Measuring groundwater-surface water interaction and its effect on wetland stream benthic productivity, Trout Lake watershed, northern Wisconsin, USA

Randall J. Hunt\(^1\), Mac Strand\(^2\), and John F. Walker\(^1\)

\(^*\)Corresponding Author
\(^1\)US Geological Survey
8505 Research Lane
Middleton, WI  53562 USA
rjhunt@usgs.gov, jfwalker@usgs.gov
fax:  (608) 821-3817

\(^2\)Department of Biology
Northern Michigan University
1401 Presque Isle Avenue
Marquette, MI  49855 USA
rostrand@nmu.edu

\(^\d\)This article is a U.S. Government work and is in the public domain in the USA
Abstract

Measurements of groundwater-surface water exchange at three wetland stream sites were related to patterns in benthic productivity as part of the U.S. Geological Survey’s Northern Temperate Lakes Water, Energy and Biogeochemical Budgets (NTL-WEBB) project. The three sites included one high groundwater discharge (HGD) site, one weak groundwater discharge (WGD) site, and one groundwater recharge (GR) site. Large upward vertical gradients were associated with smallest variation in head below the stream and fewest gradient reversals between the stream and the groundwater beneath the stream, and the stream and the adjacent streambank. The WGD site had the highest number of gradient reversals reflecting the average condition being closest to zero vertical gradient. The durations of groundwater discharge events were also related to the amount of discharge where the HGD site had the longest strong-gradient durations for both horizontal and vertical groundwater flow. Strong groundwater discharge also controlled transient temperature and chemical hyporheic conditions by limiting the infiltration of surface water.

Groundwater-surface water interactions were related to highly significant patterns in benthic invertebrate abundance, taxonomic richness, and periphyton respiration with the WGD site having greater abundance (27%), richness (19%) and periphyton respiration (39%) than the GR site. The HGD site abundance was 35% greater than in the WGD site and 53% greater than the GR site; richness and periphyton respiration were also significantly greater (p ≤ 0.001, 31% and 44% respectively) in the HGD site than in the GR site. This work suggests groundwater-surface water interactions can strongly influence benthic productivity, thus emphasizing the importance of quantitative hydrology for management of wetland-stream ecosystems in the northern temperate lakes region.

Keywords: groundwater-surface water interaction, wetland, upwelling, benthic invertebrates, periphyton respiration
Introduction

As one of five sites in the U.S. Geological Survey’s Water, Energy, and Biogeochemical Budgets (WEBB) Program, the Northern Temperate Lakes (NTL) WEBB project is working to understand processes controlling hydrologic and biogeochemical fluxes at different spatial and temporal scales (Walker and Bullen, 2000). The watershed hydrology of the NTL-WEBB site has been investigated using groundwater flow modeling (Cheng and Anderson, 1994; Hunt et al., 1998; Pint, 2002; Hunt et al., 2003) and geochemical/isotope tracing of water (Walker and Krabbenhoft, 1998; Walker et al., 2003). This work has shown that the variety of sources of water (terrestrial and lake recharge) can lead to complex flow systems with a variety of age and provenance occurring where groundwater discharges to the streams (Pint et al., 2003). While several single lake hydrogeologic studies have from this area (e.g., Wentz and Rose, 1989; Krabbenhoft and Babiarz, 1992; Rose, 1993) have demonstrated the importance of groundwater-surface water interactions (Hurley et al., 1985; Lodge et al., 1989; Hagerthey and Kerfoot, 1998), there have been fewer studies that have focused on groundwater-surface water interaction in wetland streams. These studies have addressed the centimeter-scale hydrogeochemistry of the stream hyporheic zone (Schindler and Krabbenhoft, 1998), the effect of wetland area on nutrients exported by streams in the basin (Elder et al., 2000, Elder et al., 2001), and have characterized the hydrology of a single wetland system (Marin, 1986). However, in many cases there is a need to understand how these abiotic characteristics affect biological systems (Hunt and Wilcox, 2003).

Wetland streams in the Trout Lake basin typically receive appreciable groundwater inputs, some that is focused in discrete reaches. Fisheries biologists in this region have long recognized these microhabitats as critical to brook trout (Salvelinus fontinalis), which require the relatively warm, silt-free waters of upwelling zones during winter for embryonic development (Becker, 1983) and the relatively cool summer conditions to escape temperature stress in the main channel.
as juveniles and adults (Biro, 1998; Baird and Krueger, 2003). Some bottom-dwelling invertebrates also use upwelling zones for critical development stages (Pugsley and Hynes, 1986) and many are likely to thrive in the highly stable thermal environments (Vannote and Sweeney, 1980). Although the relation of microhabitat quality for benthic invertebrates to groundwater flow is not well documented (Boulton et al., 1998), enhanced epibenthic algae growth in this region (Hagerthey and Kerfoot, 1998) and elsewhere (Coleman and Dahm, 1990; Dent et al., 2000) with increasing groundwater flow suggests that quantity of groundwater could indirectly control the abundance and species composition of grazing invertebrates, and benthic communities in general (Ward, 1989; Boulton, 1993; Brunke and Gonser, 1997).

