Spatio-temporal variations of hyporheic flow in a riffle-step-pool sequence

Daniel H. Käser,1* Andrew Binley,1 A. Louise Heathwaite,1 and Stefan Krause2

1 Lancaster Environment Centre, Lancaster University, Lancaster, UK
2 Earth Science and Geography Department, Keele University, Keele, UK

Abstract

Subsurface flow in streambeds can vary at different scales in time and space. Recognizing this variability is critical for understanding biogeochemical and ecological processes associated with the hyporheic zone. The aim of this study was to examine the variability of hydraulic conductivity (K), vertical hydraulic gradients (VHGs), and subsurface fluxes, over a riffle–step–pool sequence and at a high spatio-temporal resolution. A 20 m reach was equipped with a network of piezometers in order to determine the distribution of VHGs and K. During a summer month, temporal variations of VHGs were regularly surveyed and, for a subset of piezometers, the water level was automatically recorded at 15 min intervals by logging pressure transducers. Additionally, point-dilution tests were carried out on the same subset of piezometers. Whereas the distribution of vertical fluxes can be derived from K and VHG values, point-dilution tests allow for the estimation of horizontal fluxes where no VHG is detectable. Results indicate that, spatially, VHGs switched from upwelling to downwelling across lateral as well as longitudinal sections of the channel. Vertical fluxes appeared spatially more homogeneous than VHGs, suggesting that the latter can be a poor indicator of the intensity of flow. Finally, during flow events, some VHGs showed little or no fluctuations; this was interpreted as the result of a pressure wave propagating from upstream through highly diffusive alluvial sediments.

KEY WORDS
hyporheic exchange flow; heterogeneity; hydraulic conductivity; vertical hydraulic gradient; groundwater–surface water interactions

INTRODUCTION

Hyporheic flow variability in previous work

In this work, we define hyporheic flow as the subsurface flow beneath and adjacent to streams, whatever the origin of water. We consider hyporheic exchange flow (HEF) as a specific process by which stream water infiltrates the subsurface and returns to the stream over relatively small distances (Harvey et al., 1996). HEFs may exhibit some variability that affects the ecological functions of groundwater and lotic ecosystems. Yet, to date, hydrological variability in time, space and across scales is not well understood. In particular, different mechanisms of groundwater–surface water (GW/SW) mixing make research challenging.

A primary distinction of such mechanisms can be made between stream water infiltration caused by sedimentary transport and by flow exchanges. The former, turnover exchange (Elliott, 1990; Elliott and Brooks, 1997), results from the trapping and release of pore water, and is the response to the stream flow eroding and depositing sand bed particles as bedforms move (Packman and Bencala, 2000). The transfer of water as a result of actual flow processes can be classified into four categories according to the control: (1) turbulent diffusion is caused by the transfer of turbulent momentum between stream and pore-water flow (Zhou and Mendoza, 1993; Packman and Bencala, 2000); (2) hydrodynamically induced advection, known as bedform-/flow-induced advection, current–obstacle interaction or pumping exchange, is caused by the acceleration of flow over bedforms that gives rise to pressure variations, thus inducing flow in and out of the bed (Thibodeaux and Boyle, 1987; Elliott, 1990); (3) hydrostatically induced advection results from spatial variations of the hydraulic gradient caused either by geomorphological features such as stream meanders (Wroblicky et al., 1998; Boano et al., 2006) and in-stream structures, e.g. debris dams, step-pool sequences (Goossef et al., 2006; Lautz et al., 2006; Hester and Boyle, 2008), or hydrogeological characteristics, primarily permeability distribution (Woessner, 2000; Cardenas et al., 2004) and ambient groundwater discharge (Larkin and Sharp, 1992; Winter, 1999); and (4) what may be considered as transient exchange is the transfer driven by the fluctuations of stage and groundwater, for example through bank storage (Sauer and Pinder, 1970; Konrad, 2006).

Research has identified a wide range of time and spatial scales at which HEFs occur (Boulton et al., 1998; Hancock et al., 2005), and modelling studies in particular have emphasized their connection to a hierarchy of flow systems up to the basin scale (Wörman et al., 2007; Cardenas, 2008). Through a tracer experiment in
a mountain stream, Haggerty et al. (2002) found the residence time distribution of HEF to be a power-law covering 1.5 orders of magnitude (1.5 h to 3.5 days). It has been pointed out that this fractal scaling of flow paths can be explained by the geomorphology of the surface water–groundwater interface alone, without having to account for geological heterogeneity (Cardenas, 2008). Nevertheless, such heterogeneity also controls exchange patterns, both at the bedform scale (Salehin et al., 2004) and the reach scale (Wroblicky et al., 1998). In streambeds, permeability can be expected to exhibit strong 3D variability over a few metres (Cardenas and Zlotnik, 2003; Conant et al., 2004), and can be altered by erosion/deposition processes occurring over time scales of years (Ryan and Puckman, 2006), seasons (Generreux et al., 2007) or days (Wondzell and Swanson, 1999).

