This is the accepted manuscript version of the contribution published as:

Munz, M., Oswald, S.E., Schmidt, C. (2017): Coupled long-term simulation of reach-scale water and heat fluxes across the river-groundwater interface for retrieving hyporheic residence times and temperature dynamics
Water Resour. Res. 53 (11), 8900 - 8924

The publisher's version is available at:

http://dx.doi.org/10.1002/2017WR020667
Coupled long-term simulation of reach-scale water and heat fluxes across the river-groundwater interface for retrieving hyporheic residence times and temperature dynamics

Matthias Munz\textsuperscript{1}, Sascha E. Oswald\textsuperscript{1} and Christian Schmidt\textsuperscript{2}

\textsuperscript{1}University of Potsdam, Institute of Earth and Environmental Science, Potsdam, Germany
\textsuperscript{2}Helmholtz Centre for Environmental Research – UFZ, Department of Hydrogeology, Leipzig, Germany

Corresponding author: Matthias Munz (munz@uni-potsdam.de)

Key Points:

\begin{itemize}
  \item Calibrated and validated 3-D fully-integrated model give an accurate representation of the real system with predictive capabilities
  \item Non-submerged streambed structures cause substantial thermal heterogeneity within the saturated sediment at the reach-scale
  \item Hyporheic flow path temperature strongly depends on flow path residence time and penetration depth
\end{itemize}
Abstract

Flow patterns in conjunction with seasonal and diurnal temperature variations control ecological and biogeochemical conditions in hyporheic sediments. In particular, hyporheic temperatures have a great impact on many temperature-sensitive microbial processes. In this study, we used 3-D coupled water flow and heat transport simulations applying the HydroGeoSphere code in combination with high resolution observations of hydraulic heads and temperatures to quantify reach-scale water and heat flux across the river-groundwater interface and hyporheic temperature dynamics of a lowland gravel-bed river. The model was calibrated in order to constrain estimates of the most sensitive model parameters. The magnitude and variations of the simulated temperatures matched the observed ones, with an average mean absolute error of 0.7 °C and an average Nash Sutcliffe Efficiency of 0.87. Our results indicate that non-submerged streambed structures such as gravel bars cause substantial thermal heterogeneity within the saturated sediment at the reach-scale. Individual hyporheic flow path temperatures strongly depend on the flow path residence time, flow path depth, river and groundwater temperature. Variations in individual hyporheic flow path temperatures were up to 7.9 °C, significantly higher than the daily average (2.8 °C), but still lower than the average seasonal hyporheic temperature difference (19.2 °C). The distribution between flow path temperatures and residence times follow a power law relationship with exponent of about 0.37. Based on this empirical relation, we further estimated the influence of hyporheic flow path residence time and temperature on oxygen consumption which was found to partly increase by up to 29% in simulations.
1. Introduction

Flow and temperature patterns control ecological and biogeochemical conditions in hyporheic sediments [Findlay, 1995; Boulton et al., 1998; Krause et al., 2011]. The infiltration of river water, solutes and thermal energy, into the subsurface and its delayed exfiltration back into the river has a major influence on the transport and transformation of nutrients [Boano et al., 2010; Zarnetske et al., 2012; Krause et al., 2013] and the thermal regime within the shallow streambed [Arrigoni et al., 2008; Sawyer et al., 2012].

Solute turnover of dissolved substances in the hyporheic zone has been shown to depend on residence times [Boano et al., 2010; Zarnetske et al., 2011; Marzadri et al., 2012], penetration depth of surface water [Harvey et al., 2013] and stream discharge dynamics [Trauth et al., 2017]. For instance, long residence times promote denitrification, because usually oxygen must first be sufficiently depleted before denitrification occurs [Trauth et al., 2015]. Nevertheless, Briggs et al. [2015] showed that denitrification might also occur in oxic-saturated sediments along the entire streamline because of the presence of anoxic microzones.