The upwelling productivity effect has been attributed to microhabitat-scale differences in groundwater flux and resulting variation in stream water temperature dynamics, sediment porosity, and nutrient transport rates (Brunke and Gonser, 1997; Dent et al., 2000; Holmes, 2000). These environmental characteristics exert a combination of direct and indirect influences on stream dwelling organisms (Hynes, 1970) that could generate inter-reach productivity gradients driven by variation in hydrologic exchange. If the influence of groundwater discharge is strong enough to create localized high-quality patches for animal populations, then reach-scale hydrologic understanding (and quantitative models) could be used to predict relative biological activity at multiple scales from site to landscape.

The purpose of this study is to assess the spatial and temporal variability of local groundwater interaction with wetland-associated streams at three sites in the Trout Lake watershed, and to relate that variability to the macroinvertebrate and periphyton communities measured at the site. This work focuses on local interactions and biological communities; implications for a larger watershed context are extrapolated from this smaller scale.
Study Area

The NTL-WEBB study area is located within the Trout Lake watershed, in the Northern Highlands Lake District of north-central Wisconsin, USA. NTL-WEBB research efforts have focused primarily on Allequash Creek, a large watershed in the Trout Lake basin (Figure 1), but have also included comparison studies at two other watersheds in the basin, Stevenson and North Creeks (Figure 1). The NTL WEBB study area is co-located with one of the National Science Foundation’s Long-Term Ecological Research (LTER) sites. The North Temperate Lakes LTER site is operated by the Center for Limnology at the University of Wisconsin (Magnuson et al., 1984), and focuses primarily on seven lakes in the Trout Lake area.

The Trout Lake watershed has a surface-water drainage area of 10,500 hectares and ranges in elevation from approximately 490 to 520 meters above mean sea level. Mean monthly temperatures are minimum in January, ranging from -17 to -6 °C, and are maximum in August, ranging from +13 to +26 °C. The lakes in the area are ice-free for about 7 months per year. Precipitation averages 79 cm/yr (Cheng and Anderson, 1994), recharge to the groundwater system averages 27 cm/yr (Hunt et al., 1998); thus, evapotranspiration and canopy interception combined averages 52 cm/yr. As a result of precipitation being greater than the evapotranspiration, the streams are dominated by groundwater discharge and baseflow comprises >92% of the annual streamflow (USGS, unpublished data). In addition, the poor drainage associated with glaciated landscape results in appreciable wetland areas in the watershed, ranging from bogs to fens. The climate at the study site is affected by air masses from the North Pacific, the North Atlantic, the Gulf of Mexico, the Arctic, the High Plains, and the Ohio Valley. Located about 70 km southeast of Lake Superior, and 200 km northwest of Lake Michigan, the NTL WEBB site is also under strong climatic influence from the Great Lakes.
The principal aquifer in the watershed is comprised of a sandy outwash plain consisting of 30-50 m of unconsolidated sand and gravel overlying Precambrian igneous bedrock (Okwueze, 1983; Attig, 1985). The predominant soils are thin, poorly developed, sands, loams and silts (loamy sand, sandy loam, loamy fine sand, and sand make up 86% of the soils) with most organic content present in the upper horizon. The glacial drift is believed to be essentially carbonate free; thus, the groundwater chemistry is largely controlled by silicate hydrolysis (Kenoyer and Bowser, 1992; Bullen et al., 1996) and typically has low concentrations of ions and nutrients. However, extended water-mineral contact times associated with long groundwater flowpaths can result in appreciably higher ion concentrations (Walker et al., 2003; Pint et al., 2003). Moreover, shallow hyporheic areas can have appreciably higher localized concentrations than the larger groundwater system (Schindler and Krabbenhof, 1998).

Three stream-wetland sites in the Trout Lake watershed (Figure 1) were instrumented and operated during 2001. For clarity, sites having ground water heads above stream stage are called “discharge” sites; sites having ground water heads below stream stage are called “recharge” sites. Based on the analytic element model of Hunt et al. (1998), the sites were chosen to encompass a high groundwater discharge (HGD) at the Allequash Creek, a weak groundwater discharge (WGD) at the North Creek, and a groundwater recharge (GR) site at the Stevenson Creek (Figure 1). The wetland plant communities at the sites are shrub-carr and fen communities dominated primarily by Carex spp. and Alnus rugosa. Mean baseflows in the wetland streams are similar, typically ranging from about 0.09 m$^3$/s at the GR Site to approximately 0.13 m$^3$/s at the HGD Site (USGS, unpublished data).

Previous work has quantified stream water nutrient concentrations at these sites. The NTL-WEBB project routinely sampled the streams during 1991-1995, and sampled quarterly thereafter. Using these data, Elder et al. (2001) reported similar total phosphorus and SRP yields
and retentions in the Allequash, North, and Stevenson watersheds that were attributed to similar land use, geologic setting, and dominance of baseflow over storm-related runoff. Dissolved organic and inorganic carbon concentrations in stream water were also similar, and showed no trends when Allequash Creek was sampled longitudinally (Elder et al., 2000). Nitrogen concentrations were also similar at the three sites (USGS, unpublished data).