Since the formulation of the patch dynamics concept (Townsend, 1989), lotic ecologists and biogeochemists have developed an increased interest in these dynamics and the heterogeneity of hyporheic flows (Hendricks, 1993; Stanley et al., 1997; Lake, 2000; Brunke et al., 2003). Whereas patchiness makes hyporheic research difficult (Palmer, 1993), ecologists mostly identify such complexity as an enhancer of biogeochemical processes, e.g. nutrient recycling (Grimm et al., 2005) and biodiversity (White, 1990). Therefore it is likely that our understanding of hyporheic flow processes will mature by addressing, rather than bypassing, these issues of dynamics and heterogeneity.

**Purpose of this study**

Previous studies have not exploited the continuous logging of hydraulic heads to study simultaneously spatial and temporal variations of hyporheic flow over longitudinal and cross sections of a riffle–pool sequence. Additionally, although temporal variability has been addressed by inter-seasonal comparisons, little is known about the dynamics of hyporheic flow over shorter periods. The aim of this research was to gain a better understanding of such variability through the case study of a riffle–step–pool sequence and its associated HEF system. Specific objectives were the characterization of: (1) the spatial distribution of streambed permeability; and (2) the spatial and temporal variability of vertical hydraulic gradients (VHGs). Furthermore the aim was to compare the use of VHGs, VHG-derived fluxes and dilution test-derived fluxes to characterize flow patterns. Although this work investigates the controls on HEF patterns, it does not seek to confirm the actual infiltration of stream water through artificial or environmental tracers.

**SITE AND METHODS**

**Field site**

This study was conducted at the River Leith, within the Eden catchment, in the northwest of England (Figure 1). The Leith’s catchment covers an area of 54 km² and elevation ranges between 105 and 370 m above sea level. Annual precipitation, mainly rain, is approximately 900 mm near the outlet of the Leith, and 1400 mm at its source. In summer, discharge in the vicinity of the outlet can be as low as 0.03 m³s⁻¹. From 2004 to 2007, the mean daily discharge in summer (1 June–1 September) ranged from 0.1 to 2.1 m³s⁻¹, and during the period concerned by this study, from 27 June to 5 August 2008, it fluctuated between 0.08 and 0.7 m³s⁻¹ (Figure 2).

The field site is located in the lower part of the catchment, and belongs to a 3 km stretch that is known to be significantly groundwater-fed. Stream flow measurements carried out on 12 different dates over this longer reach indicate that groundwater discharge causes an average of 70% increase of stream flow when flow is lower than 0.15 m³ s⁻¹ (Seymour et al., 2008). The underlying bedrock, the aeolian *Penrith Sandstone*, is a major aquifer and is part of the Permo-Triassic Sandstone (Allen et al., 1997). The lithological log of a borehole, located 150 m downstream of the field site, indicates that the sandstone bedrock extends at least 50 m deep beneath the channel, and that it is overlaid by 1 to 2 m of soft sediments. Five
auger profiles in the floodplain show that the depositional pattern of this alluvium consists mainly of silty soil above a mix of gravels, sands and silts (augering locations in Figure 1).

The investigated reach is 20 m long. Morphologically it represents the middle section of a river (Owen et al., 2005); at this location, the Leith is neither a headwater stream nor a lowland river, but it meanders within a narrow floodplain (<100 m), and its bed is characterized by riffle–pool sequences and the predominance of gravels and cobbles. Although the regional surface water slope is 0.5%, the study reach has a higher slope (1%), as it covers a riffle–step–pool sequence (Figure 1). During low flows, the step’s slope on its own is higher than 2.4%. Note that this journal issue contains the study of the Leith as it covers a riffle–step–pool sequence (Figure 1).

Methods

**Piezometers.** Channel hydraulic heads were obtained from 28 piezometers that were arranged in pairs or triplets in order to cover different depths per location (Figure 1), with depths ranging from 23 to 73 cm below the sediment–water interface. An additional piezometer was installed on each bank and one at the outer edge of the floodplain (50–80 cm and 25 cm deep below stream stage at low flows, respectively). Piezometers were constructed from 32 mm (inner) diameter PVC pipes. Their bottom end was open, perforated over 12 cm by approximately 60 × 7 mm diameter holes, and covered by a protective nylon mesh. On site they were installed in holes (approximately 5 cm diameter) dug by a petrol-powered auger; as the auger was retrieved, loose sediments collapsed immediately. The first measurements were carried out several days after installation in order to ensure a total recovery of hydraulic heads.

**Hydraulic conductivity, vertical hydraulic gradients and hydraulic heads.** Saturated hydraulic conductivity was estimated through slug tests, analysed using the Hvorslev method (Schwartz and Zhang, 2003). Most slug tests were replicated two to three times with different slug lengths (5, 10 and 20 cm) and hydraulic conductivity was then calculated using the arithmetic mean. As we found no evidence of skin effects induced by the nylon mesh, these were assumed not to affect the analyses.