In particular, denitrification is of great importance because it enables the permanent removal of nitrogen from the aquatic system [Harvey and Bencala, 1993; Zarnetzke et al. 2011; 2012]. Whether the hyporheic zone acts as a nitrate source or sink also depends strongly on the ratio between the in-stream concentrations of ammonium and nitrate [Marzadri et al., 2011; Zheng et al., 2016]. In general, hyporheic exchange flows have a crucial impact on the health of our riparian ecosystems by increasing solute residence times and thus solute exposure to microbial communities, which in turn fosters biogeochemical cycling of nutrients and contaminants [Mulholland et al., 2008; Battin et al. 2016]. Besides subsurface residence times also temperature has been shown to control microbial processes in hyporheic sediments [Thamdrup and Fleischer, 1998; Acuna et al., 2008]. Microbially mediated solute transformation rates in gravel bed rivers have been observed to vary by a factor of 10 for temperature differences of 10 °C [Vieweg et al., 2016]. Diurnal temperature variations in the stream water propagate into the hyporheic zone by heat advection and conduction [Anderson, 2005]. These variations are increasingly damped at increasing depths in the sediment [Arrigoni et al., 2008], and hence also with increased flow path lengths and residence times [Marzadri et al., 2013]. Thus, we hypothesize that there is a coupled effect of residence times and hyporheic temperatures on solute processing in the hyporheic zone.
To study the spatial and temporal variability of water and heat flux in riverbed sediments, a broad range of experimental methods exist [e.g. Kalbus et al., 2006; Rau et al., 2014]. Despite recent improvements in logger technology and measurement methods, the experimental approaches are limited to device-dependent temporal and spatial resolutions and to their ability to cover entire river reaches. Using heat as a natural tracer offers quantitative methods to estimate one-dimensional vertical water fluxes in near surface sediments [e.g. Hatch et al., 2006; Munz et al., 2017]. The presence of horizontal flows, including hyporheic exchange, will cause uncertainty in these analytical solution results [Brookfield et al., 2013; Cuthbert and Mackay, 2013; Revers and Hatch, 2016]. In contrast, three-dimensional flow and heat transport models can be used to simulate the complex river-groundwater exchange and thermal transport process in river-riparian systems, spatially distributed and at a high temporal resolution [Brookfield et al., 2009; Nützmann et al., 2014; Karan et al., 2014].

The application of distributed, fully-integrated models requires the parametrization of the flow and heat transport equations. Improvements in model calibration have been made by coupling both temperature and pressure to constrain estimates of hydraulic and thermal properties [Doussan et al., 1994; Naranjo et al., 2012; Karan et al., 2014]. Especially the streambed hydraulic conductivity and its spatial representation strongly affect simulated river-groundwater exchange, hyporheic flow paths and residence times across dune-like streambed structures [Salehin et al., 2004; Sawyer and Cardenas, 2009; Tonina et al., 2016] and along the river reach [Naranjo et al., 2012; Karan et al., 2014].

At any scale, the large number of input parameters for spatially distributed, fully-integrated models and their interactions is a challenge for an effective model parameterization. Thus, there is a strong need to identify the key parameters controlling hydrological and thermal processes, as well as a subset of parameters to be targeted in more computationally-intensive parameter estimation processes. Global sensitivity analysis, accounting for model non-linearity and parameter interactions, helps to prioritize the parameters to be included in subsequent parameter estimation and to identify the dominant processes within the study domain [Van Griensven et al., 2006; Saltelli et al., 2008; Srivastava et al., 2014].

Detailed process-based numerical models were developed at the scale of single geomorphological structures, also coupled with heat and solute transport, including multiple reacting and interacting species (e.g. Zheng et al. 2016). These models allow for a holistic investigation of exchange and transformation processes in the hyporheic zone. Zheng et al. 2016 conducted a series of numerical reactive solute transport simulations of hyporheic
processes within a dune with different uniform temperatures. Their simulations showed that increased sediment temperatures resulted in shallower oxic-anoxic zone boundaries as a consequence of increased aerobic respiration and nitrification rates. While this study is one of the first to couple and integrate fluid flow and temperature dependent reactive solute transport, their analysis did not consider heterogeneous hyporheic temperature patterns.

The aim of our study was to investigate reach-scale hyporheic exchange fluxes as well as the associated heterogeneous hyporheic temperatures and their influence on temperature-sensitive biogeochemical processes, such as hyporheic oxygen consumption. We addressed this goal by transient 3-D coupled water flow and heat transport simulations in combination with established particle tracking techniques to quantify hyporheic transport and thermal properties and by showing the impact of temperature variations on oxygen consumption along hyporheic flow path.

