottom of meadow	1246	1003	243	1000	1200	1100	220	-920	-11000	-5700	1100	11000	6000
7[a]	meadow	1559	1246	313		1000				-2400				

[a]upstream discharge not available for Reach 7

Figures

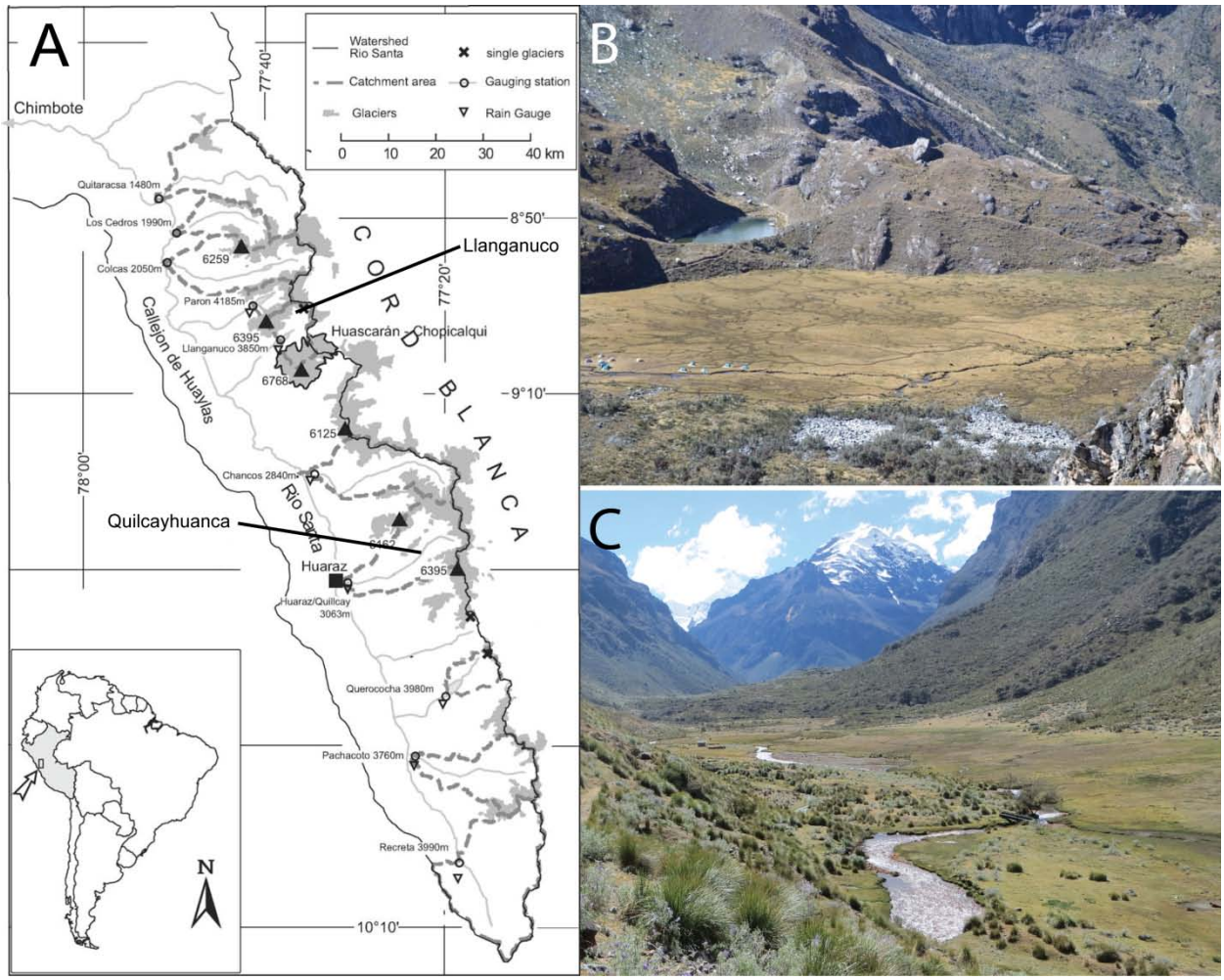

Figure 1. (a) The Cordillera Blanca and upper Rio Santa valley, with arrows indicating study locations in the Llanganuco and Quilcayhuanca watersheds. (b) Photograph of the studied meadow at Llanganuco, looking from west to east down from the top of the Laguna 69 fan. Water flow in the stream is from left to right. (c) Photograph of the Casa del Agua meadow at Quilcayhunaca, looking upstream at the valley-crossing moraine.

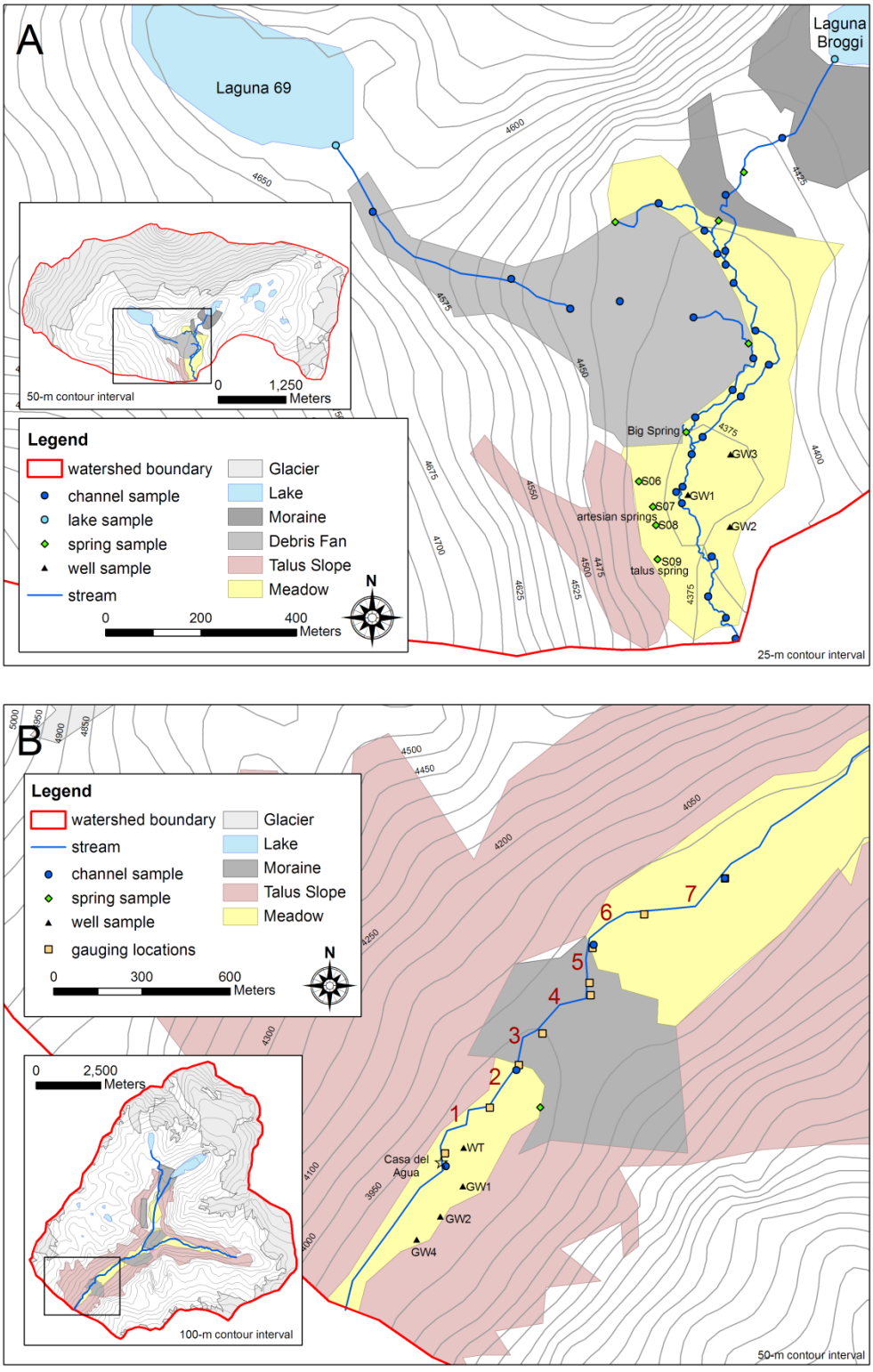

Figure 2. Maps of water samples collected in 2012 at study sites in the (a) Llanganuco and (b) Quilcayhuanca valleys, with geomorphic units shown. At Quilcahuanca, fluorometer locations for discharge gauging measurements and stream reach numbers are also indicated.