Water levels were measured using a graduated electrical contact meter with a precision of 0.3 cm. Over the period of study (27 June to 3 August 2008), we carried out 11 surveys, during which subsurface heads and the stream stage were measured (Figure 2a). We refer to this dataset as discrete measurements, as opposed to the continuous ones described hereafter.

The in-stream VHG (%) was calculated as 100 · Δh/Δl, where Δh is the elevation difference between water levels of the stream and the piezometer, and Δl is the distance between the mid-screen depth and the stream–sediment interface (Kalbus et al., 2006). Because Δh is calculated as the subsurface head elevation minus the stage elevation, positive values reflect the potential for upwelling and negative values for downwelling. In order to map the head distribution over the entire reach, the location and elevation of piezometers were surveyed using a total station.

Sixteen pressure transducers were used to log water levels at a time step of 15 min. Twelve of them were spatially spread out in channel piezometers, three were installed in the bank, and one was located in the downstream pool, protected by a perforated steel tube, for recording stream stage. Although a second stage logger was placed in the riffle, a technical failure made it impossible to access the data. An additional logger was kept out of water to allow for barometric compensation. In this paper we refer to these logged water levels as the continuous measurements.

**Determination of flux.** The flux, i.e. the specific discharge, through the streambed was characterized through two approaches. First, vertical fluxes were calculated using Darcy’s law as \( q_v = K \cdot \frac{\Delta h}{\Delta l} \), where \( K \) is the hydraulic conductivity derived from the slug tests. Second, fluxes were derived by point-dilution (or borehole dilution) tests. This technique can provide an estimate of the horizontal average flux near the well screen (Freeze and Cherry, 1979). The method consists of injecting a tracer at screen depth and logging the exponential decrease of concentration caused by the through-flow. In order to analytically determine the flux, three assumptions must be made about the interstitial flow: (1) water...
upstream of the borehole is at uniform background (tracer) concentration; (2) there is no vertical flow in the aquifer; and (3) the flow is steady. Additionally, four stipulations must be implemented in the experimental design: (1) the concentration within the borehole remains uniform and equal to the concentration leaving the borehole; (2) the concentration at time zero is instantaneously raised to $C_0$; (3) the tracer does not increase the density of the water; and (4) the head is not affected by the injection. From these assumptions (Hazell, 1998), the horizontal flux can be derived as:

$$q_h = -\frac{\pi r}{2\Delta t} \ln \left( \frac{C_t}{C_0} \right)$$

Where $t$ is the time; $C_0$ is the peak concentration following the injection minus the background concentration; $C_t$ is the concentration of the tracer at time $t$ minus the background concentration; $r$ is the radius of the piezometer; and $\alpha$ is an adjustment factor that depends on the hydraulic characteristics of the aquifer, the gravel or sand pack and the well screen. The slope $\ln(C_t/C_0)/t$ can be obtained by a simple linear regression.

Dilution tests were performed at the River Leith in 17 piezometers, on 27 June and 31 July, with a conductive tracer (KCl) that was selected for its high electrical conductivity to density ratio. The injection was performed instantaneously with a syringe containing a 3 ml solution at 220 ms cm$^{-1}$, equivalent to 0.35 g KCl, in order to limit head elevation (lower than 0.4 cm) and density effects. The injection tube was carefully flushed with air immediately after injection to ensure uniform mixing in the 96 ml volume screened by the piezometer, but was not removed from the well in order to prevent vertical mixing above the screened section. The experimental design (Figure 2b) was first tested in the laboratory by injecting a coloured tracer in a transparent tube. In the field, self-made four-electrode probes were connected to a data logger (Campbell CR10X) that recorded electrical conductance at 60 s intervals, prior to and after the injection.

Data analysis did not require any calibration of the probes with solutions of known electrical conductivity, as we assume a linear relation between conductivity and KCl concentration and no change in the geometric factor that relates conductance to electrical conductivity. For calculating the flux with the dilution-test equation, the coefficient $\alpha$ was set to 2, which is a good approximation where there is no gravel pack (Hazell, 1998). Although a key assumption is the dominance of lateral flow, this method was applied to piezometers with VHGs equal to zero as well as to those with positive or negative VHGs. In the second case, the calculated flux can only serve as an indicator of potential flux.

RESULTS

Spatial variability

Hydraulic conductivity. Slug test results were carefully examined and several values were excluded from the hydraulic conductivity ($K$) dataset. Out of the 31 piezometers, I was totally unresponsive and discarded from this study, and 12 of them had their test interrupted because the response was obviously too low ($K < 10^{-7}$ m s$^{-1}$) for the sandy to gravelly media that we observed when augering (Fetter, 2001). Some of these tubes, when removed, showed deposits of fine sediments over the screen length, caused by the rupture of the nylon mesh during installation. Therefore only 17 piezometers (including the two bank ones), which were found to be clear of internal deposits, were retained.