2. Methods

2.1. Study Site and Data Collection

The study site is located within the catchment of the Selke River, a third-order perennial river (mean low discharge = 0.23 m$^3$ s$^{-1}$, mean discharge = 1.51 m$^3$ s$^{-1}$, mean high discharge = 15.6 m$^3$ s$^{-1}$) in the northern part of the Harz Mountains in central Germany (51°43'39.9"N 11°18'53.2"E) as high-intensity test-site of a TERENO observatory [Wollschläger et al., 2017]. The studied river reach is 250 m long, encompassing a variety of natural fluvial morphological structures, such as a point bar (27 m long and 3.5 m wide during base flow) and an in-stream gravel bar (20 m long and 6 m wide during base flow), which are typical for gravel bed rivers [Montgomery et al., 1995; Wondzell and Gooseff, 2013]. The dominant streambed structures present were formed during flood events in 2010 and persisted without major sediment redistributions until June 2013. Periodic inspection at the logger locations during regularly performed data read-out did not indicate scour or erosion in proximity to the sensor. Minor sediment redistribution might have occurred within the river reach during the observation period, nevertheless, this was neither observed nor could it be included in the simulation set-up. The topography of the study site was surveyed in 2011 using differential GPS in combination with a laser tachymeter (Trimble GPS R8), with 1633 data points. Average aerial resolution of the survey was 0.7 m (resolution was increased at locations of high variability in riverbed topography). The local aquifer consists of up to 8 m-thick fluvial
sediments, with grain sizes ranging from medium sands to coarse gravels underlain by less permeable clay and silt deposits which form the bottom of the alluvial aquifer. The studied river reach is characterized by overall losing conditions within the river channel, whereby the water flow direction within the streambed often has a strong horizontal flow component oblique to the riverbed interface [Munz et al., 2016]. Patterns of hyporheic infiltration and exfiltration were observed at the streambed around the in-stream gravel bar [Trauth et al., 2015]. The flow patterns within the gravel bar are highly variable in time, depending on the water level in the river (discharge) and the ambient groundwater flow direction [Trauth et al., 2015; Munz et al., 2016].

Continuous measurements of hydraulic heads and temperatures at different depths along the river bank (1.7 m, 2.7 m, 4.7 m) and within the river at a temporal resolution of 10 min were conducted over a two-year period from May 2011 until June 2013. Additionally, riverbed temperature profiles were recorded at 20 locations with seven points each in the depth interval 0 - 0.65 m below the streambed surface every 10 minutes using Multi Level Temperature Probes (Fig. 1). The vertical temperature profiles were measured along the thalweg, as well as at the head, crest and tail of the dominant fluvial morphological structures. A detailed description of the experimental setup, the temperature sensors and their operation times can be found in Munz et al. [2011a; 2016].

### 2.2. Hydraulic Conductivity and Thermal Diffusivity

A detailed characterization of the hydraulic conductivity of the streambed and the aquifer is required to accurately simulate hydrological fluxes [Naranjo et al., 2012; Karan et al., 2014]. Slug tests were carried out at 70 locations distributed across the river reach to estimate the horizontal hydraulic conductivity ($K_h$). Individual slug tests were conducted at four different depth intervals below the streambed surface (0 - 0.2 m, >0.2 - 0.4 m, >0.4 - 0.8 m, and >0.8 m) at the same location. Major measurement campaigns were conducted in September 2011 – 2013 (detail about the slug test procedure can be found in Munz et al. [2016]). To estimate the hydraulic conductivity of the aquifer, slug tests were also carried out in the existing groundwater observation wells. The data was analysed using the case G, basic time lag equation from Hvorslev [1951]. The calculated $K_h$ were corrected to a reference temperature of 10 °C. Streambed hydraulic conductivities were grouped based on observation year (2011,
streambed depth (0 - 0.2 m, >0.2 - 0.4 m, >0.4 m - 0.8 m, and >0.8 m), and their relative location with respect to the geomorphological structures (head, crest, tail).

In order to estimate the thermal diffusivity ($K_e$) of the streambed sediments we applied the field method proposed by Luce et al. [2013] (see equation 64c) to the one-dimensional streambed temperature profiles. Streambed thermal diffusivity was grouped for the relative location at the study side (thalweg, head, tail). The temperature profiles at the crest of the point bar and the in-stream gravel bar were excluded from these analysis as the subsurface flow direction is known to have a significant horizontal component; i.e. could not be assumed to be one-dimensional (for details see analysis of vertical temperature profiles presented in Munz et al. [2016]). The analysis of variance (MATLAB Statistics Toolbox) was used to test for differences of $K_h$ and $K_e$ in space and time. All tested variations with a significance level (p-value) smaller than 0.05 were taken as being significant.