Stream water samples were collected during the 2001 study period (n=4) and showed similar nutrient concentrations (Figure 2) during June and July (Univ. of Wisconsin-Madison Center for Limnology, unpublished data). Soluble reactive phosphorus (SRP) concentrations – often considered the limiting nutrient in freshwater systems – are low in the water of all three streams (Figure 2), though concentrations at the HGD site appear to be slightly higher than at the GR site. This difference may be an artifact of having only two samples available for this site versus three samplings at the other two sites. Stream waters are thought to be phosphorus and nitrogen limited based on a cursory nutrient limitation study at the three sites during June-August 2001 (Jacques Finley, Univ. of Minnesota, written communication, Sept. 5, 2003).

Nutrient chemistry in porewater 60 cm below the streams at these three sites during June and July 2001 (Univ. of Wisconsin Center for Limnology, unpublished data) was more variable between sites, with nutrient concentrations in the groundwater higher or near the concentrations measured in the stream (Figure 2). Moreover, a relation between the groundwater velocity and groundwater porewater chemistry was observed such that the HGD site was associated with higher concentrations of nitrate plus nitrite and lower concentrations of ammonium and SRP. While relatively high, the concentrations of ammonium measured at the GR site were not great enough to be acutely toxic to stream biota.

The porewater results from our three sites can be compared to porewater nutrient chemistry reported by Hagerthey and Kerfoot (1998) below a lake located in the Trout Lake
watershed about 7 km away. Lake water had nutrient concentrations lower than porewaters regardless of groundwater flow direction; lake nutrient concentrations were also lower than measured in the streams during this study, which likely reflects the longer residence time of lake water. They also reported lower porewater SRP concentrations in weak discharge areas and areas where lake water recharges the groundwater, which may reflect the low concentrations of the lake-water source. Similar to the relation shown in Figure 2, they found lower ammonium porewater concentrations associated with high groundwater discharges. For comparison purposes, the GR site porewater concentrations of ammonium and SRP (Figure 2) are higher than any value reported by Hagerthey and Kerfoot (1998); the porewater values at the HGD and WGD sites are comparable to their reported concentrations.

Methods

Site Instrumentation Methods: At each of the three sites, hydrologic instrumentation collected stream stage and temperature, water table elevation in the streambank, groundwater elevation from below the streambed, shallow groundwater temperature below the streambed, and groundwater specific conductance from a monitoring well below the streambed. Biological sampling equipment included macroinvertebrate and periphyton samplers placed on the streambed. The specifics of the installations are given below.

Stream stages were measured using a shaft potentiometer and float-counterweight system (HGD site - Allequash Creek), RDS WL-40 capacitance rod (WGD site - North Creek), and a nitrogen bubble-pressure transducer system (Paroscientific PS-1 pressure transducer; GR site - Stevenson Creek). Stream stages were collected continuously over the study period by dataloggers; the interval of measurement was variable depending on time of year, and ranged between one and six hours. Stream temperatures were measured approximately 1 cm above the stream bottom using Hobo Tidbit dataloggers; the interval of measurement for temperature
ranged between 15 minutes and one hour during the study period. Shallow groundwater temperatures beneath the stream were collected from thermocouples located 0.15 m, 0.3 m, and 1 m below the streambed with an interval of measurement between one and three hours.

Streambank water-table wells consisting of 5.1-cm-diameter PVC pipe with a 1-m screened interval were installed using a hand auger and placed so that the screen intersected the water table. At each site, two piezometers were installed below the streambed consisting of 3.2-cm-diameter stainless steel pipe with 0.3-m stainless steel drive point screens. The piezometers were hand-driven to depth such that the bottom of the screen was placed 1 m below the top of the streambed sediments. Water-level measurements from the streambank and below the streambed were made at the sites using RDS WL-40 capacitance probes and Solinst Levelogger non-vented pressure transducers, respectively. One Solinst Barologger was placed at the GR site to record barometric pressure for compensating the non-vented pressure transducers. The datalogger recording interval ranged between two and six hours depending on time of year. In the second piezometer at each site, a Campbell Scientific temperature-specific conductance probe was installed near the screen midpoint. This instrument measured the non-purged/ambient specific conductance within the well over time. The recording interval for specific conductance measurements ranged between one and three hours.

The biota samplers used a passive approach that provided artificial substrate for colonization by macroinvertebrates, largely arthropods, mollusks, and annelids in these streams, and periphyton, a complex of epibenthic algae, fungi, bacteria, protists, and micrometazoans in a polysaccharide matrix that coats exposed streambed surfaces (Lock et al., 1984). Multiple plate-style samplers (Flannagan and Rosenberg, 1982) anchored to the streambed with bricks were used to sample macroinvertebrates in moderately flowing sites in the three study reaches (HGD, WGD, GR sites – Figure 1). Five samplers were employed in each reach for one-month colonization periods. Samplers were placed in different locations within each reach from month to month. The recharge and weak discharge sites were sampled monthly from early spring through mid-autumn (April 1 – November 4, 2001). The strong discharge site was sampled from late spring
through mid-autumn (May 29 – November 4, 2001). Invertebrate samples were transported to the lab in 95% ethanol and picked from the samplers under 10X magnification. Invertebrates were identified to family following keys in Merritt and Cummins (1996) and Smith (2001). Thus, taxonomic richness in this study refers to the number of families present in each sample.