Where piezometers were responsive, results of the multiple slug tests showed that the coefficient of variation of $K$, for each piezometer, ranged from 9% to 72% with an average of 35%. The two deepest channel piezometers (65 and 55 cm) had lower $K$ values ($0.7 \times 10^{-5}$ and $1.9 \times 10^{-5}$ m s$^{-1}$) because they reached the sandstone bedrock. These were excluded from the subsequent flux calculations as two values were not judged sufficient to enable comparisons with the rest of the network. For all other points, $K$ was measured at depths of 15 to 40 cm. Values span over an order of magnitude between $2 \times 10^{-5}$ and $12 \times 10^{-5}$ m s$^{-1}$, which corresponds to the well-sorted gravels and sandy gravels observed in sediment core samples from the site. Because several piezometers were discarded, most locations are characterized by measurements at a single depth. For the two locations with two operational piezometers, we used the geometric mean of $K$.

Upstream of the step, a relatively low permeability area appears on the right-hand side of the channel, whereas downstream of the step, the low permeability area is on the left-hand side (Figure 3a). On average, permeability is lower in the pool. On the left bank, $K$ is similar to the channel values ($2.5 \times 10^{-5}$ m s$^{-1}$), whereas on the right bank it is an order of magnitude lower ($0.1 \times 10^{-5}$ m s$^{-1}$). Although these values only represent a single location per bank, the latter agrees with the lithological profile of the eroded bank that consists of silty material, and the former agrees with the core profile that consists, at screen depth and above, of sandy gravel.

Vertical hydraulic gradients. Here we present the spatial variability of VHGs during lower stream flows, relative to the entire month of survey. Temporal variations of VHGs are presented in the Section on the Discrete Time Series of Hydraulic Heads and VHGs. Those piezometers characterized by a permeability that was low for the range anticipated, given the sediment characteristics, were included in VHG analysis provided they were responsive. At these locations, changes of VHG were evaluated for long-term variability only ($\pm$1 day).

During typical baseflow conditions, gradients were mostly positive, and ranged from $-1$ to 18% (e.g. 27 June depicted in Figure 3b). Here ‘typical’ is relative to the period of study, and includes stage variations up to 15 cm above the lowest observed stage (Figure 2a). Negative VHGs indicate that the subsurface head is below stream level, and that potential flow is downwelling.
Inversely, positive VHGs imply that piezometric heads are higher than stage and subsurface water discharge possibly occurs. On 27 June, the maximum value (18%) corresponds to a 5 cm difference between the piezometric level and stage. One nest, consisting of two piezometers that almost repeatedly indicated ‘no vertical flow’ and downwelling (−2%), was located directly upstream of the step (Figure 3b).

The spatial distribution of VHGs has been found to (1) mirror, to some extent, the distribution of \( K \) in an inverse manner (Figures 3a,b, 4); (2) exhibit downwelling and no vertical flow conditions upstream of the step, probably as a result of a step-induced HEF system; and (3) show lateral variations of VHGs, both in magnitude (over the whole reach), and direction (upstream of the step).

**VHG-derived flux.** The distribution of vertical fluxes, calculated with Darcy’s law, is presented for the representative baseflow conditions defined in the previous paragraph (27 June), whereas time series are presented in the Section on Continuous Time Series of Hydraulic Heads. Unlike \( K \) and VHGs, values of vertical flux cover a range less than an order of magnitude: from 1 to \( 6 \times 10^{-6} \text{ m s}^{-1} \) (Figure 3c). Consequently, fluxes appear more homogeneous over most of the reach. The centre of the pool (\( 10^{-6} \text{ m s}^{-1} \)) and the right side of the riffle (0 to \( 10^{-6} \text{ m s}^{-1} \)), however, are characterized by lower values. Notice that no flux was calculated at the downwelling location next to the step, as it responded poorly to the slug test and was therefore discarded from the permeability dataset.

**Point-dilution-derived flux.** Half of the injections were performed simultaneously on 27 June, and the other half simultaneously on 31 July, in similar stream flow and antecedent conditions (Figure 2a). The difference in stage elevation between both dates was 3 cm, and stage variations were less than 2 cm over the 36 h preceding the injections; therefore, fluxes derived from both experiments are assumed to be comparable. In order to obtain a reading of the natural background, electrical conductance was logged during several hours prior to the injection, and the logging was interrupted approximately 3 days after the injection. For the first test, background conductance was set equal to the final plateau value, which, in all cases, remained constant for at least 36 h. For the second test, background conductance was averaged from the 20 h of logged conductance prior to the injection because a sharp rise in stage occurred before readings were stable; the analysis was restricted to the 12 h prior to the event.

Dilution responses were smooth and enabled a good to very good estimation of the slope \( ln(C_i/C_0)/t \) (mean \( R^2 = 0.98 \)). The slightly imperfect fit, which occurs in most tests (Figure 5), is possibly a result of changing hydrologic conditions during the logging period, presumably associated to the receding stage.