2.3. Integrated Surface-Subsurface Flow and Heat Transport Simulations

2.3.1. HydroGeoSphere Model

The fully-integrated surface-subsurface flow and transport model HydroGeoSphere (HGS) includes simulations of water flow and heat transport [Therrien et al., 2010]. HGS is a finite element code that solves Richard’s equation for variable saturated flow in 3-D. Surface water flow is simulated using the numerical solution of the diffusion-wave equation (an approximation of the Saint-Venant equations) together with Manning’s equation. The Manning formula is an empirical formula estimating the average stream velocity depending on the streambed roughness ($n$). Details for the governing flow and transport equations can be found in Therrien et al. [2010] and Brookfield et al. [2009].

Established coupling methods in fully integrated codes are the continuity of pressure or the dual conductance concept (dual node approach), based on a first-order exchange coefficient ($l_e$) as a numerical parameter. Within the continuity of pressure approach the surface water and porous media flow and transport equations are solved simultaneously at a single interface node; however, this approach requires small time steps in order to guarantee numerical stability leading to extremely long simulation times [Ebel et al., 2009; Huang and Yeh, 2009]. For the dual conductance concept the flow and transport equations are solved for separate surface and subsurface nodes assuming that the exchange flux depends on the hydraulic head gradient across a coupling interface as well as the vertical saturated hydraulic conductivity and the first-order exchange coefficient [Therrien et al., 2010]. The dual node approach can
approximate the continuity of pressure approach, if \( l_e \) is sufficiently small [Huang and Yeh, 2009, Liggett et al., 2011]. To guarantee numerical stability in the coupled water flow and heat transport simulation, the reach scale model was set up using the dual node approach to couple the surface and subsurface flow domains. The HydroGeoSphere code has been successfully applied using the dual node approach in simulating combined flow of water, heat and solute across the hillslope [e.g. Frei and Fleckenstein, 2014] and catchment scale [e.g. Liggett et al., 2015] and for surface water groundwater interaction at lakes [e.g. Ala-aho et al., 2015] and river stretches [e.g. Brookfield et al., 2009].

2.3.2. Model Setup (Baseline Model)

We applied a transient 3-D coupled water flow and heat transport model to simulate reach-scale water and heat flux across the river-groundwater interface and hyporheic temperature dynamics for a two-years' time-series of naturally occurring hydrological and thermal conditions. The dimensions of the model domain were 80 m x 250 m x 9 m (width, length, thickness; Fig. 1). The domain thickness was set according to the less permeable clay and silt deposits that form the bottom of the alluvial aquifer. A 2-D triangular finite element mesh was generated for the top of the domain, representing the ground and riverbed surface (GRID BUILDER, McLaren, 2008). Element size was set to 10 m in the floodplain, but subsequently refined to 0.25 m in the riverbed. The generated 2-D mesh was superimposed onto a 3-D finite element mesh consisting of 28 model layers, with a vertical discretization of 0.05 m for the upper 12 layers, which increased to 0.4 m at the base of the 3-D mesh, located at 151.2 m a.s.l. The whole model domain was discretized into approximately 250,000 elements. The chosen node spacing provided an adequate spatial resolution for representing riverbed topography, in order to capture steep hydraulic head and temperature gradients, both horizontally and vertically. The surface topography of the domain was interpolated based on the riverbed and riparian floodplain DGPS survey with a high vertical resolution to the elevation map, including all locations of the numerical mesh nodes (Fig. 1).

The hydraulic conductivity of the streambed (0 - 0.8 m below the streambed surface) was defined as heterogeneous and anisotropic according to the field observations (cp. section 3.1). To implement the observed spatial variability of streambed \( K_h \) we used conditional, sequential Gaussian simulation (SGS) with histogram reproduction [Hansen and Mosegaard, 2008] using the geostatistical MATLAB toolbox mGstat. In SGS, the measured hydraulic conductivities
were transformed to standard Gaussian values and the semivariogram of the transformed data was estimated. Measured $K_h$ values were replicated at their locations, and values at the remaining locations were randomly drawn from the conditional cumulative distribution function, defined by the kriging mean and variance. SGS preserved the observed mean, variance and the spatial structure of the streambed hydraulic conductivities. The aquifer (~0.8 m – 9 m) was divided into three zones of homogeneous and anisotropic $K_h$ corresponding to the depth of the groundwater monitoring wells (Fig. 1, Table 1). The anisotropy ratio was obtained through calibration. The thermal parameters of the streambed were defined as spatially homogeneous according to the field observations (cp. section 3.1) (Table 1).