Periphyton samplers consisted of 4.8 x 4.8 cm ceramic tiles affixed to concrete bricks with hook and loop fasteners and placed on the streambed near the invertebrate samplers. Seven to ten tiles per site per month were employed. After one-month colonization periods, tiles were cleaned of macroscopic invertebrates in the field before they were placed in 240 ml jars containing approximately 150 ml of stream water and transported to the lab on ice. Once in the lab, jars were sealed with respirometry lids and placed in dark environmental chambers set at 14°C. CO₂ efflux rate (uL CO₂ • hr⁻¹) was determined by drawing a 5 ml headspace gas sample through a self-sealing membrane on the respirometry lid, once approximately 24 hours after sealing and again approximately 24 hours later. CO₂ concentration was determined by injecting headspace gas samples into a CID 301 Infra-red Gas Analyzer. Five stream-water samples from each site were also analyzed each month as were ten deionized water controls. Efflux values from water samples were negligible in all cases and thus were excluded from the present study.

Statistical Methods: In addition to simple statistics (mean, median, standard deviation), time-series of water-level measurements were characterized using an approach similar to the root-zone residence-time analysis of Hunt et al. (1999). The difference between the water level in a piezometer and the water level of the stream was computed for each time step. Positive differences indicate groundwater discharge to the stream, and negative differences indicate groundwater recharge. In an effort to minimize small fluctuations in the water-level difference time series a 24-hour moving average was computed. The resulting smoothed time series was analyzed to determine the length of contiguous periods where strong groundwater discharge conditions (defined by a water-level difference exceeding 4 cm) existed. This resulted in a series
of durations corresponding to strong discharge conditions for each piezometer. Using the standard Weibull plotting position formula the cumulative probabilities for each strong-gradient duration were determined and used to fit a log-normal distribution to the duration series. The results were plotted using a cumulative probability scale. This presentation is similar to flow duration or grain size distribution curves, but in this case steep curves reflect frequent strong-gradient events at the duration given by the x-axis. The duration at a given probability refers to the probability that a value will be less than the duration specified on the x-axis.

Ranked data were analyzed for all biotic measurement due to deviations from normality and/or unequal variances. Macroinvertebrate abundance (individuals per sampler), macroinvertebrate richness (number of families per sampler) and periphyton respiration data (uL CO₂ per hour) for April through October samples (GR site vs WGD site) were analyzed with Mann-Whitney tests. Richness, abundance, and respiration data for June-October samples from all three sites were analyzed with Kruskal-Wallis tests. In all tests, the null hypothesis, i.e., samples came from the same population, was rejected when \( p \leq 0.05 \). Parametric alternatives to these procedures (two-sample t-tests and one-way ANOVAs) conducted on raw and transformed data produced identical results to the nonparametric tests.

**Results and Discussion**

*Hydrologic Differences Between Sites:* The sites had a range of vertical gradient between the groundwater system beneath the stream and the stream stage (Table 1), agreeing with the analytic element model results used to identify the site locations. The HGD site in Allequash Creek was the strongest discharge site as it had the largest gradient (Table 1); heads measured approximately 75 cm below the streambed residing on average, 27 cm above stream level. The GR site in Stevenson Creek, on the other hand, was a recharge site where heads beneath the stream were on
average 5 cm below the stream level. The WGD site in North Creek was a weaker discharge site (average head 1.8 cm above stream stage), and was intermediate to the HGD and GR sites. Although actual groundwater flow will be a function of the streambed sediment properties as well as hydraulic gradient, groundwater flow is difficult to accurately measure directly. Nevertheless, temperature-profile modeling (e.g., Lapham, 1989) at the HGD and GR sites demonstrated that groundwater flow magnitude is proportional to gradient (Spitzer-List, 2003).

In addition to vertical gradients which measure groundwater exchange through the stream bottom, water level data were also collected from an adjacent stream bank well (“bank well”) to assess the horizontal gradient between the stream and the groundwater system. This flow can be an important component because the sediments intervening along the side of a stream channel are commonly more transmissive than the sediments underneath the stream channel. While the gradient magnitude is not directly comparable because the distance to the stream edge is not constant at all sites, the variability of the water levels again is associated with groundwater flow direction (Table 1), with the GR site (the site with lowest water table relative to stream stage) having larger variability than either of the groundwater discharge sites.

In addition to the average gradients, aquatic systems may also be affected by the transience in the system. That is, how often does the system change from the average state? Two approaches are discussed below to characterize transience. The first is a simple look at the standard deviation of the measurements, and the number of gradient reversals that occurred during the study period. While generally useful, this approach does not provide information on the frequency and duration of the transient events. A second approach characterizes the distribution of periods of strong groundwater discharge to each site using techniques developed by Hunt et al. (1999).