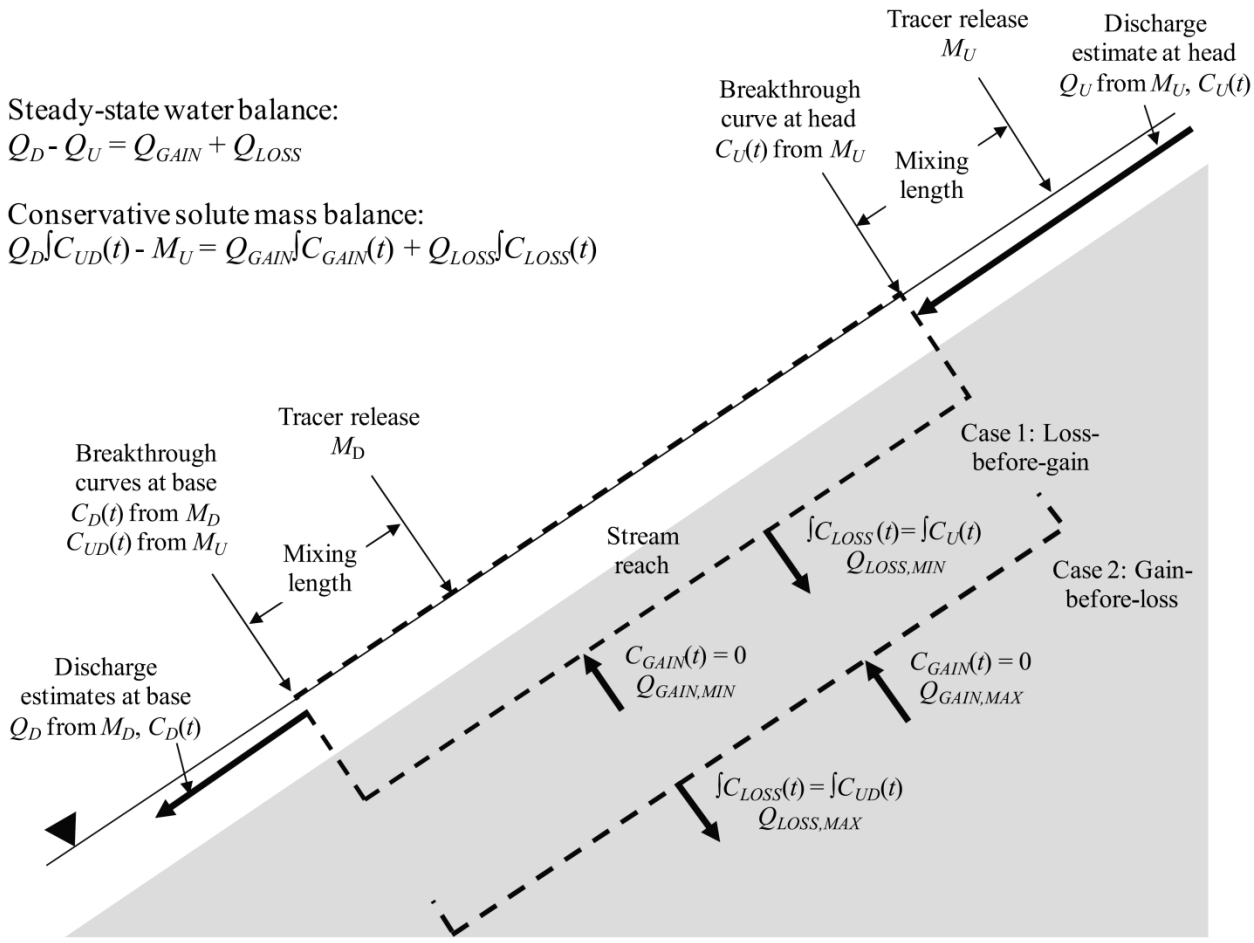

Figure 3. Schematic diagram of the "double slug" method over a hypothetical stream reach. Figure from Payne et al. 2009.

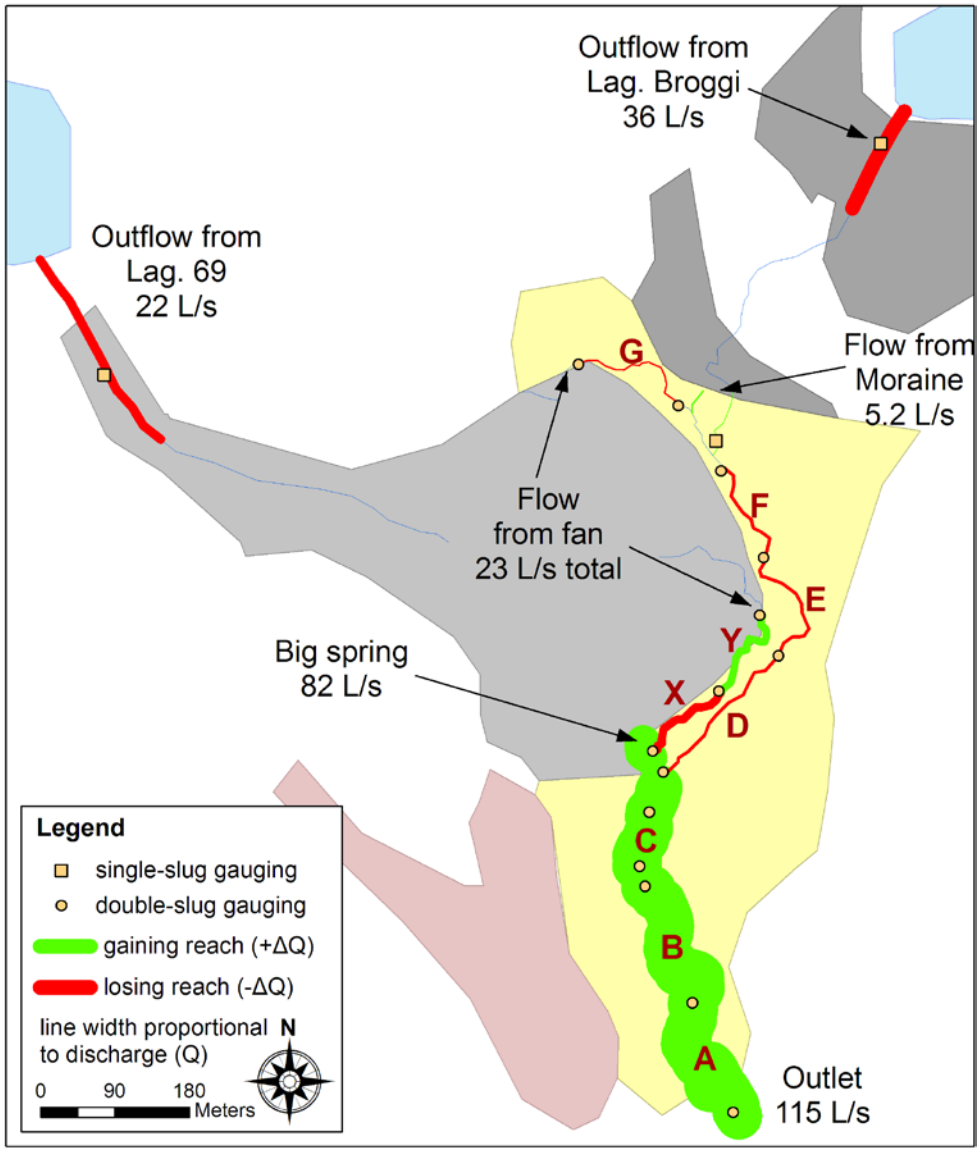

Figure 4. Simplified map of Llanganuco showing fluorometer locations for double- and single-slug discharge gauging measurements, as well as discharge and water balance data. Reach color indicates whether the reach was gaining or losing (positive or negative ΔQ) based on water balance calculations or clear subjective observations (for single-slug reaches only). Width of the colored lines are proportional to reach discharge (Q_{av}). Surface discharge (Q) at select locations is also indicated in L s^{-1}. Geomorphic units are as described in Figure 2a.

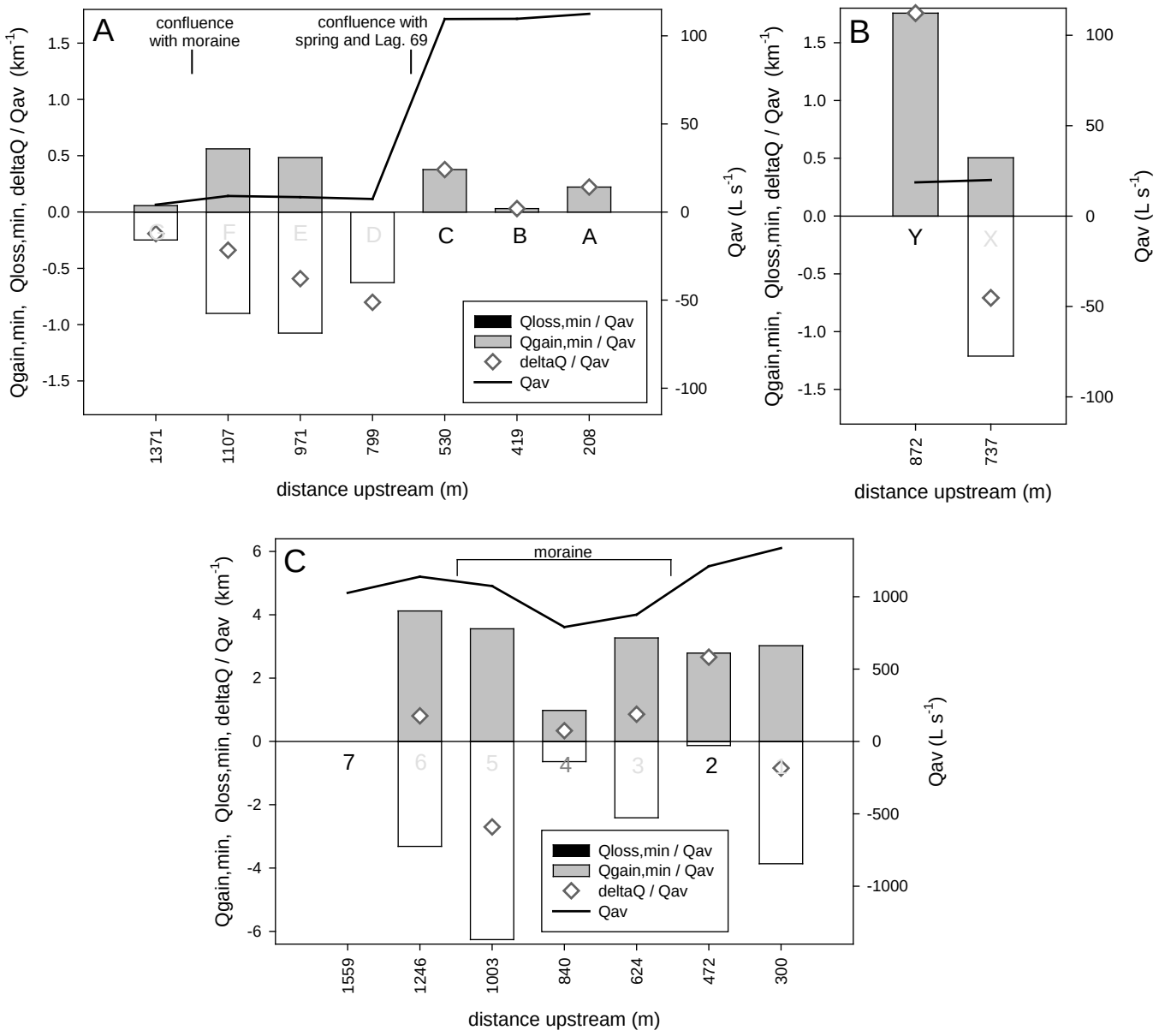

Figure 5. Stream reach discharges and water balances for (a) Llanganuco main stem, (b) Llanganuco Laguna 69 tributary, and (c) Quilcayhuanca. Distance upstream is the distance of the top of each reach (labeled with reach letter or number) from an arbitrary point near the most down-stream part of the study areas. Q_{av} (black line) is the average of Q at the top and bottom of each reach, in L s^{-1} (right axis). Bars show the minimum calculated gross gain and loss ($Q_{gain,min}$ and $Q_{loss,min}$) for each reach as a fraction of Q_{av} normalized by stream length (km^{-1}) (left axis). Diamond symbols represent the ΔQ over each reach as a fraction of Q_{av} normalized by stream length (km^{-1}) (left axis). For Reach 7 at Quilcayhuanca, only $Q_{loss,max}$ (not shown) and Q at the bottom of the reach (shown in place of Q_{av}) can be calculated.