Fluxes derived from point-dilution tests are generally lower than VHG-derived fluxes (Figure 3c vs 3d), and a distinction is apparent between the upstream riffle, with lower fluxes, and the downstream pool. However, these values must only be considered as indicators as, in most cases, the assumption of ‘absence of vertical flow’ is invalid.
Temporal variability

Discrete time series of hydraulic heads and VHGs. Among the 11 survey dates, 8 are located in the tail of a falling limb, and 4 are representative of high flow events during the survey (Figure 2a). Depending on the location, discrete measurements captured stream stage variations of 13 to 20 cm. Subsurface heads span over a slightly narrower range, from 14 to 19 cm. The highest observed VHG was 27% and the lowest, −6%. Overall, these values remained relatively constant, varying on average of ±7% and at some locations no more than ±2%.

Despite their apparent stability, however, the discrete measurements of VHG show a coherent response to stream stage when comparing the whole range of flow conditions. Through a comparison of riffle and pool piezometers, two main observations can be made: (1) all subsurface heads increase with stage, although, for a given stage rise, subsurface heads increase more in the pool than upstream (Figure 6a); and (2) VHGs tend to increase with stream stage upstream of the step, whereas they do not exhibit any trend in the pool (Figure 6b). In other words, the amplitude of subsurface head fluctuations is larger downstream than upstream, with downstream VHGs remaining constant because stage and heads change at a similar rate. In effect, the amplitude of stream stage fluctuations is also higher downstream than upstream (Figure 6c).

Continuous time series of hydraulic heads. Water levels, logged at 15 min intervals, were calibrated with the discrete measurements. After calibration the residual error between logged and dipped data was estimated at ±1.1 cm, certainly as a result of the logger’s sensitivity to temperature fluctuations. Stream stage values were
adjusted to each location by using polynomial regressions based on the relationship between dipped values at the location of the stage logger and at all other locations. As a result of this transformation, the maximum difference between logged and dipped stage increased to ±1.5 cm. Nevertheless, the error between two successive manual measurements usually did not exceed 0.5 cm. The propagation of this error in estimates of VHGs means that uncertainty is higher at shallow depths and low VHGs.

At a first glance, the relative constancy of most VHG time series is striking given that several flow events occurred in July 2008 (Figures 7, 8). Particularly for those piezometers with low uncertainties on VHG measurements, time series clearly exhibit little variation, even during high flow events (Figure 7: P1, P6b, P7b), thus implying that subsurface heads peak and recede at the same time and the same rate as stage. It could be argued that there is a hydraulic connection between the piezometer’s screen and the sediment–water interface, but the fact that these gradients are positive, and remain so during the rapid elevation of water levels, eliminates this possibility.

A robust interpretation of VHG time series with a higher level of uncertainty is more delicate (in particular P2, P7a, P9). It appears that an event-related pattern repeats itself throughout several datasets: in the rising limb of flow events, gradients decrease and in the falling limb they start increasing for a period of 2 days at most, indicating that subsurface heads rise and recede slower than stage (P2, P7a during the 12/7 event). These variations appear to cover a relatively short period, relative to the month of survey. Finally the head time series recorded at the floodplain’s edge shows some evidence of bank storage, as the recession rates are not as steep as those of the stream hydrograph (Figure 8). Although, strictly speaking, a piezometer does not measure the water table elevation, it is assumed here to provide a reliable proxy given the distance to the channel (40 m) and the relatively shallow depth of the screen (1-1 m below ground level).

Figure 7. (a) Continuous and discrete VHG time series; the top left of each plot displays the logger’s ID and the mid-screen depth; error bars on discrete data reflect a ±6 mm error on Δh, and error bars on continuous data reflect ±10 mm error on Δh, in addition to the instrument’s noise; (b) location of the loggers
DISCUSSION

Spatial variability of permeability

Streambed permeability can control the pattern and intensity of subsurface fluxes. Although it may be impossible to infer permeability from geomorphological data alone (e.g., Conant, 2004), the research reported here suggests two potential relationships between permeability and geomorphology: (1) the lowest permeability zones are adjacent to the eroding banks; and (2) average permeability is lower in the pool than in the riffle. In some cases, the distribution of bed permeability may be attributed to sediment sorting (e.g., Cardenas and Zlotnik, 2003; Ryan and Boufadel, 2007). This can hold for the riffle versus pool pattern (discussed later), but the deposition of finer sediments next to eroding banks seems counter-intuitive, as the velocity of stream water is indeed higher in the outer part of the bend, where erosion shapes the bank. However, the zone of lower streambed permeability located upstream (Figure 3a) is probably caused by the apparent large slumps of silty bank material mixed with bed gravels. In this area, bed hydraulic conductivity values ($2.7 - 2.8 \times 10^{-5}$ m s$^{-1}$) lie between the low value of the adjacent bank ($0.1 - 1 \times 10^{-5}$ m s$^{-1}$) and the higher values of the left side of the channel ($6 - 12 \times 10^{-5}$ m s$^{-1}$). Additionally, more piezometers were clogged in this area than anywhere else. In the pool we assume that the same process takes place as the left bank shows a clear concave eroding profile. Caution, of course, is needed in generalizing these observations as the present study concerns a short reach. For example, bank slumps would only affect channel permeability where there is a sufficient contrast in grain size between channel and bank materials. Bank erosion itself is controlled by several factors including the degree of basal erosion and over-steepening, the resistance of the bank to slumping and the ability of the flow to remove slumped material (Bridge, 2003).