The strength of the groundwater driving force is relatively constant over time, as evidenced by the lowest standard deviation in the groundwater system underneath the stream (2.3 cm) being measured at the HGD site (Table 1). Whereas the groundwater system is a large reservoir that provides consistent upward pressure, the GR site does not have this strong,
consistent flow direction and is more easily affected by changes in stream stage. This is shown by
the larger standard deviation (4.7 cm) in the streambed groundwater data (Table 1). The WGD
site is again intermediate between the two other sites (Table 1). The relation of the strength of
groundwater flow to the variability in the streambed heads was linear (Figure 3), reflecting the
range of stability resulting from the large reservoir of the groundwater system. The number of
vertical gradient reversals measured at the sites during 2001 was not proportional to groundwater
discharge (Table 1). At the HGD site there were no reversals, at the GR site there were 48
reversals, and the WGD site had the most reversals (238) reflecting its proximity to a zero vertical
gradient (Table 1, Figure 3) – a condition where small changes in water levels can cause reversals
in groundwater flow direction. A very similar ranking also is observed in the horizontal gradient
reversals (Table 1). These data suggest that the ability of the surface water to horizontally and
vertically infiltrate the adjacent aquifer is largest at the GR site, but the reversals of flow direction
were most often found in the site with the weakest gradient, in this case the WGD site.

The ranking of strong-gradient durations, defined here as continuous time that the 24-hour
moving average head difference was greater than 4 cm, were dependent on whether horizontal or
vertical flow was considered (Figure 4). For the bank wells (horizontal flow), the WGD site had
the shortest strong-gradient durations, followed by the GR and HGD sites. Although the median
durations (50% probability) for GR and WGD sites are similar, the WGD is more likely to have
short-duration periods than the GR site as evidenced by the divergence at the higher probability
values (Figure 4a). For the streambed piezometers, the ranking of strong-vertical gradient
duration is related to the amount of groundwater discharge (Figure 4b). Note that the streambed
piezometer distribution at the HGD site plots as a vertical line with a duration equal to the
monitoring period demonstrating that strong-gradient conditions existed for the entire monitoring
period; the streambed piezometer distribution at GR site plots as a vertical line with a duration
equal to 0 days showing strong-gradient conditions never existed for the entire monitoring period.
While the GR site is characterized by small horizontal gradients (Table 1), the transient response
of the site is characterized by longer duration periods of strong horizontal discharge than the WGD site (Figure 4a).

The strength of groundwater discharge also controls the degree of groundwater-surface water interaction, as well as the transient porewater chemistry response, as shown by the measurements taken during a June 19, 2001 storm event (Figure 5). At the WGD site the storm has little effect on the porewater temperature or specific conductance. While porewater temperature data were not available at the HGD site for this storm, the specific conductance time series also shows little effect from the storm. At the GR site, however, warm/dilute surface water derived from the storm event to penetrate into the hyporheic zone, as shown by deflections in specific conductance in the streambed well and the porewater temperature profile. Interestingly, a peat layer present below the streambed appears to bifurcate the lateral flow of infiltrating surface water as shown by the temperature response at 15 and 91 cm depth, and less so at the 30 cm depth. The temperature and specific conductance time series are consistent with the hydrographs measured during this time period in that there is a downward gradient at the recharge site and upward gradients measured at the groundwater discharge sites during the storm (Figure 5).

In addition to controlling the exchange between the hyporheic zone and the stream, groundwater discharge also affects the longer-term thermal environment of the stream system, as shown by the temperate ranges for the stream (Table 2, Figure 6) and three depths (Table 2) below the streambed (15 cm, 30 cm, and 91 cm). Because data were not collected at the HGD site streambed during 2001, 2002 data are included for comparative purposes. The HGD and WGD sites are associated with colder temperatures (Table 2, Figure 6) and less temperature variability in the stream and the streambed beneath the stream (Table 2). The GR site is warmer at all depths with the exception of 91 cm where it is slightly cooler than the WGD site. We attribute this to the presence of a peat bed between the 30 cm and 91 cm depths at the GR site (USGS, unpublished field notes). Peat sediments have relatively high insulation capacity (Hunt et al., 1996); thus, heat transport from the surface through the sediment matrix is likely reduced.
**Biological Differences between Sites:** A total of 23,090 macroinvertebrates representing 50 taxonomic families were collected from 90 samplers (mean = 4.74 samplers per site per month). A total of 141 periphyton respiration samples were collected and analyzed (mean = 7.42 per site per month). Noticeable differences were noted for all measures (Table 3), although the degree of significance depended on whether the entire April-October or shortened June-October data set was used. For the entire April through October period collected at the WGD and GR sites, abundance (246 vs 180 individuals per sampler, \( U = 837 \) vs 285, \( p = 0.0005 \)), richness (11 vs 9 families per sampler, \( U = 830 \) vs 292, \( p = 0.0000 \)), and respiration (11 vs 7 \( \mu L \) CO\(_2\) per hour; \( U = 1591 \) vs 609; \( p = 0.0002 \)) means were larger at the WGD site than at the GR site. For the June through October period when data were collected concurrently at all three sites, abundance was much greater at the HGD site than in the WGD and GR sites (means = 388, 252, and 184 individuals per sampler respectively, \( H = 31.847, p = 0.0000 \)). Abundance at the WGD site was not significantly different from GR site abundance when the shorter June-October data were used at the critical rejection level of \( p \leq 0.050 \). However, at \( p \leq 0.055 \), abundance at the WGD site could be determined to be significantly greater than at the GR site, which would be consistent with the strong pattern produced when April and May data are included. Richness was much greater at the HGD site than at the WGD and GR sites during the June – October period (means = 13, 10, and 9 families per sampler respectively, \( H = 830, p = 0.0000 \)). Respiration for the June through October period was much greater at the HGD and WGD sites than at the GR site (12, 12, vs 7 \( \mu L \) CO\(_2\) per hour respectively, \( H = 19.142, p = 0.0001 \)).