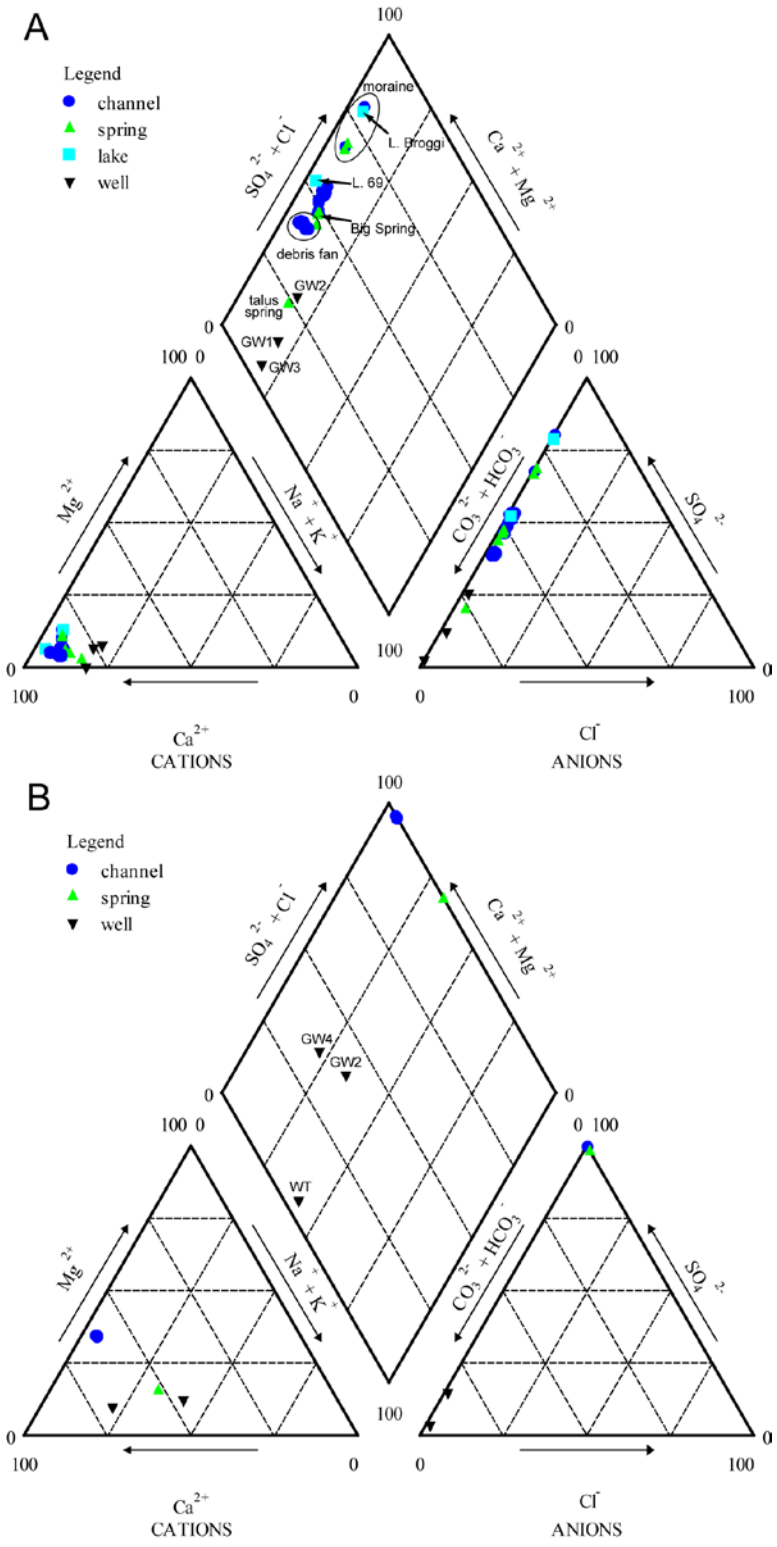

Figure 6. Piper plots of sample geochemistry from (a) Llangancuo and (b) Quilcayhuanca. Chemistry of GW3 at Quilcayhuanca is not shown due to likely contamination of the sample.

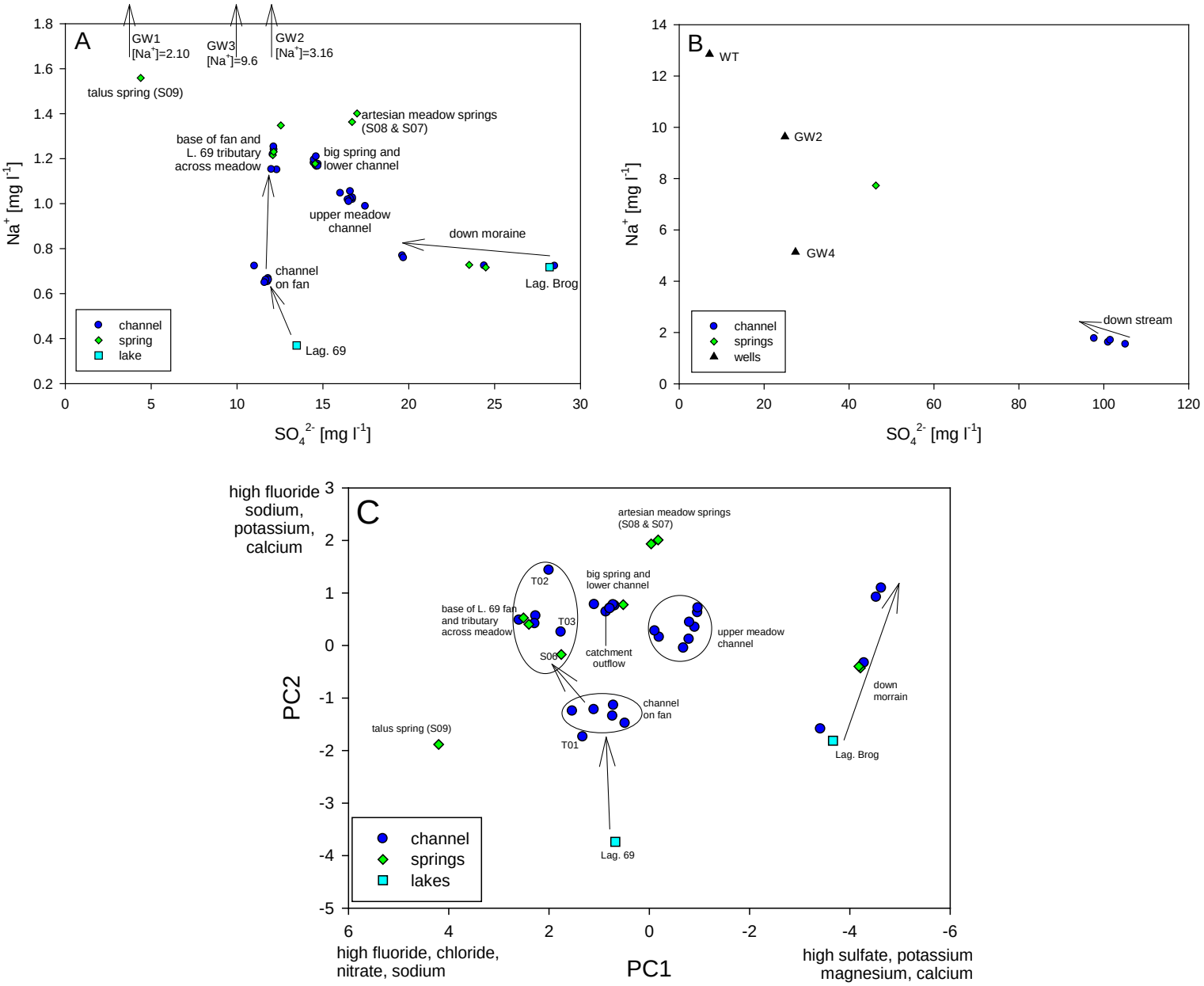

Figure 7. Plots of selected geochemistry, including sodium versus sulfate plots for (a) Llanganuco and (b) Quilcayhuanca, and (c) a plot of the principle components PC2 versus PC1 for Llanganuco.