Our results show that this process of bank slumping can be superimposed by other hydraulic processes. Indeed the geometric mean of hydraulic conductivity is lower in the pool ($3.6 \times 10^{-5}$ m s$^{-1}$) than in the riffle ($5.7 \times 10^{-5}$ m s$^{-1}$)—a pattern that can be related to excess fine sediments settling preferentially in pools where stream velocity is lower (Sable and Wohl, 2006). Similar conditions of lower permeability in eroding areas and pools has been previously observed by Rouch (1992), in Malard et al. (2002).

Spatial variability of VHGs

At low stream flows, average VHGs are significantly higher in the pool than in the riffle—respectively 3% and 9% on 27 June. However, among the six piezometers located directly upstream of the step, only the two closest to the inner side of the bend show a potential for downwelling flow, probably reflecting a HEF system induced by the step. This area appears on the right-hand side of the stream, where the slope of the stream’s water surface above the step was observed to be the steepest. This suggests that on the other side of the channel, where the step’s slope is lower, discharge of ambient groundwater reduces the penetration depth of surface water. The relationship between discharging groundwater and the depth of HEF systems has been recognized for longitudinal profiles of the streambed, from the dune scale (Cardenas and Wilson, 2007) to the channel-unit scale (Wroblicky et al., 1998; Storey et al., 2003). Our observations point out that lateral variation may also control HEF patterns, and possibly more at riffles and steps located at the apex of a bend or at flow-through reaches (sensu Woessner, 2000). This adds to the fact that subsurface flow in the bed of a meander can also be controlled laterally by the supererevation of the water surface along the outer bank (Cardenas et al., 2004; Boano et al., 2006).

Spatial variability of fluxes derived from VHGs and dilution tests

Interestingly VHG-derived fluxes do not vary significantly in space because the two terms of Darcy’s equation ($K$ and $V_h$) are, despite a low coefficient of determination ($R^2 = 0.28$), inversely correlated (Figure 4). The relative homogeneity of groundwater discharge appears to be maintained by the streambed’s permeability distribution, which causes an increase of VHGs in low $K$ areas. These results prompt the following conclusion: VHGs are sometimes used to provide an estimate of the potential strength of vertical hydrological exchange (e.g. Boulton et al., 1999; Edwardson et al., 2003; LeFebvre et al., 2004; Pretty et al., 2006; Wondzell, 2006; Schmidt et al., 2007); yet depending on the permeability structure, the actual flux can be relatively homogeneous despite a high variability of VHGs. In the present settings, VHGs would be misleading indicators of the hydrological conditions.

Therefore when measuring VHGs, we recommend not only a systematic characterization of $K$, but also
analysing relationships between the vertical component of flow and any other parameter using the Darcian velocity, which is a more realistic estimate of the potential intensity of flux. In some limited cases only, VHGs can be used as self-sufficient estimators of flux; these include: (1) settings where permeability is assumed to be homogeneous; (2) estimations of temporal changes at a given location, where permeability is assumed constant e.g. Hanrahan, 2008; and (3) the use of VHG solely for estimating directional changes of flux, i.e. negative, positive, or no vertical flow, e.g. Anderson et al. (2005), Tillman et al. (2003) or Brunke and Gonser (1999). To complete these comments it is worth adding two elementary points: a positive or negative VH does not imply a hydrologic connection — although it may be reasonable to assume so, neither does it discard the possible occurrence of an inverse hydraulic gradient between the sediment–water interface and the screen depth. For example, in this study we acknowledge that surface water may infiltrate the channel bed at locations and times where VHG indicate upwelling. Particularly in the upstream riffle, the roughness of the gravel bed may cause surface water to infiltrate by turbulent diffusion and advective flow. In a flume study carried out by Packman et al. (2004) with 6 mm diameter gravels, turbulent diffusion seemed to affect the first 4–6 cm of the subsurface, even with a flat bed. Such flow patterns are very likely to occur in the River Leith but were not captured by VHG that integrated the head difference over deeper sections.

Where VHG are equal to zero, point-dilution tests provide a direct estimate of the horizontal flux (Figure 3d). Between the two locations where this assumption is valid (Figure 3b), the flux ratio ($q_h/q_v$) is 0.2, possibly as a result of the similar change in $K$ ($K_h/K_v = 1.25$). Where VHG are different from zero and vertical flow contributes to the dilution of the tracer, the actual horizontal flux will be lower than the calculated flux. Over the whole reach, this implies that horizontal fluxes are, in general, significantly lower than vertical fluxes. This type of measurements could constitute a reliable source of information to constrain a hydrological model of hyporheic flow. In this study, we cannot distinguish the relative contributions of horizontal and vertical flow, it is wiser to avoid any further interpretation.