These results were consistent with a hypothesis relating biological activity to variation in groundwater-surface water exchange. However, the causal mechanisms underlying the observed pattern are unclear at present and a topic of on-going research. Because periphyton responses generally tracked macroinvertebrate community characteristics (Table 3, Figure 7), it is conceivable that inter-site differences in animal community composition and abundance were driven by localized increases in epibenthic productivity. If true, environmental properties resulting from quantitative variation in groundwater-surface water exchange may exert strong effects on
stream life. While our data show that macroinvertebrate and periphyton quantity and quality correspond to the degree of groundwater discharge, it is well know that higher trophic levels such as brook trout would also respond to local groundwater-surface water conditions. For example, monthly degree-day temperature conditions (Figure 6) show that average streamwater temperatures increase with decreasing groundwater discharge. Thermal stress for brook trout occurs when temperatures exceed 20°C (Becker 1983). Average daily temperatures occasionally exceeded this mark in all three sites. However, this effect was greater at the GR site (no. days = 86) than at the WGD site (69 days) and HGD site (34 days). Temperatures greater than 25°C are lethal to brook trout (Becker 1983), a condition that did not occur at the HGD site, occurred 6 times at the WGD site, and 20 times at the GR site. Brook trout seek out upwelling sites during these stressful periods, which allows them to maintain body temperatures that are as much as 4°C cooler than stream water averages (Baird and Krueger 2003). This suggests that local availability of upwelling zones could exert control on top predator abundance in these communities.

**Implications for Watershed Water Resources Management:** Understanding the association of wetland stream benthic productivity and groundwater-surface water interaction will likely be critical as water managers move toward a more integrated approach to the management of watershed ecology. In the work presented here, not all wetland streams provided the same habitat (and associated biodiversity and richness) even though the wetland plant communities were similar among the sites. Rather, the site’s relation to the harder-to-characterize groundwater system appeared to be an important driver. Moreover, this work demonstrates the importance of understanding the transient nature of local groundwater-surface water interaction in order to accurately characterize the drivers that control the ecological system. While this would be important for assessing wetland-specific objectives such as the loss of one wetland from the basin compared to another (and related cumulative adverse effects), it also has ramifications for scaling up results from the field scale to the watershed scale. That is, numerical groundwater flow modeling has demonstrated that the Trout Lake watershed is dominated by groundwater
discharge conditions. Clearly, extrapolating results from the Stevenson Creek groundwater recharge site would not be representative of the watershed as a whole. The Allequash Creek HGD site, however, is located in the area of largest hydraulic gradient in the basin. Thus, it too is likely a poor representative for the “watershed average” condition. Although it is not known if the intermediate WGD site at North Creek is the site that represents average conditions for the Trout Lake watershed, knowledge of the groundwater flow system, and the range of possible groundwater-surface water interactions, can narrow the uncertainty and give bounds for what can be expected for the system as a whole. Although it is encouraging that the groundwater system can be simulated numerically (giving rise to quantitative predictions throughout the watershed), more information regarding the controlling processes that link groundwater-surface water interaction to the biotic system is needed to attain true integrated watershed management. This “ecohydrologic” understanding of the system will likely become of primary importance as society demands more quantitative and spatially distributed answers from watershed managers.

Conclusions

As part of research conducted at one of the U.S. Geological Survey’s Water, Energy and Biogeochemical Budgets (WEBB) sites, work was undertaken to determine if variation in groundwater-surface water exchange can be used to explain observed patterns in benthic productivity. The major findings of this work were:

- Groundwater-surface water measurements were taken at three wetland stream sites that included high groundwater discharge (HGD site on Allequash Creek), weak groundwater discharge (WGD site on North Creek), and groundwater recharge (GR site on Stevenson Creek). High groundwater discharge was associated with the lowest variability in groundwater head below the stream; the groundwater recharge site had the highest
variability in head beneath the stream. Horizontal and vertical gradient reversals were fewest at the HGD site, but largest at the WGD site (the site closest to zero gradient).