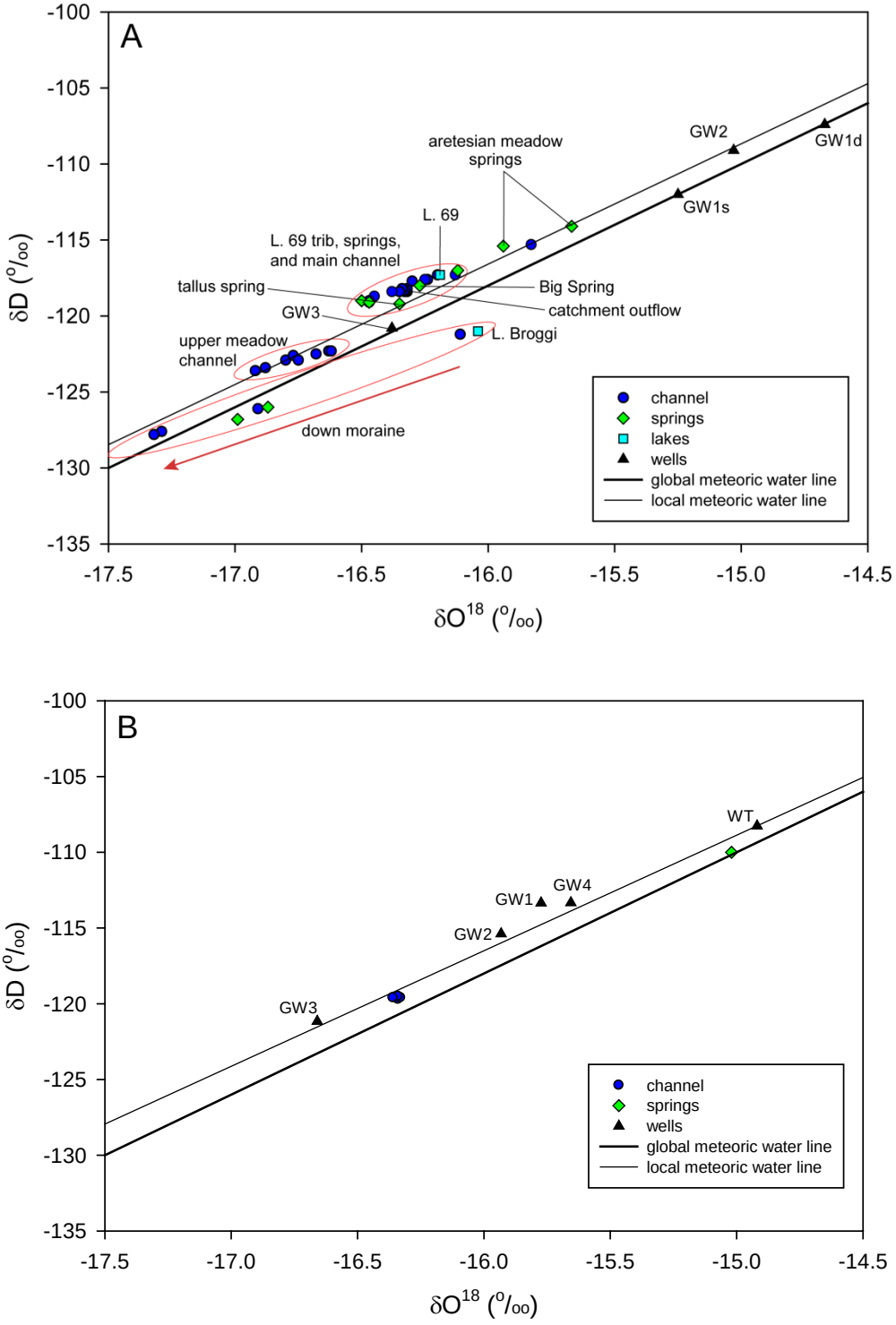

Figure 8. Plots of stable water isotopic composition for (a) Llanganuco and (b) Quilcayhuanca.

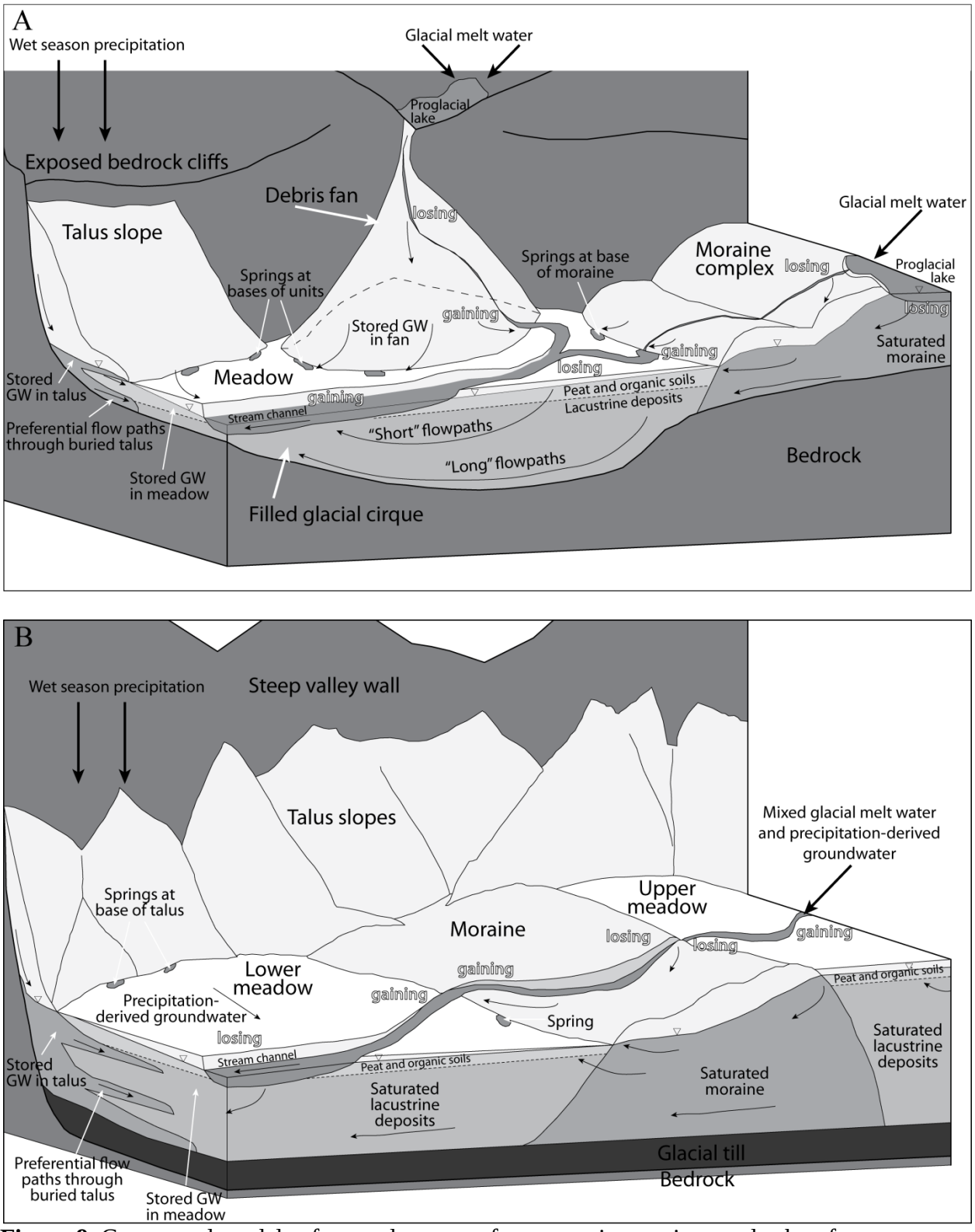

Figure 9. Conceptual models of groundwater-surface water interactions and subsurface flowpaths (black arrows) at (a) Llanganuco and (b) Quilcayhuanca. Locations where surface water is gaining from or losing to the subsurface are indicated in outlined text.

Summary and Significance of Research

The research projects described in this proposal investigate groundwater-surface water interactions (GWSWI) from several distinct human perspectives. The first chapter looks at hyporheic exchange in a mechanical and mathematical context and introduces a new method that provides improved spatial and temporal resolution to flux calculations, but it also supplies a tool to the hydrological community that makes heat transport modeling a more approachable, consistent, and hopefully more widely-used technique for calculating vertical water flux rates using temperature. The second chapter investigates GWSWI around constructed stream restoration structures in restored streams in Central New York. This study significantly contributes to the body of field data concerning hyporheic flow around structures, and also provides evidence to modify pre-existing conceptions of how stream restoration impacts hyporheic exchange and stream health in morphologically complex stream reaches. This research is important for understanding how the human activity of stream restoration affects hyporheic flow and biogeochemical cycling of nutrients and organic matter in stream ecosystems, which can impact nutrient pollution remediation and ecosystem health. It is important that ecologists, hydrologists, and engineers who design stream restoration projects understand how cross-vanes and other structures work to alter hyporheic flow patterns. The results of this study will directly benefit the environment by improving the design of future stream restoration projects in New York State and elsewhere, and by informing the assessment of existing projects. This work will increase the awareness of hyporheic exchange and its impact on nutrient dynamics and ecological health in restored stream ecosystems, and I believe that it already has, through two published papers (Gordon *et al.*, 2012; Gordon *et al.*, 2013) and four national and international conference presentations.

The third chapter investigates GWSWI in proglacial, alpine catchments of the Peruvian Andes. This research addressed unknown aspects of proglacial watershed hydrology, and the results expand the current state of knowledge about the role of groundwater and its interaction with surface water in tropical glacierized catchments. The context of this project is specific, but the results will be applicable to many regions around the world where mountains are a source of water for human use. The substantial influence of groundwater on water resources in alpine environments has been previously underestimated or ignored, and these results will be of interest to modelers and resource managers. Measurements of GWSWI can be used to model changes in surface water resources available to downstream users under changing climate scenarios. My results will allow more accurate predictions of future water availability in this critical region and other water-stressed areas globally.

The research projects described in this dissertation are united by important linkages between GWSWI and human activities. On one hand, I investigate how human restoration activities alter GWSWI, which in turn affect water quality and ecosystem services. On the other, I address how GWSWI may interact with an anthropogenically altered climate to affect water supplies for human users. All the projects look at the fluxes of water between small streams and underlying sediments, and use similar field and analytical methods to quantify these fluxes and their effects. It is my hope that these projects can improve the understanding of GWSWI as well as improve human lives.

References

Ames A, Dolores S, Valverde A, Evangelista P, Javier D, Gavnini W, Zuniga J, Gómez V. 1989. Clacier inventory of Peru, Part 1. Hidrandina.

Anderson MP. 2005. Heat as a ground water tracer. Ground Water, **43**: 951-968. DOI: 10.1111/j.1745-6584.2005.00052.x.

Anderson WP, Anderson JL, Thaxton CS, Babyak CM. 2010. Changes in stream temperatures in response to restoration of groundwater discharge and solar heating in a culverted, urban stream. J. Hydrol., **393**: 309-320. DOI: 10.1016/j.jhydrol.2010.08.030.

Appelo CAJPD. 1993. Geochemistry, groundwater and pollution. Balkema.

Baker MA, Dahm CN, Valett HM. 2000. Anoxia, anaerobic metabolism, and biogeochemistry of the stream-water-groundwater interface. In: Streams and Ground Waters, Jones JB, Mulholland PJ (eds.) Academic Press, pp: 425.