A first estimate of the initial conditions for hydraulic head and temperature was based on interpolated field data. To determine a realistic initial condition, the transient model was run over a full seasonal temperature cycle as a spin-up period with the boundary condition from May 2011 to May 2012 (when hydrological and thermal regimes were comparable). The head and temperature distributions at the end of this one-year spin-up were used as updated initial conditions. According to the updated initial condition of the heads, the first order exchange coefficient was optimized in a way that the river groundwater exchange is insensitive to further reductions of $l_e$ (the coupled system was near continuity), according to the approach proposed by Liggett et al. [2011]. The transient simulation was carried out for the period from May 2011 to June 2013. The boundary conditions of the transient simulation were updated every hour based on hourly averaged measured hydraulic heads and temperatures. Boundary conditions were set as follows:

- **Prescribed hydraulic head** (Dirichlet) at the sides of the model domain. Based on measured hydraulic heads along the river and in the wider area of the floodplain (four observation wells outside the study site), the heads were linearly interpolated for each input time step and a specific hydraulic head was extracted at the location of each boundary node. These extracted hydraulic heads were extended homogeneously for the entire domain depth (green lines in Fig. 1).

- **Prescribed water flux** (Neumann) at the river inlet, based on discharge measurements at the nearby gauging station Meisdorf (51°41'29.1"N 11°17'02.2"E). Several manual discharge measurements at the study site have shown that the discharge at the gauging station is practically equal to the discharge at the domain boundary.

- **Critical-depth** boundary is applied at the river outlet, forcing the water depth at the boundary to be equal to the critical depth, and discharge to be calculated by the model.
Precipitation and potential evapotranspiration applied to the surface. The precipitation was measured at the nearby climate station Aschersleben (51°43'33.6"N 11°30'39.6"E) of the German weather service. Potential evapotranspiration was calculated with external routines based on measured climate conditions and crop vegetation (annual course of leaf area index) using the Penman Monteith equation [Monteith, 1981].

No flow boundaries were taken for the bottom of the model domain, representing the top of the confining layer.

Prescribed groundwater temperature at the sides of the model domain, based on measured GW temperature at the flood plain (1.7 m, 2.7 m, 4.7 m below surface). Based on hourly average borehole temperatures, vertical temperature profiles were interpolated for each simulation time step and assigned to the boundary nodes (orange lines in Fig. 1). This ensured that the water entering over these boundaries was consistent with the measured groundwater temperature at the study site. At the bottom of the domain a constant value of 9.03 °C (mean groundwater temperatures over the simulation period) was assigned to all bottom nodes (dashed orange line in Fig. 1).

Prescribed river temperature at the river inlet. Measured river water temperature (10 min frequency) was aggregated into hourly values and assigned to the nodes at the river inlet.

Heat flux at the river outlet is calculated by the model based on calculated river discharge and temperature at the boundary nodes.

Atmospheric energy input calculated by means of the ambient air temperature and incoming short wave and long wave solar radiation using the HGS atmospheric module [Therrien et al., 2010, Verseghy et al., 1991, supporting information]. Air temperature was measured at the study site (0.2 m above ground). Incoming short wave and long wave solar radiation was measured at the nearby climate station Aschersleben and applied to all surface nodes. Ground surface albedo depends on water content of the surface elements [Therrien et al., 2010]. Specific values for dry and saturated (water) conditions are shown in table 1. The atmospheric thermal inputs are calculated explicitly, and are treated as a source/sink term in the thermal energy transport equation for the surface regime [Therrien et al., 2010]. This approach allowed the simulation of river water temperature based on atmospheric and subsurface energy input (river temperature not fixed at the nodes of the surface domain).
2.4. Parameter Sensitivity Analysis and Calibration Strategy

We conducted a global sensitivity analysis to understand the effect of model parameterization on the simulation results in order to effectively calibrate the fully-integrated model and prove the benefits of including temperature in addition to hydraulic heads for model parameterization of fully-integrated numerical models. Based on these analyses, we identified the most relevant model parameters and output variables for model calibration (parameter prioritization).

All parameters classified as sensitive to the simulation results were calibrated using PEST [John Doherty, Watermark Numerical Computing, 2005], a model-independent software package often applied for parameter estimation and uncertainty analysis of hydrological and environmental computer models. PEST reduces the discrepancies between model outputs and field observations to a minimum in the weighted-least-squares sense by iteratively adjusting the calibration parameters. For practicability, calibration and sensitivity analysis were limited to a 90-day period from October 2011 to January 2012. This period was chosen because it represents a considerable dynamics in river discharge and temperature. Each of the sensitivity and calibration model runs was initialized by starting with the base line pressure head and temperature distribution for May 2011, followed by half of a temperature cycle as model spin-up period for the perturbed parameter setup.