Nevertheless, the results have general implications on similar applications of point-dilution tests. For example, the conductometric standpipe approach estimates the interstitial flow velocity by comparing field dilution rates to a set of laboratory-calibrated curves (Carling and Boole, 1986; Greig et al., 2005). Calibration is carried out in a flume with horizontal interstitial flow and repeated with a range of grain size distributions. Values of interstitial velocity in a streambed are then obtained by relating the field dilution rate to a curve calibrated with sediments of similar grain size characteristics. In their field application, Greig et al. (2005) found velocities that were comparable with values reported in other studies. However, our conclusions are that dilution tests will produce accurate results only when the assumption of horizontal flow is not violated. At the River Leith, for example, the flux derived from the dilution test on the deepest piezometer of the experimental set (mid-screen depth: 38 cm; VH: $+7 \pm 2\%$) ‘underestimated’ the vertical flux with a $1 : 3$ ratio $(9.9 \times 10^{-7} / 2.6 \times 10^{-6}$ ms$^{-1}$).

Finally, the sensitivity analysis performed on a HEF model by Wroblicky et al. (1998) suggests that low values of streambed hydraulic conductance in groundwater-fed streams tend to increase hydraulic gradients between the channel and the floodplain, limiting the extent of stream water infiltration. Therefore the assumption of horizontal flow is more likely to be transgressed in settings of lower permeability.

**Temporal variations of VHG: discrete and continuous measurements**

We have seen that riffle and pool VHG respond differently to fluctuations of stream stage. With an increase of stage, riffle VHG increase whereas pool VHG remain relatively constant. This may seem counter-intuitive as subsurface heads in the pool increase 20% more than in the riffle. We propose, however, three interpretations of such behaviour:

1. the channel’s morphology (banks are steeper in the pool) is such that the stage rises 25% more downstream than upstream (Figure 6c); therefore, the sharper rise of subsurface levels in the pool is not apparent from VHG because of an equally high rise of stage;
2. above the step, the slope of the water surface is lower at high flows; therefore, the presence of the step does not deflect upwelling flow as much, and VHG tend to be spatially more homogeneous;
3. in riffle–step–pool sequences, the upstream downwelling area tends to extend longitudinally further than the downstream upwelling area (Gooseff et al., 2006); therefore, temporal variations of HEFs may affect a larger area and have more influence on averaged heads upstream than downstream.

These mechanisms can, of course, act together. It is recognized that the intensity of ambient groundwater discharge controls the depth, flux and mean residence time of HEF systems (e.g. Cardenas and Wilson, 2007). Likewise our results show that the shape of such systems may also undergo asymmetrical variations, as a response to changing boundary conditions. A better understanding of these flow dynamics, which are 3D by nature, still requires research.

In this study, the asymmetrical response was actually subtle and the averaged VHG did not change, over time, more than 3%. Likewise, the continuous time series showed that VHG can also remain constant throughout flow events. This is more obvious where the measurement error is low (P1, P6b, P7b in Figure 7). Surprisingly, despite fast transitions from base to peak flow (a few hours, Figure 8), stream stage and subsurface heads can increase at the same rate. This synchronous response
seems to go against the ‘conventional wisdom’ that expects a switch between upwelling (gaining) and downwelling (loosing) conditions, in response to increased stage relative to the riparian water table (Figure 9a), e.g. Soulsby et al., 2001; Arntzen et al., 2006; Malcolm et al., 2006; Puckett et al., 2008. One rapid fall and rise of VHG appeared in the time series P1, associated with the last and highest flow peak on 3 August. However, it may be asked how VHGs could remain constant, given that peak flows are not caused by a general rise of the groundwater table. The answer probably lies in the subsurface head time series recorded at the edge of the floodplain. During the rising limb and for most events, these show a ‘synchronous’, yet dampened, response to stage fluctuations (given the logging time step and uncertainty), so if floodplain sediments are diffusive enough (see Sophocleous, 1991) to enable the propagation of a pressure wave over 40 m, it is possible that the upstream rise of stage propagates in the subsurface by a similar process, across the meander and/or along the stream, thus increasing heads in the streambed proportionally to stage (Figure 9b). The fact that the study reach was located downstream of a riffle may enhance this process, as it implies that a zone of elevated stage is close relative to the average stream slope. Such synchronous behaviour between stage and floodplain heads, at similar distances, was also reported by Jung et al. (2004).