Baraer M, Mark BG, McKenzie JM, Condom T, Bury J, Huh KI, Portocarrero C, Gomez J, Rathay S. 2012. Glacier recession and water resources in Peru's Cordillera Blanca. J. Glaciol., **58**: 134-150. DOI: 10.3189/2012JoG11J186.

Baraer M, McKenzie JM, Mark BG, Bury J, Knox S. 2009. Characterizing contributions of glacier melt and groundwater during the dry season in a poorly gauged catchment of the Cordillera Blanca (Peru). Advances in Geosciences, **22**: 41-49.

Barnett TP, Adam JC, Lettenmaier DP. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. Nature, **438**: 303-309. DOI: 10.1038/nature04141.

Barry R, Seimon A. 2000. Research for mountain area development: Climatic fluctuations in the mountains of the Americas and their significance. Ambio, **29**: 364-370.

Barry RG. 2006. The status of research on glaciers and global glacier recession: a review. Prog Phys Geog, **30**: 285-306. DOI: 10.1191/0309133306pp478ra.

Barshad I. 1954. Cation exchange in micaceous minerals: II. replaceability of ammonium and potassium from vermiculite, biotite, and montmorillonite. Soil Sci., **78**: 57-76.

Barshad I, Kishk FM. 1968. Oxidation of ferrous iron in vermiculite and biotite alters fixation and replaceability of potassium. Science, **162**: 1401-1402.

Bencala KE, Gooseff MN, Kimball BA. 2011. Rethinking hyporheic flow and transient storage to advance understanding of stream-catchment connections. Water Resour. Res., **47**: W00H03. DOI: 10.1029/2010wr010066.

Bernhardt ES, Palmer MA, Allan JD, Alexander G, Barnas K, Brooks S, Carr J, Clayton S, Dahm C, Follstad-Shah J, Galat D, Gloss S, Goodwin P, Hart D, Hassett B, Jenkinson R, Katz S, Kondolf GM, Lake PS, Lave R, Meyer JL, O'Donnell TK, Pagano L, Powell B, Sudduth E. 2005. Ecology - Synthesizing US river restoration efforts. Science, **308**: 636-637. DOI: 10.1126/science.1109769.

Bernhardt ES, Sudduth EB, Palmer MA, Allan JD, Meyer JL, Alexander G, Follastad-Shah J, Hassett B, Jenkinson R, Lave R, Rumps J, Pagano L. 2007. Restoring rivers one reach at a time: Results from a survey of US river restoration practitioners. Restor. Ecol., **15**: 482-493.

Blair JM. 1988. Nutrient release from decomposing foliar litter of three tree species with spicial reference to calcium, magnesium and potassium dynamics. Plant and Soil, **110**: 49-55.

Boas ML. 1983. Mathematical methods in the physical sciences. 2nd Edn., Wiley.

Boulton AJ. 2007. Hyporheic rehabilitation in rivers: restoring vertical connectivity. Freshw. Biol., **52**: 632-650. DOI: 10.1111/j.1365-2427.2006.01710.x.

Boulton AJ, Datry T, Kasahara T, Mutz M, Stanford JA. 2010. Ecology and management of the hyporheic zone: stream-groundwater interactions of running waters and their floodplains. J. N. Am. Benthol. Soc., **29**: 26-40. DOI: 10.1899/08-017.1.

Boulton AJ, Findlay S, Marmonier P, Stanley EH, Valett HM. 1998. The functional significance of the hyporheic zone in streams and rivers. Annu. Rev. Ecol. Syst., **29**: 59-81.

Box GEP, Jenkins GM, Reinsel GC. 1994. Time series analysis: forecasting and control. 3rd Edn., Prentice Hall.

Bredehoeft JD, Papadopulos IS. 1965. Rates of vertical groundwater movement estimated from the Earth's thermal profile. Water Resour. Res., **1**: 325-328. DOI: 10.1029/WR001i002p00325.

Briggs MA, Lautz LK, Hare DK, Gonzalez-Pinzon R. in revision. The influence of spatial and temporal hyporheic flux patterns on streambed biogeochemistry. Freshwater Science.

Briggs MA, Lautz LK, McKenzie JM, Gordon RP, Hare DK. 2012. Using high-resolution distributed temperature sensing to quantify spatial and temporal variability in vertical hyporheic flux. Water Resour. Res., **48**: W02527. DOI: 10.1029/2011WR011227.

Brown LE, Hannah DM, Milner AM, Soulsby C, Hodson AJ, Brewer MJ. 2006. Water source dynamics in a glacierized alpine river basin (Taillon-Gabietous, French Pyrenees). Water Resour. Res., **42**: W08404. DOI: 10.1029/2005wr004268.

Brunke M, Gonser T. 1997. The ecological significance of exchange processes between rivers and groundwater. Freshw. Biol., **37**: 1-33.

Buffington JM, Tonina D. 2009. Hyporheic exchange in mountain rivers II: Effects of channel morphology on mechanics, scales, and rates of exchange. Geography Compass, **3**: 1038-1062.

Bury JT, Mark BG, McKenzie JM, French A, Baraer M, Huh KI, Luyo MAZ, Lopez RJG. 2011. Glacier recession and human vulnerability in the Yanamarey watershed of the Cordillera Blanca, Peru. Climatic Change, **105**: 179-206. DOI: 10.1007/s10584-010-9870-1.

Buttle JM. 1994. Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins. Prog Phys Geog, **18**: 16-41. DOI: 10.1177/030913339401800102.

Caballero Y, Chevallier P, Gallaire R, Pillco R. 2004. Flow modelling in a high mountain valley equipped with hydropower plants: Rio zongo valley, Cordillera Real, Bolivia. Hydrol. Process., **18**: 939-957. DOI: 10.1002/hyp.1339.

Caissie D. 2006. The thermal regime of rivers: a review. Freshw. Biol., **51**: 1389-1406. DOI: 10.1111/j.1365-2427.2006.01597.x.

Champ DR, Gulens J, Jackson RE. 1979. Oxidation-reduction sequences in ground-water flow systems. Can. J. Earth Sci., **16**: 12-23.

Chemical Rubber Company (CRC). 2011. CRC Handbook of Chemistry and Physics (Online). CRC Press.

Clark ID, Fritz P. 1997. Environmental Isotopes in Hydrogeology. Lewis Publishers.

Clow DW, Schrott L, Webb R, Campbell DH, Torizzo A, Dornblaser M. 2003. Ground Water Occurrence and Contributions to Streamflow in an Alpine Catchment, Colorado Front Range. Ground Water, **41**: 937-950. DOI: 10.1111/j.1745-6584.2003.tb02436.x.

Conant B. 2004. Delineating and quantifying ground water discharge zones using streambed temperatures. Ground Water, **42**: 243-257.

Constantz J. 2008. Heat as a tracer to determine streambed water exchanges. Water Resour. Res., **44**: W00D10. DOI: 10.1029/2008wr006996.

Covino TP, McGlynn BL. 2007. Stream gains and losses across a mountain-to-valley transition: Impacts on watershed hydrology and stream water chemistry. Water Resour. Res., **43**: W10431. DOI: 10.1029/2006wr005544.

Covino TP, McGlynn BL, Mallard JM. 2011. Stream-groundwater exchange and hydrologic turnover at the network scale. Water Resour. Res., **47**: W12521. DOI: 10.1029/2011wr010942.

Crispell JK, Endreny TA. 2009. Hyporheic exchange flow around constructed in-channel structures and implications for restoration design. Hydrol. Process., **23**: 1158-1168. DOI: 10.1002/Hyp.7230.

Dakin JP, Pratt DJ, Bibby GW, Ross JN. 1985. Distributed Optical Fiber Raman Temperature Sensor Using a Semiconductor Light-Source and Detector. Electron. Lett., **21**: 569-570.

Day TJ. 1977. Field procedures and evaluation of a slug dilution gauging method in mountain streams. Journal of Hydrology New Zealand, **16**: 113-133.

Dethier DP. 1986. Weathering rates and chemical flux from catchments in the Pacific Northwest, USA. In: Rates of Chemical Weathering of Rocks and Minerals, Coleman SM, Dethier DP (eds.) Academic Press.

Doll BA, Grabow GL, Hall KR, Halley J, Harman WA, Jennings GD, Wise DE. 2003. Stream Restoration: A Natural Channel Design Handbook. N. C. Stream Restoration Inst., N. C. State Univ.

Duff JH, Triska FJ. 2000. Nitrogen biogeochemistry and surface-subsurface exchange in streams. In: Streams and Ground Waters, Jones JB, Mulholland PJ (eds.) Academic Press, pp: 425.

Earle LR, Warner BG, Aravena R. 2003. Rapid development of an unusual peat-accumulating ecosystem in the Chilean Altiplano. Quaternary Research, **59**: 2-11.

Elliott AH, Brooks NH. 1997. Transfer of nonsorbing solutes to a streambed with bed forms: Theory. Water Resour. Res., **33**: 123-136.

Endreny TE, Lautz LK, Siegel DI. 2011. Hyporheic flow path response to hydraulic jumps at river steps: Flume and hydrodynamic models. Water Resour. Res., **47**: W02517. DOI: 10.1029/2009wr008631.

Endreny TE, Lautz LK, Siegel DI. 2011. Hyporheic flow path response to hydraulic jumps at river steps: Hydrostatic model simulations. Water Resour. Res., **47**: W02518. DOI: 10.1029/2010wr010014.

Engelhardt I, Piepenbrink M, Trauth N, Stadler S, Kludt C, Schulz M, Schuth C, Ternes TA. 2011. Comparison of tracer methods to quantify hydrodynamic exchange within the hyporheic zone. J. Hydrol., **400**: 255-266. DOI: 10.1016/j.jhydrol.2011.01.033.

Evangelou VP, Wang J, Phillips RE. 1994. New eevelopments and perspectives on soil potassium quantity/intensity relationships. Adv. Agron., **52**: 173-227.

Fanelli RM, Lautz LK. 2008. Patterns of water, heat, and solute flux through streambeds around small dams. Ground Water, **46**: 671-687. DOI: 10.1111/j.1745-6584.2008.00461.x.

Ferguson G, Bense V. 2011. Uncertainty in 1D Heat-Flow Analysis to Estimate Groundwater Discharge to a Stream. Ground Water, **49**: 336-347. DOI: 10.1111/j.1745-6584.2010.00735.x.