For the global sensitivity analysis, we applied Morris' elementary effects screening method [Morris, 1991] using the R software package sensitivity [R Development Core Team, 2011; Pujol, 2009]. The Morris method was selected because of avoiding otherwise too high computational requirements for simulating transient, coupled water flow and heat transport. The Morris screening method is designed to work at low computational cost and provides the average ($\mu$) and standard deviations ($\sigma$) of local sensitivities obtained at different locations in parameter space. A high value of $\mu$ indicates a large influence of the corresponding input parameter on model output and a large value of $\sigma$ indicates that the corresponding input parameter is attributed to non-linearity and/or interactions with other model parameters [Morris, 1991; Saltelli et al., 2004; Srivastava et al., 2014]. The mean value of $\mu$ over all parameters considered ($\bar{\mu}$) serves as criterion to identify more influential parameters over less influential ones, namely all observed parameters with $\mu > \bar{\mu}$ are selected as most sensitive parameters and are considered in model calibration.
The results of Morris’ method are presented as scatter plots of $\mu$ versus $\sigma$ and a linear function representing double the standard error of the mean (SEM) of all elementary effects $\mu = 2 \times SEM$; with $SEM = \frac{\sigma}{\sqrt{r}}$ [Morris, 1991; Srivastava et al., 2014]. Parameters lying below that dotted line in figure 2 correspond to parameters indicating high sensitivity but little non-linearity and/or interaction effects. Such parameters are sensitive over the entire parameter space without much variability in the individual sensitivity. Based on initial sensitivity evaluations carried out for the transient coupled water flow and heat transport model, a setting with 60 repetitions and 5 levels was chosen for the Morris screening (requiring 780 model runs). These values correspond to settings proposed by Morris [1991] and Campolongo et al. [2007] to achieve meaningful analysis. With these settings, stable results were obtained for extended analysis.

The calibrated model was run from May 2011 to June 2013. To quantify the model quality, we used the average water balance error, the mean absolute error (MAE) and the Nash Sutcliffe efficiency (NSE) criteria. The model results were evaluated with Tecplot 360, Version 2012 R1 (TecPlot Inc.), a model-independent visualization and data analysis tool.

2.5. Flow Path Analysis via Particle Tracking

A particle tracking analysis was performed to evaluate hyporheic flow path lengths, penetration depth (maximum vertical extend), travel times and hyporheic flow path temperatures. Hyporheic flow path are defined by particle tracks that first enter and then again exit the subsurface domain via the top boundary. Forward advective particle tracking (based on a second-order Runge-Kutta integration) was conducted using Tecplot 360 [TECPLOT User’s Manual, 2012]. Massless particles were released from each boundary node of the wetted streambed area. Thus, the number of particles released varied between 891 and 10,713, depending on the size of the area that is fully submerged in the HGS-simulations. Particle transport followed the pore water velocity field extracted from the HGS simulations for selected time steps. This approach implicitly assumes that the actual discharge variations in the relevant time window are sufficiently slow to allow hydraulic head in the surface and the sediment to reach a quasi-equilibrium, (i.e. subsurface residence time was short relative the transition of discharge). Boano et al. [2010, 2013] showed that the temporal variability of hyporheic residence times is relatively limited compared to the variations in stream discharge. In the subsequent analysis, particle tracks following hyporheic flow paths were extracted to
calculate the length, depth, residence time and temperature for each individual hyporheic flow path. Average particle track (flow path) temperature was calculated along each individual hyporheic flow path in order to get a single flow path specific temperature estimate (flow path temperature). To calculate the hyporheic exchange fluxes within the study site, the hyporheic flow paths were flux weighted using the simulated river-groundwater exchange flux by the HGS model at the infiltration point of each particle.

2.6. Heat Transport and Subsurface Temperature Patterns

To specify heat transport processes at the study site we calculated average advective and conductive heat fluxes across the river groundwater interface. Therefore the cell-centered advective and conductive heat fluxes of the river cells are multiplied by the volume of the particular cell, the results are summed over all cells and divided by the river area. Furthermore, we calculated the thermal Peclet number for each cell of the model domain, as the ratio between heat transported by advection and heat transported by conduction [Anderson, 2005]. Snap shots of subsurface temperature were taken at low river discharges (0.25 m$^3$s$^{-1}$), as well as spring-summer (20 May 2012) and autumn-winter (22 November 2011) temperature conditions characterised by strong temperature gradients between river and groundwater. For the calculation of daily average subsurface temperatures and daily amplitudes, the simulation results were extracted every 30 minutes over the course of one day. On this basis, daily average temperatures and amplitudes were calculated and extracted along the river channel transect (Fig. 1).