- The duration of strong discharge events (head difference > 4 cm) differed between sites as well. The HGD site had the longest duration of strong discharge events, both for horizontal and vertical flow. The GR site had no strong discharge events for vertical flow but was intermediate for horizontal flow events. While the GR site is characterized by the smallest average horizontal gradients, the transient response of the GR site is characterized by longer duration periods of strong horizontal discharge than the WGD site.

- Groundwater discharge creates distinct temperature/chemical environs where the stability in the system is related to the amount of groundwater discharge. The hyporheic conditions are cooler and are not greatly affected by changes in the overlying surface water. At the GR site, however, changes in surface water temperature and specific conductance were transported to the hyporheic zone very quickly, presumably by infiltrating surface water.

- Macroinvertebrate abundance, taxonomic richness, and periphyton respiration data from these stream reaches fit the predicted gradient of relatively high values at the HGD site and relatively low values at the GR site.

- Not all wetlands provided the same habitat even though the wetland plant communities were similar among the sites. Rather, the site’s relation to the harder-to-characterize groundwater system appeared to be an important driver. This work also demonstrates the importance of understanding the transient nature of local groundwater-surface water interactions in order to accurately characterize the drivers that control the ecological system. While important for assessing local wetland functions, knowledge of the range of groundwater-surface water interactions operating in a watershed could also be critical for
applying site-specific results to watershed-scales most often used in integrated water resources management.

These results suggest that groundwater-surface interaction can have appreciable transient characteristics and can exert significant effects on the biology of a local stream-wetland system. Moreover, this interaction is expected to vary on the watershed scale. Assessing and managing groundwater-surface water interaction within a watershed may be required to accurately assess and manage the ecological habitat in a watershed context.

Acknowledgements

The manuscript was improved by review from Emily Stanley, Donald Rosenberry, Charles Dunning, Tom Winter, and one anonymous reviewer. Matthew P. Ayres provided helpful discussions early in the project, helped develop the respiration methods, and lent the equipment used for making early CO₂ measurements. Joseph Scanlan is acknowledged for assistance in the field and laboratory, Steve Connolly for assistance in the laboratory. This work was funded by the U.S. Geological Survey’s Northern Temperate Lakes Water, Energy, and Biogeochemical Budgets (NTL-WEBB) program.

Authors’ note: Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government

References


Figure Captions

Figure 1. Map showing location of the study area in Wisconsin. Wetland-stream sites HGD, WGD, and GR refer to high groundwater discharge, weak groundwater discharge, and groundwater recharge conditions, respectively.

Figure 2. Results of stream and porewater nutrient concentrations June 2001 (UW-Madison Center for Limnology, unpublished data). SW refers to “surface water”; GW refers to “groundwater”.

Figure 3. Average vertical gradient between the stream and the groundwater beneath the stream is plotted against the standard deviation in the groundwater head beneath the stream.

Figure 4. Cumulative probability of the duration of strong-discharge events for horizontal (Fig. 4a) and vertical (Fig. 4b) groundwater flow.

Figure 5. Hydrographs (top row), temperature profiles (middle row), and streambed well specific conductance (bottom row) for a transient precipitation event occurring on June 18 through 19, 2001.

Figure 6. Cumulative degree days (base = 0 degrees C) for each site for May through October 2001.

Figure 7. Abundance and richness (macroinvertebrate samples) and respiration (periphyton samples) for June-October 2001 sampling.
EXPLANATION

- Surface-water basin boundary
- Surface-water sub-basin boundary
- Wetland-stream site
- LTER study lakes
- Other lakes


Hunt et al. Figure 1
Hunt et al. Figure 2

Error bars = \pm \left( \frac{\text{std dev}}{\sqrt{n}} \right)

- LOG CONCENTRATION (\mu g/L)

- NO\textsubscript{3} + NO\textsubscript{2} - N
- NH\textsubscript{4} - N
- SRP

- HGD site sw
- WGD site sw
- GR site sw
- HGD site gw
- WGD site gw
- GR site gw

n=2
Hunt et al. Figure 4
Hunt et al. Figure 6

Total degree days accumulated
- Allequash HGD site = 2,897
- North WGD site = 3,262
- Stevenson GR site = 3,461
Hunt et al. Figure 7
Table 1 - Summary of 2001 Hydrologic Data.
HGD = high groundwater discharge; WGD = weak groundwater discharge; GR = groundwater recharge
Positive values refer to flow toward the stream, negative values are flow away from the stream