Fetter CW. 2001. Applied Hydrogeology. 4th Edn., Prentice Hall.

Findlay S. 1995. Importance of Surface-Subsurface Exchange in Stream Ecosystems - the Hyporheic Zone. Limnol. Oceanogr., **40**: 159-164.

Gooseff MN. 2010. Defining hyporheic zones - advancing our conceptual and operational definitions of where stream water and groundwater meet. Geography Compass, **4**: 945-955.

Gooseff MN, Anderson JK, Wondzell S, LaNier J, Haggerty R. 2006. A modeling study of hyporheic exchange pattern and the sequence, size, and spacing of stream bedforms in mountain stream networks, Oregon, USA. Hydrol. Process., **20**: 2443-2457. DOI: 10.1002/Hyp.6350.

Gordon RP, Lautz LK, Briggs MA, McKenzie JM. 2012. Automated calculation of vertical pore-water flux from field temperature time series using the VFLUX method and computer program. J. Hydrol., **420–421**: 142-158.

Gordon RP, Lautz LK, Daniluk TL. 2013. Spatial patterns of hyporheic exchange and biogeochemical cycling around cross-vane restoration structures: Implications for stream restoration design. Water Resour. Res., **49**: 2040-2055. DOI: 10.1002/wrcr.20185.

Goto S, Yamano M, Kinoshita M. 2005. Thermal response of sediment with vertical fluid flow to periodic temperature variation at the surface. J. Geophys. Res.-Sol. Ea., **110**: B01106. DOI: 10.1029/2004jb003419.

Harvey JW, Fuller CC. 1998. Effect of enhanced manganese oxidation in the hyporheic zone on basin-scale geochemical mass balance. Water Resour. Res., **34**: 623-636. DOI: 10.1029/97wr03606.

Harvey JW, Wagner BJ. 2000. Quantifying hydrologic interactions between streams and their subsurface hyporheic zones. In: Streams and Ground Waters, Jones JB, Mulholland PJ (eds.) Academic Press, pp: 425.

Hatch CE, Fisher AT, Revenaugh JS, Constantz J, Ruehl C. 2006. Quantifying surface water-groundwater interactions using time series analysis of streambed thermal records: Method development. Water Resour. Res., **42**: W10410. DOI: 10.1029/2005wr004787.

Healy RW, Ronan AD. 1996. Documentation of computer program VS2DH for simulation of energy transport in variably saturated porous media—Modification of the U.S. Geological Survey's computer program VS2DT. In: U.S. Geological Survey Water-Resources Investigations Report 96-4230, pp: 36.

Hester ET, Doyle MW. 2008. In-stream geomorphic structures as drivers of hyporheic exchange. Water Resour. Res., **44**: W03417. DOI: 10.1029/2006wr005810.

Hester ET, Gooseff MN. 2010. Moving beyond the banks: Hyporheic restoration is fundamental to restoring ecological services and functions of streams. Environ. Sci. Technol., **44**: 1521-1525. DOI: 10.1021/Es902988n.

Hester ET, Gooseff MN. 2011. Hyporheic Restoration in Streams and Rivers. In: Stream Restoration in Dynamic Fluvial Systems: Scientific Approaches, Analyses, and Tools, AGU, pp: 167-187.

Hey RD. 2006. Fluvial geomorphological methodology for natural stable channel design. J. Am. Water Resour. Assoc., **42**: 357-374.

Hood JL, Roy JW, Hayashi M. 2006. Importance of groundwater in the water balance of an alpine headwater lake. Geophys. Res. Lett., **33**: L13405.

Hooper RP. 2003. Diagnostic tools for mixing models of stream water chemistry. Water Resour. Res., **39**: 1055. DOI: 10.1029/2002wr001528.

Hubbart J, Link T, Campbell C, Cobos D. 2005. Evaluation of a low-cost temperature measurement system for environmental applications. Hydrol. Process., **19**: 1517-1523. DOI: 10.1002/hyp.5861.

Huntington JL, Niswonger RG. 2012. Role of surface-water and groundwater interactions on projected summertime streamflow in snow dominated regions: An integrated modeling approach. Water Resour. Res., **48**: W11524. DOI: 10.1029/2012wr012319.

Huth AK, Leydecker A, Sickman JO, Bales RC. 2004. A two-component hydrograph separation for three high-elevation catchments in the Sierra Nevada, California. Hydrol. Process., **18**: 1721-1733. DOI: 10.1002/hyp.1414.

IAEA. 2013. Global Network of Isotopes in Precipitation (GNIP) Database. Accessible at: http://www.iaea.org/water.

Jensen JK, Engesgaard P. 2011. Nonuniform groundwater discharge across a streambed: heat as a tracer. Vadose Zone J., **10**: 98-109. DOI: 10.2136/vzj2010.0005.

Johnson AN, Boer BR, Woessner WW, Stanford JA, Poole GC, Thomas SA, O'Daniel SJ. 2005. Evaluation of an inexpensive small-diameter temperature logger for documenting ground water-river interactions. Ground Water Monit. Remed., **25**: 68-74.

Jones JB, Fisher SG, Grimm NB. 1995. Nitrification in the hyporheic zone of a desert stream ecosystem. J. N. Am. Benthol. Soc., **14**: 249-258. DOI: 10.2307/1467777.

Kasahara T, Hill AR. 2006. Effects of riffle-step restoration on hyporheic zone chemistry in N-rich lowland streams. Can. J. Fish. Aquat. Sci., **63**: 120-133. DOI: 10.1139/F05-199.

Kasahara T, Hill AR. 2006. Hyporheic exchange flows induced by constructed riffles and steps in lowland streams in southern Ontario, Canada. Hydrol. Process., **20**: 4287-4305. DOI: 10.1002/Hyp.6174.

Kasahara T, Hill AR. 2008. Modeling the effects of lowland stream restoration projects on stream-subsurface water exchange. Ecol. Eng., **32**: 310-319. DOI: 10.1016/j.ecoleng.2007.12.006.

Kasahara T, Wondzell SM. 2003. Geomorphic controls on hyporheic exchange flow in mountain streams. Water Resour. Res., **39**: 1005. DOI: 10.1029/2002wr001386.

Kaser DH, Binley A, Heathwaite AL. 2012. On the importance of considering channel microforms in groundwater models of hyporheic exchange. River Res. Appl.: Published online, doi:10.1002/rra.1618. DOI: 10.1002/rra.1618.

Kaser G, Ames A, Zamora M. 1990. Glacier fluctuations and climate in the Cordillera Blanca, Peru. Annals of Glaciology, **14**: 136-140.

Kaser G, Cogley JG, Dyurgerov MB, Meier MF, Ohmura A. 2006. Mass balance of glaciers and ice caps: Consensus estimates for 1961-2004. Geophys. Res. Lett., **33**: L19501. DOI: 10.1029/2006gl027511.

Kaser G, Juen I, Georges C, Gomez J, Tamayo W. 2003. The impact of glaciers on the runoff and the reconstruction of mass balance history from hydrological data in the tropical Cordillera Blanca, Peru. J. Hydrol., **282**: 130-144. DOI: 10.1016/s0022-1694(03)00259-2.

Kaser G, Osmaston H. 2002. Tropical glaciers. Cambridge University Press.

Keery J, Binley A, Crook N, Smith JWN. 2007. Temporal and spatial variability of groundwater-surface water fluxes: Development and application of an analytical method using temperature time series. J. Hydrol., **336**: 1-16. DOI: 10.1016/j.jhydrol.2006.12.003.

Kilpatrick FA, Cobb ED. 1985. Measurement of discharge using tracers. In: Techniques of Water Resources Investigations of the U.S. Geological Survey, United States Government Printing Office, pp: 52.

Langston G, Bentley LR, Hayashi M, McClymont A, Pidlisecky A. 2011. Internal structure and hydrological functions of an alpine proglacial moraine. Hydrol. Process., **25**: 2967-2982. DOI: 10.1002/hyp.8144.

Lapham WW. 1989. Use of temperature profiles beneath streams to determine rates of vertical ground-water flow and vertical hydraulic conductivity. U.S. Geol. Surv. Water Supply Pap. 2337.

Lautz LK. 2010. Impacts of nonideal field conditions on vertical water velocity estimates from streambed temperature time series. Water Resour. Res., **46**: W01509. DOI: 10.1029/2009wr007917.

Lautz LK. 2012. Observing temporal patterns of vertical flux through streambed sediments using time-series analysis of temperature records. J. Hydrol., **464-465**: 199-215. DOI: 10.1016/j.jhydrol.2012.07.006.

Lautz LK, Fanelli RM. 2008. Seasonal biogeochemical hotspots in the streambed around restoration structures. Biogeochemistry, **91**: 85-104. DOI: 10.1007/s10533-008-9235-2.

Lautz LK, Kranes NT, Siegel DI. 2010. Heat tracing of heterogeneous hyporheic exchange adjacent to in-stream geomorphic features. Hydrol. Process., **24**: 3074-3086. DOI: 10.1002/hyp.7723.

Lautz LK, Ribaudo RE. 2012. Scaling up point-in-space heat tracing of seepage flux using bed temperatures as a quantitative proxy. Hydrogeol. J., **20**: 1223-1238. DOI: 10.1007/s10040-012-0870-2.

Lautz LK, Siegel DI. 2006. Modeling surface and ground water mixing in the hyporheic zone using MODFLOW and MT3D. Adv. Water Resour., **29**: 1618-1633. DOI: 10.1016/j.advwatres.2005.12.003.

Lave R. 2009. The Controversy Over Natural Channel Design: Substantive Explanations and Potential Avenues for Resolution1. J. Am. Water Resour. Assoc., **45**: 1519-1532. DOI: 10.1111/j.1752-1688.2009.00385.x.

Liu FJ, Williams MW, Caine N. 2004. Source waters and flow paths in an alpine catchment, Colorado Front Range, United States. Water Resour. Res., **40**: W09401. DOI: 10.1029/2004wr003076.

Love DA, Clark AH, Glover JK. 2004. The lithologic, stratigraphic, and structural setting of the giant antamina copper-zinc skarn deposit, Ancash, Peru. Economic Geology, **99**: 887-916. DOI: 10.2113/econgeo.99.5.887.