In order to normalize different temperature conditions, a normalized flow path temperature difference ($T_{norm}$) was defined and evaluated for each simulation time step with an absolute temperature difference between river and groundwater temperatures larger than 3 °C. The normalization was based on the average flow path temperature ($T_{flow\ path}$), the river temperature ($T_{river}$) of the particular simulation time $t$ and the long term average groundwater temperature ($T_{GW}$) as:

$$T(t)_{norm} = \frac{T(t)_{flow\ path} - T(t)_{river}}{T_{GW} - T(t)_{river}}; \text{if } |T_{GW} - T(t)_{river}| > \Delta T_{min} > 0 \degree C \ (1).$$

Here $\Delta T_{min}$ was chosen to be 3°C. The value of $T_{norm}$ is usually, but not necessarily between 0, if flow path temperature is equal to river temperature, and 1, if flow path temperature is...
equal to groundwater temperature. The general relation was applied to selected snapshots, highlighting the dependence of hyporheic residence times on hyporheic flow path temperatures for positive and negative temperature gradients between the river and groundwater.

2.7. Temperature dependent Biogeochemical Processes in the Streambed

To improve our understanding of spatially variable hyporheic temperature patterns on biogeochemical processes in the streambed we choose to focus on oxygen consumption along hyporheic flow path because of its importance for the metabolism of hyporheic organisms and in identifying prevailing redox conditions which also control nitrogen transformations [Marzadri et al., 2012; Zarnetzke et al., 2013]. We assume that transport of dissolved oxygen is dominated by advection (diffusion and dispersion along and between hyporheic flow paths is neglected) and that the respiration rate coefficient is temperature dependent. The fraction of consumed oxygen along hyporheic flow paths (oxygen consumption) is described by first order reaction kinetics (González-Pinzón et al., 2012: Vieweg et al., 2016):

\[
\text{fraction of consumed oxygen} = 1 - e^{(-k(t) \times t_{res}(t))} \quad (2)
\]

where \( k \) is the first order respiration rate coefficient and \( t_{res} \) is the residence time. The temperature dependent first-order respiration rate coefficient is calculated based on an empirical relation derived at the in-stream gravel bar (Vieweg et al., 2016):

\[
k(t) = 0.187 \, d^{-1} \times e^{0.232 \, ^{\circ}C^{-1} \times T_{\text{effective}}(t)} \quad (3)
\]

where \( T_{\text{effective}}(t) \) is the mean effective flow path temperature. To estimate the influence of residence time-dependent flow path temperatures on oxygen consumption, we calculated the relative error (%) between fraction of consumed oxygen using river temperature as the effective temperature \( (T_{\text{effective}}(t)) \), and residence time-dependent flow path temperatures \( (T_{\text{flow path}}, \text{see Section 3.5, Equation 4}) \), as the effective temperature in Equation (3), respectively.

3. Results

3.1. Hydraulic Conductivity and Thermal Diffusivity
The mean hydraulic conductivity of the streambed was $1.4 \times 10^{-3}$ m s$^{-1}$ ranging from $1.1 \times 10^{-4}$ m s$^{-1}$ to $9.2 \times 10^{-3}$ m s$^{-1}$. We tested if the variance of hydraulic conductivity can be explained by spatial and temporal properties of the measurements (Table 2). Significant variation (p-value = 0.009) have only been found for the position along the streambed structure (head, crest, tail). The median ($1.2 \times 10^{-3}$ m s$^{-1}$) and the range ($9.0 \times 10^{-3}$ m s$^{-1}$) of hydraulic conductivity at the crest was higher than the median ($0.8 \times 10^{-3}$ m s$^{-1}$ / $0.7 \times 10^{-3}$ m s$^{-1}$) and range ($0.4 \times 10^{-3}$ m s$^{-1}$ / $0.6 \times 10^{-3}$ m s$^{-1}$) of the hydraulic conductivity at the head and center of the geomorphological structures (Fig. S1 in the supporting information). We did not observe significant variation of streambed hydraulic conductivity with increasing depth or over time (p-value > 0.05). In contrast, the measured hydraulic conductivity within the aquifer significantly decreased with depth (Table 3, p-value = 8.25 $\times$ 10$^{-5}$).