<table>
<thead>
<tr>
<th></th>
<th>HGD Site</th>
<th>WGD Site</th>
<th>GR Site</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Streambed well - Stream data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean vertical gradient</td>
<td>0.35</td>
<td>0.02</td>
<td>-0.06</td>
</tr>
<tr>
<td>median vertical gradient</td>
<td>0.36</td>
<td>0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td>std deviation vertical gradient</td>
<td>0.04</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>std deviation stream bed gw level (cm)</td>
<td>2.3</td>
<td>4.2</td>
<td>4.7</td>
</tr>
<tr>
<td>std deviation stream stage (cm)</td>
<td>4.2</td>
<td>4.3</td>
<td>3.2</td>
</tr>
<tr>
<td>number of gradient reversals</td>
<td>0</td>
<td>238</td>
<td>48</td>
</tr>
<tr>
<td>number of measurements</td>
<td>2511</td>
<td>2434</td>
<td>3854</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>HGD Site</th>
<th>WGD Site</th>
<th>GR Site</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bank well - Stream data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean horizontal gradient</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>median horizontal gradient</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>std deviation horizontal gradient</td>
<td>0.01</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td>std deviation stream bank gw level (cm)</td>
<td>3.3</td>
<td>4.9</td>
<td>6.1</td>
</tr>
<tr>
<td>std deviation stream stage (cm)</td>
<td>4.5</td>
<td>4.3</td>
<td>3.8</td>
</tr>
<tr>
<td>number of gradient reversals</td>
<td>8</td>
<td>248</td>
<td>124</td>
</tr>
<tr>
<td>number of measurements</td>
<td>1270</td>
<td>2434</td>
<td>2260</td>
</tr>
</tbody>
</table>

site classification: high gw dischchg weak gw dischchg gw recharge
Table 2 - Summary of 2001 and 2002 Temperature Data
HGD = high groundwater discharge; WGD = weak groundwater discharge; GR = groundwater recharge

<table>
<thead>
<tr>
<th>Stream</th>
<th>May through December 2001</th>
<th>April through November 2002</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HGD Site</td>
<td>WGD Site</td>
</tr>
<tr>
<td>mean (degrees C)</td>
<td>14.4</td>
<td>16.3</td>
</tr>
<tr>
<td>median (degrees C)</td>
<td>14.8</td>
<td>17.5</td>
</tr>
<tr>
<td>std deviation (degrees C)</td>
<td>6.1</td>
<td>5.9</td>
</tr>
</tbody>
</table>

15 cm depth below stream
mean (degrees C) | NM | 13.9 | 14.4 | 10.3 | 11.8 | 12.6 |
median (degrees C) | NM | 14.5 | 15.6 | 10.7 | 12.8 | 13.4 |
std deviation (degrees C) | NM | 4.5 | 5.6 | 4.0 | 5.5 | 6.5 |

30 cm depth below stream
mean (degrees C) | NM | 13.6 | 13.6 | 9.1 | 11.4 | 11.7 |
median (degrees C) | NM | 15.2 | 14.6 | 9.3 | 12.3 | 12.4 |
std deviation (degrees C) | NM | 4.1 | 4.3 | 2.7 | 5.1 | 5.4 |

91 cm depth below stream
mean (degrees C) | NM | 12.6 | 11.3 | 7.5 | 10.5 | 9.7 |
median (degrees C) | NM | 13.4 | 11.5 | 7.7 | 11.2 | 9.9 |
std deviation (degrees C) | NM | 2.9 | 2.2 | 0.9 | 4.2 | 3.3 |

minimum number of measurements distributed over time period
5136 | 4323 | 4927 | 1877 | 1878 | 5630 |

site classification
high gw dischg weak gw dischg gw recharge high gw dischg weak gw dischg gw recharge
NM = not measured
Table 3. Means (± 1 standard error) for macroinvertebrate abundance (individuals per sampler), macroinvertebrate richness (mean families per sampler), and periphyton respiration (μL CO₂ • hr⁻¹). Asterisks presented to the right of the larger mean in each pair-wise comparison indicate significance levels for each comparison (from Mann-Whitney U tests in recharge vs weak discharge and Kruskal-Wallis H tests in full gradient analyses), *** indicates p ≤ 0.001 * p ≤ 0.10 (= 0.055 in only case), n.s. indicates non-significant differences. June-October 2001 data are plotted in Figure 7 for illustrative purposes.

<table>
<thead>
<tr>
<th></th>
<th>HGD Site (strong discharge)</th>
<th>WGD Site (weak discharge)</th>
<th>GR Site (recharge)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abundance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>April-October</td>
<td>246 ± 15 ***</td>
<td>180 ± 30</td>
<td></td>
</tr>
<tr>
<td>June - October</td>
<td>388 ± 34 ***</td>
<td>252 ± 19 *</td>
<td>183 ± 37</td>
</tr>
<tr>
<td><strong>Richness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>April-October</td>
<td>11.2 ± 0.4 ***</td>
<td>9.0 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>June - October</td>
<td>12.9 ± 0.4 ***</td>
<td>10.2 ± 0.4 (n.s.)</td>
<td>8.8 ± 0.3</td>
</tr>
<tr>
<td><strong>Periphyton</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>April-October</td>
<td>11.05 ± 1.00 ***</td>
<td>6.80 ± 0.56</td>
<td></td>
</tr>
<tr>
<td>June - October</td>
<td>12.10 ± 0.86 (n.s.)</td>
<td>11.82 ± 1.44 ***</td>
<td>6.75 ± 0.73</td>
</tr>
</tbody>
</table>