Lumbanraja J, Evangelou VP. 1990. Binary and ternary exchange behavior of potassium and ammonium on Kentucky subsoils. Soil Sci. Soc. Am. J., **54**: 698-705.

Maassen S, Balla D. 2010. Impact of hydrodynamics (ex- and infiltration) on the microbially controlled phosphorus mobility in running water sediments of a cultivated northeast German wetland. Ecol. Eng., **36**: 1146-1155. DOI: 10.1016/j.ecoleng.2010.01.009.

Malakoff D. 2004. Profile - Dave Rosgen - The river doctor. Science, **305**: 937-939.

Mark BG, Bury J, McKenzie JM, French A, Baraer M. 2010. Climate Change and Tropical Andean Glacier Recession: Evaluating Hydrologic Changes and Livelihood Vulnerability in the Cordillera Blanca, Peru. Ann. Assoc. Am. Geogr., **100**: 794-805.

Mark BG, McKenzie JM. 2007. Tracing increasing tropical Andean glacier melt with stable isotopes in water. Environ. Sci. Technol., **41**: 6955-6960. DOI: 10.1021/Es071099d.

Mark BG, McKenzie JM, Gomez J. 2005. Hydrochemical evaluation of changing glacier meltwater contribution to stream discharge: Callejon de Huaylas, Peru. Hydrolog Sci J, **50**: 975-987.

Mark BG, Seltzer GO. 2003. Tropical glacier meltwater contribution to stream discharge: a case study in the Cordillera Blanca, Peru. J. Glaciol., **49**: 271-281.

McClymont AF, Hayashi M, Bentley LR, Muir D, Ernst E. 2010. Groundwater flow and storage within an alpine meadow-talus complex. Hydrol. Earth Syst. Sci., **14**: 859-872. DOI: 10.5194/hess-14-859-2010.

McClymont AF, Roy JW, Hayashi M, Bentley LR, Maurer H, Langston G. 2011. Investigating groundwater flow paths within proglacial moraine using multiple geophysical methods. J. Hydrol., **399**: 57-69. DOI: 10.1016/j.jhydrol.2010.12.036.

McKnight DM, Runkel RL, Tate CM, Duff JH, Moorhead DL. 2004. Inorganic N and P dynamics of Antarctic glacial meltwater streams as controlled by hyporheic exchange and benthic autotrophic communities. J. N. Am. Benthol. Soc., **23**: 171-188.

Mencio A, Mas-Pla J. 2008. Assessment by multivariate analysis of groundwater-surface water interactions in urbanized Mediterranean streams. J. Hydrol., **352**: 355-366. DOI: 10.1016/j.jhydrol.2008.01.014.

Messerli B. 2001. Editorial: The International Year of Mountains (IYM), the Mountain Research Initiative (MRI) and PAGES. PAGES News, **9**: 2.

Milner AM, Brown LE, Hannah DM. 2009. Hydroecological response of river systems to shrinking glaciers. Hydrol. Process., **23**: 62-77. DOI: 10.1002/Hyp.7197.

Morrice JA, Dahm CN, Valett HM, Unnikrishna PV, Campana ME. 2000. Terminal electron accepting processes in the alluvial sediments of a headwater stream. J. N. Am. Benthol. Soc., **19**: 593-608. DOI: 10.2307/1468119.

Muir DL, Hayashi M, McClymont AF. 2011. Hydrological storage and transmission characteristics of an alpine talus. Hydrol. Process., **25**: 2954-2966. DOI: 10.1002/hyp.8060.

Nagle G. 2007. Evaluating 'natural channel design' stream projects. Hydrol. Process., **21**: 2539-2545. DOI: 10.1002/hyp.6840.

Niswonger R, Prudic DE. 2003. Modeling heat as a tracer to estimate streambed seepage and hydraulic conductivity. In: Heat as a Tool for Studying the Movement of Ground Water Near Streams, U. S. Geological Servey Circular 1260, Constantz J, Stonestrom DA (eds.), pp: 81-89.

Nogues-Bravo D, Araujo MB, Errea MP, Martinez-Rica JP. 2007. Exposure of global mountain systems to climate warming during the 21st Century. Global Environ Chang, **17**: 420-428. DOI: 10.1016/j.gloenvcha.2006.11.007.

NRCS. 2007. Stream Restoration Design: National Engineering Handbook. U.S. Department of Agriculture National Resources Conservation Service.

Payn RA, Gooseff MN, McGlynn BL, Bencala KE, Wondzell SM. 2009. Channel water balance and exchange with subsurface flow along a mountain headwater stream in Montana, United States. Water Resour. Res., **45**: W11427. DOI: 10.1029/2008wr007644.

Poole GC. 2010. Stream hydrogeomorphology as a physical science basis for advances in stream ecology. J. N. Am. Benthol. Soc., **29**: 12-25. DOI: 10.1899/08-070.1.

Porta JM, Verbeek JJ, Krose BJA. 2005. Active appearance-based robot localization using stereo vision. Autonomous Robots, **18**: 59-80. DOI: 10.1023/B:AURO.0000047287.00119.b6.

Pretty JL, Hildrew AG, Trimmer M. 2006. Nutrient dynamics in relation to surface-subsurface hydrological exchange in a groundwater fed chalk stream. J. Hydrol., **330**: 84-100. DOI: 10.1016/j.jhydrol.2006.04.013.

Rau GC, Andersen MS, McCallum AM, Acworth RI. 2010. Analytical methods that use natural heat as a tracer to quantify surface water-groundwater exchange, evaluated using field temperature records. Hydrogeol. J., **18**: 1093-1110. DOI: 10.1007/s10040-010-0586-0.

Ronan AD, Prudic DE, Thodal CE, Constantz J. 1998. Field study and simulation of diurnal temperature effects on infiltration and variably saturated flow beneath an ephemeral stream. Water Resour. Res., **34**: 2137-2153. DOI: 10.1029/98wr01572.

Rosgen DL. 2001. The cross-vane, W-weir, and J-hook vane structures: Their description, design and application for stream stabilization and river restoration. In: Wetlands Engineering and River Restoration Conference, Hayes DF (ed.) American Society of Civil Engineers, pp: 22.

Roweis ST. 1998. EM Algorithms for PCA and SPCA. In: Advances in Neural Information Processing Systems, Jordan MI, Kearns MJ, Solla SA (eds.) MIT Press, pp: 626-632.

Roy JW, Hayashi M. 2009. Multiple, distinct groundwater flow systems of a single moraine-talus feature in an alpine watershed. J. Hydrol., **373**: 139-150. DOI: 10.1016/j.jhydrol.2009.04.018.

Sawyer AH, Cardenas MB, Buttles J. 2011. Hyporheic exchange due to channel-spanning logs. Water Resour. Res., **47**: W08502. DOI: 10.1029/2011wr010484.

Schmadel NM, Neilson BT, Stevens DK. 2010. Approaches to estimate uncertainty in longitudinal channel water balances. J. Hydrol., **394**: 357-369. DOI: 10.1016/j.jhydrol.2010.09.011.

Schmidt C, Conant B, Bayer-Raich M, Schirmer M. 2007. Evaluation and field-scale application of an analytical method to quantify groundwater discharge using mapped streambed temperatures. J. Hydrol., **347**: 292-307. DOI: 10.1016/j.jhydrol.2007.08.022.

Schmidt C, Martienssen M, Kalbus E. 2011. Influence of water flux and redox conditions on chlorobenzene concentrations in a contaminated streambed. Hydrol. Process., **25**: 234-245. DOI: 10.1002/hyp.7839.

Schornberg C, Schmidt C, Kalbus E, Fleckenstein JH. 2010. Simulating the effects of geologic heterogeneity and transient boundary conditions on streambed temperatures - Implications for temperature-based water flux calculations. Adv. Water Resour., **33**: 1309-1319. DOI: 10.1016/j.advwatres.2010.04.007.

Selker JS, Thevenaz L, Huwald H, Mallet A, Luxemburg W, de Giesen NV, Stejskal M, Zeman J, Westhoff M, Parlange MB. 2006. Distributed fiber-optic temperature sensing for hydrologic systems. Water Resour. Res., **42**: W12202. DOI: 10.1029/2006wr005326.

Shanafield M, Hatch C, Pohll G. 2011. Uncertainty in thermal time series analysis estimates of streambed water flux. Water Resour. Res., **47**. DOI: 10.1029/2010wr009574.

Squeo FA, Warner BG, Aravena R, Espinoza D. 2006. Bofedales: high altitude peatlands of the central Andes. Rev Chil Hist Nat, **79**: 245-255.

Stallman RW. 1965. Steady one-dimensional fluid flow in a semi-infinite porous medium with sinusoidal surface temperature. J. Geophys. Res., **70**: 2821-2827.

Stonestrom DA, Constantz J. 2003. Heat as a tool for studying the movement of ground water near streams. In: U.S. Geol. Surv. Circ. 1260, pp: 96.

Stream Solute Workshop. 1990. Concepts and methods for assessing solute dynamics in stream ecosystems. J. N. Am. Benthol. Soc., **9**: 95-119.

Sueker JK, Ryan JN, Kendall C, Jarrett RD. 2000. Determination of hydrologic pathways during snowmelt for alpine/subalpine basins, Rocky Mountain National Park, Colorado. Water Resour. Res., **36**: 63-75.

Swanson TE, Cardenas MB. 2010. Diel heat transport within the hyporheic zone of a pool-riffle-pool sequence of a losing stream and evaluation of models for fluid flux estimation using heat. Limnol. Oceanogr., **55**: 1741-1754. DOI: 10.4319/lo.2010.55.4.1741.

Swanson TE, Cardenas MB. 2011. Ex-Stream: A MATLAB program for calculating fluid flux through sediment-water interfaces based on steady and transient temperature profiles. Computers & Geosciences, **37**: 1664-1669. DOI: 10.1016/j.cageo.2010.12.001.

Tipping ME, Bishop CM. 1999. Probabilistic principal component analysis. Journal of the Royal Statistical Society, Series B (Statistical Methodology), **61**: 611-622.

Tonina D, Buffington JM. 2009. Hyporheic exchange in mountain rivers I: Mechanics and environmental effects. Geography Compass, **3**: 1063-1086.