Among those VHG time series characterized by a higher level of uncertainty, some appear to exhibit event-related variations, such as a rapid fall during the rising limb. In some cases, VHGs increased immediately after the peak flow that occurred on 12 July, presumably because stream stage receded faster than subsurface heads (e.g. P2–P5 in Figure 7). This increase lasted approximately 2 days, until the stream discharge reached the level of the seasonal recession curve (Figure 8) and VHGs started returning to their initial level. In a similar study where streamed permeability varied over three orders of magnitude, Arntzen et al. (2006) observed an inverse relationship between K and the temporal variability of VHG. In the present work, however, given the uncertainty of most time series, these variations could not be reliably quantified, and therefore no relationship between temporal patterns and other measured parameters (e.g. K, piezometer location) was assessed.

**Flow-event versus inter-seasonal dynamics: a hypothesis**

In many settings, as the wet season unfolds, the alluvial groundwater table rises more than the stream stage, and hydraulic gradients increase predominantly towards the stream (Figure 9c), e.g. Devito et al., 1996; Dahn et al., 1998; Wroblicky et al., 1998; Soulsby et al., 2001; Konrad, 2006; Malcolm et al., 2006; Opsahl et al., 2007. It is a common observation, although Fraser and Williams (1998) also observed higher streamward VHGs in seasons dominated by baseflow—in summer and winter.

In our work, which focused on single flow events, the stage appeared to rise more than the groundwater table at the outer edge of the floodplain (Figure 8), and VHGs either did not respond or showed a slight decrease during the rising limb. This decrease is relatively common (references in the Section on Temporal Variations of VHGs), again despite the study by Alexander and Caisse (2003) who observed, together with flow events, an increase of groundwater discharge to the stream.

We hypothesize that high stream flows are likely to be associated with: (1) high streamward VHGs when considering inter-seasonal fluctuations, and (2) constant or lower VHGs when considering event-flow fluctuations. In the second case, we further hypothesize that highly diffusive (or high K) sediments will favour constant VHGs by enabling the propagation of pressure waves, whereas low-diffusivity (or low K) materials will favour a reduction of VHGs towards the stream, and possibly a gradient reversal.

**The interaction of multiple flow processes**

With respect to mechanisms controlling GW/SW mixing in the near stream environment, our study examined two different processes: (1) hydrostatically induced advection associated with a riffle–step–pool sequence; (2) and transient processes associated with stream stage variations. Potentially, both transient and steady processes can control GW/SW mixing. However, on this particular reach and within the hydrological conditions of this study, transient processes did not appear to significantly affect the time-averaged VHGs. Future research should attempt to better understand in which conditions VHGs or are not affected by stream stage variations.

**Figure 9. Conceptual cross sections of (a) a channel and its adjacent bank, representing three types of VHG variations associated with an increase of stream stage, for a reach that is gaining in low stage conditions. The relative amplitude of groundwater level (ΔGW level) and stream stage (ΔStage) fluctuations is different for each section. Sections’ left and right panes depict, by symmetry, the same profile in low and high stage conditions, respectively. Subsurface water is shown in grey, stream stage is represented by a horizontal line, arrows show the general flow direction, and an in-stream piezometer indicates the relative intensity and direction of VHG.**
CONCLUSIONS

This paper describes the spatial and temporal variability of hyporheic flow observed in a riffle–step–pool sequence. The study site is located in a groundwater-fed reach and field work was carried out during low flow conditions. Data were collected through a network of piezometers, a subset of which was equipped with high-frequency logging pressure transducers.

We showed that hydraulic conductivity and VHGs can exhibit significant lateral changes over a riffle–step–pool sequence. Although such variations have been previously reported, they are perhaps still underestimated compared to longitudinal patterns. Lateral and longitudinal variations of permeability were related to geomorphology in two ways: the average permeability was lower in the pool than in the riffle, a fact that is commonly attributed to grain size sorting; and permeability was lower in the vicinity of the eroding banks, probably as a result of the degrading slumps of low permeability bank material. Most importantly the downwelling area, probably induced by the step, only covered a minor fraction of the transversal stream section. This lateral inversion of the VHG was presumably controlled by the interaction between the ambient groundwater discharge and the slope of the stream surface above the step. Our study stressed the limits of VHGs as indicators of flow exchange intensity, as VHGs displayed spatially more variability than the vertical fluxes.

The temporal analysis of data indicated that changes in stage did not have the same effect on VHGs in the pool and the riffle. Additionally several continuous VHG time series demonstrated that stream stage and subsurface heads could fluctuate at the same rate, with the result that VHGs remain constant during stream flow events. We observed that stage variations and the response of the subsurface water level 40 m away from the channel could fluctuate at the same rate, with the result that VHGs remain constant during stream flow events. We showed that hydraulic conductivity and VHGs can

ACKNOWLEDGEMENTS

We are grateful to Kat Giddings and Kelly Cooper for their help with field work, and to Geoff Johnston and Phil Smith for their technical support. This work was supported by Environment Agency funding awarded to the Lancaster Environment Centre. We greatly appreciate the comments from two anonymous reviewers, which helped strengthen the manuscript.

REFERENCES