Triska FJ, Kennedy VC, Avanzino RJ, Zellweger GW, Bencala KE. 1989. Retention and transport of nutrients in a 3rd-order stream in Northwestern California: Hyporheic processes. Ecology, **70**: 1893-1905.

Valett HM, Morrice JA, Dahm CN, Campana ME. 1996. Parent lithology, surface-groundwater exchange, and nitrate retention in headwater streams. Limnol. Oceanogr., **41**: 333-345.

Viviroli D, Durr HH, Messerli B, Meybeck M, Weingartner R. 2007. Mountains of the world, water towers for humanity: Typology, mapping, and global significance. Water Resour. Res., **43**. DOI: 10.1029/2006wr005653.

Vogel T, Brezina J, Dohnal M, Dusek J. 2010. Physical and Numerical Coupling in Dual-Continuum Modeling of Preferential Flow. Vadose Zone J., **9**: 260-267. DOI: 10.2136/vzj2009.0091.

Vogel T, Dohnal M, Votrubova J. 2011. Modeling heat fluxes in macroporous soil under sparse young forest of temperate humid climate. J. Hydrol., **402**: 367-376. DOI: 10.1016/j.jhydrol.2011.03.030.

Vogt T, Schneider P, Hahn-Woernle L, Cirpka OA. 2010. Estimation of seepage rates in a losing stream by means of fiber-optic high-resolution vertical temperature profiling. J. Hydrol., **380**: 154-164. DOI: 10.1016/j.jhydrol.2009.10.033.

Vuille M, Francou B, Wagnon P, Juen I, Kaser G, Mark BG, Bradley RS. 2008. Climate change and tropical Andean glaciers: Past, present and future. Earth-Sci Rev, **89**: 79-96. DOI: 10.1016/j.earscirev.2008.04.002.

Wagnon P, Ribstein P, Schuler T, Francou B. 1998. Flow separation on Zongo Glacier, Cordillera Real, Bolivia. Hydrol. Process., **12**: 1911-1926. DOI: 10.1002/(sici)1099-1085(19981015)12:12<1911::aid-hyp673>3.0.co;2-h.

White DS. 1993. Perspectives on defining and delineating hyporheic zones. J. N. Am. Benthol. Soc., **12**: 61-69.

White DS, Elzinga CH, Hendricks SP. 1987. Temperature patterns within the hyporheic zone of a northern Michigan river. J. N. Am. Benthol. Soc., **6**: 85-91.

Wilson J, Molina O, Fernandez AWS. 1995. Mapa geologico del cuadrangulo de Carhuaz. Instituto Geologico Minero y Metalurgico, Ministerio de Energia y Minas, Republica del Peru.

Winter TC, Harvey JW, Franke OL, Alley WM. 1998. Ground water and surface water: a single resource. U.S. Geological Survey.

Wondzell SM. 2011. The role of the hyporheic zone across stream networks. Hydrol. Process., **25**: 3525-3532. DOI: 10.1002/hyp.8119.

Woocay A, Walton J. 2008. Multivariate analyses of water chemistry: Surface and ground water interactions. Ground Water, **46**: 437-449. DOI: 10.1111/j.1745-6584.2007.00404.x.

Young PC, Pedregal DJ, Tych W. 1999. Dynamic harmonic regression. J. Forecast., **18**: 369-394.

Young PC, Taylor CJ, Tych W, Pegregal DJ, McKenna PG. 2010. The Captain Toolbox. Centre for Research on Environmental Systems and Statistics, Lancaster University, UK. Internet: http://www.es.lancs.ac.uk/cres/captain.

Zarnetske JP, Haggerty R, Wondzell SM, Baker MA. 2011. Dynamics of nitrate production and removal as a function of residence time in the hyporheic zone. J. Geophys. Res.-Biogeo., **116**. DOI: 10.1029/2010jg001356.

Biographical Information

Author: Ryan P. Gordon

EDUCATION

Williams College, Williamstown, Massachusetts *2005*
B.A. *Summa cum Laude* with Highest Honors in Geosciences

EMPLOYMENT and ACADEMIC APPOINTMENTS

Water Science Fellow, Department of Earth Sciences, Syracuse University	*2013 – present*
NSF Graduate Research Fellow, Department of Earth Sciences, Syracuse University	*2010 – 2013*
Research Assistant, Department of Earth Sciences, Syracuse University	*2009 –2010*
GIS Assistant, The Nature Conservancy, Brunswick, Maine	*2008*
Naturalist, Appalachian Mountain Club, Gorham, New Hampshire	*2007*
Staff Geologist, GeoDesign, Inc., Portland, Oregon	*2005 –2006*
Wilderness Ranger, U.S. Forest Service, Enterprise, Oregon	*2005*
Summer Science Research Fellow, Department of Geosciences, Williams College	*2004 –2005*
Research Assistant, Hopkins Memorial Forest, Williams College	*2003*

AWARDS and HONORS

GSA Outstanding Student Research Award and Hydrogeology Division Research Award	*2012*
GSA Graduate Student Research Grant	*2012*
Student Publication Award, Syracuse University Earth Sciences	*2012*
CUAHSI Pathfinder Graduate Student Fellowship	*2011*
NSF Graduate Research Fellowship	*2010*
Marjorie Hooker Award, Syracuse University Earth Sciences	*2010*
Sigma Xi Scientific Honor Society	*2005*
Phi Beta Kappa Honor Society	*2004*
Willis I. Milham Scholarship, Williams College	*2004 – 2005*
Class of 1960 Scholar, Williams College	*2005*

PUBLICATIONS

2013, **Gordon, R.P.**, L.K. Lautz, and T.L. Daniluk, Spatial patterns of hyporheic exchange and biogeochemical cycling around cross-vane restoration structures: Implications for stream restoration design, Water Resources Research, 49: 1-16, doi: 10.1002/wrcr.20185.

2012, Daniluk, T.L., L.K. Lautz, and **R.P. Gordon**, Surface water–groundwater interaction at restored streams and associated reference reaches, Hydrological Processes, published online Aug. 30, doi: 10.1002/hyp.9501.

2012, **Gordon, R.P.**, L.K. Lautz, M.A. Briggs, and J.M. McKenzie, Automated calculation of vertical pore-water flux from field temperature time series using the VFLUX method and computer program, Journal of Hydrology, 420-421: 142-158, doi: 10.1016/j.jhydrol.2011.11.053.

2012, Briggs, M.A., L.K. Lautz, J.M. McKenzie, **R.P. Gordon**, and D.K. Hare, Using high-resolution distributed temperature sensing to quantify spatial and temporal variability in vertical hyporheic flux, Water Resources Research, 48: W02527, doi: 10.1029/2011WR011227.

2005, **Gordon, R.P.**, Kinematics and Chronology of Faulting in the Western Berkshire Massif, Massachusetts, B.A. Honors Thesis, Williams College, Williamstown, Massachusetts. Advised by Dr. Paul Karabinos.

PROFESSIONAL PRESENTATIONS and CONFERENCE PROCEEDINGS

2013, **Gordon, R.P.**, L.K. Lautz, J.M. McKenzie, B.G. Mark, and D. Chavez, Groundwater, springs, and stream flow generation in an alpine meadow of a tropical glacierized catchment, Abstract presented at 2013 Fall Meeting, AGU, San Francisco, Calif., 9-13 Dec.

2013, **Gordon, R.P.**, and L.K. Lautz, Hyporheic exchange and biogeochemical cycling around stream restoration structures: Implications for the health of restored streams, Geological Society of America Abstracts with Programs, v. 45, p. 133.

2012, **Gordon, R.P.**, L.K. Lautz, J.M. McKenzie, and B.G. Mark, Groundwater-surface water interactions in a glacierized catchment and their influence on proglacial water supply, Abstract GC52B-06 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.

2011, **Gordon, R.P.**, L.K. Lautz, M.A. Briggs, and J.M. McKenzie, Automated calculation of vertical pore-water flux from real-world temperature time series using the VFLUX method and computer program, Geological Society of America Abstracts with Programs, v. 43, p. 390.

2011, Briggs, M.A., L.K. Lautz, **R.P. Gordon**, M.J. McKenzie, and D.K. Hare, Evolving hyporheic exchange flux during baseflow recession: using high-resolution heat data to quantitatively assess temporal patterns, Geological Society of America Abstracts with Programs, v. 43, p. 389.

2010, **Gordon, R.P.**, L.K. Lautz, and T. Daniluk, Spatial patterns of hyporheic flow and biogeochemical cycling around cross-vane restoration structures, Abstract H21B-1034 presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 Dec.

2010, Daniluk, T., L.K. Lautz, and **R.P. Gordon**, Water, heat and solute fluxes through hyporheic zones at stream restoration sites and their associated reference stream, Abstract H21B-1045 presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 Dec.

2010, **Gordon, R.**, L. Lautz, and T. Daniluk, Hyporheic flow, solute transport, and heat flux in the stream bed around cross-vane restoration structures, Geophysical Research Abstracts, v. 12, EGU2010-13113, presented at 2010 EGU General Assembly, Vienna, Austria, 3-7 May.

2010, Lautz, L., R. Fanelli, **R. Gordon**, and T. Daniluk, Spatial patterns of water, heat and solute fluxes through the hyporheic zone at stream restoration sites, Geophysical Research Abstracts, v. 12, EGU2010-1730, presented at 2010 EGU General Assembly, Vienna, Austria, 3-7 May.

2007, Karabinos, P., D.J. Morris, **R.P. Gordon**, and J.M. Pyle, Intracrustal wedging and emplacement of external basement massifs in the northern Appalachians, Geological Society of America Abstracts with Programs, v. 39, p. 37.

2006, Karabinos, P., D.J. Morris, **R.P. Gordon**, and J.M. Pyle, Geochronological constraints of the emplacement of basement uplifts in the northern Appalachians, Geological Society of America Abstracts with Programs, v. 38, p. 63.

2005, Karabinos, P., **R.P. Gordon**, and J.M. Pyle, Age and style of emplacement of basement uplifts in the northern Appalachians, Geological Society of America Abstracts with Programs, v. 37, p. 20.

2005, **Gordon, R.P.**, and P. Karabinos, Thrusting in the Berkshire Massif, western New England, Geological Society of America Abstracts with Programs, v. 37, p. 72-73.