The mean thermal diffusivity of the streambed was $1.5 \times 10^{-6}$ m$^2$ s$^{-1}$ ranging from $3.9 \times 10^{-7}$ m$^2$ s$^{-1}$ to $2.9 \times 10^{-6}$ m$^2$ s$^{-1}$ (detailed distribution of $K_e$ is provided in the supporting information S2). We tested if the variance of thermal diffusivity can be explained by spatial properties of the measurements. We did not observe significant variation of streambed thermal diffusivity in space (p-value = 0.23), i.e. there is a rather homogeneous distribution of the streambed thermal parameters.

### 3.2. Parameter Sensitivity Analysis

River-groundwater exchange derived from the coupled water flow and heat transport simulations was substantially influenced by the properties characterizing the distribution of the hydraulic conductivity (mean, variance and anisotropy) and was also sensitive to the bulk density of the sediment, thermal conductivity and dispersion length (Fig. 2a). The hydraulic conductivity and dispersion length are directly related to the subsurface heat flux (advection-conduction equation), whereas the bulk density and thermal conductivity of the sediment are used to describe the temperature dependence of the hydraulic conductivity (Molson et al., 1992, supporting information).

River and groundwater hydraulic heads were found to be more sensitive to properties assigned to the surface domain, especially to the riverbed friction, than to the properties of the sediment (Fig. 2b). The riverbed friction mainly controls the water level in the river, according to the Manning Strickler formula [Therrien et al., 2010]. Changes in river level also altered the simulated hydraulic heads in the adjacent groundwater, but the hydraulic head was insensitive to the mean and variance of the hydraulic conductivity.
Besides the properties characterizing the distribution of the hydraulic conductivity (mean, variance and anisotropy), the most sensitive parameters affecting the groundwater and hyporheic temperatures were the thermal conductivity, bulk density and heat capacity (Fig. 2c, d). These parameters are directly related to the general equation for variably saturated subsurface heat transport (Molson et al., 1992, supporting information). In general, subsurface temperatures across the domain were influenced by the same parameters affecting the river-groundwater exchange flux (with some differences in the order of importance and apart from the dispersion length), highlighting the benefits of temperature in constraining fully-integrated models in order to accurately simulate river-groundwater exchange (Fig. 2). Most of the parameters observed had substantial direct effects on the simulation results, with weak non-linear/interaction effects. Only the dispersion length was characterized by strong non-linear/interaction effects on the simulated river-groundwater exchange flux (Fig. 2).

Figure 3a demonstrates the relation between hydraulic head bias, temperature bias and river-groundwater exchange flux for all sensitivity simulations. We observed a significant change in the mean and variance of the simulated exchange flux when temperature was used in addition to the hydraulic head information for model validation (normalized hydraulic head and temperature BIAS < 0.1). By including heat transport in the model validation, the model non-uniqueness could be reasonably constrained and the accuracy of flux estimation could be significantly improved (Fig. 3b).

### 3.3. Model Validation

The model was calibrated to the river and groundwater hydraulic heads, as well as to streambed and groundwater temperatures. Values for the auto-calibrated parameters are presented in Table 1. The simulated hydraulic heads in the river clearly represented the observed hydraulic heads (MAE = 0.03, NSE = 0.92) and also matched their seasonal dynamics (Fig. 4a). MAE and NSE comparing observed and simulated hydraulic heads in the groundwater ranged from 0.02 to 0.08 and from 0.63 to 0.95, respectively (Fig. 5, first row). There is one exception, observation point ML_5, where the simulated hydraulic head was systematically overestimated (MAE = 0.2).

Simulated temperatures in the river were clearly able to closely represent the observed temperatures (MAE = 0.3, NSE = 1) and also matched well their seasonal and daily dynamics.
proving the accurate model setup of the surface domain, where the temperature dynamics were mainly driven by river boundary conditions and atmospheric input. MAE and NSE comparing observed and simulated groundwater and riverbed temperatures ranged from 0.3 to 1.75 and from 0.56 to 1, respectively (Fig. 5). In general, the riverbed temperatures were reproduced accurately for all selected locations within the riverbed, the head, crest and tail of the intensively monitored geomorphological structures. The differences between observed and simulated temperatures slightly increased with increasing aquifer depth. This deviation was a mismatch in the simulated amplitude of the temperature signal. In most cases, the simulated annual temperature extremes were lower compared to the observed. That deviation was most obvious at ML_3, where the summer maximum temperatures were systematically underestimated and the winter minimum temperatures were systematically overestimated (Fig. 5). Overall, while using heads only leaves an ambiguity for fluxes, by including temperature data, which reflect also advective heat transport, the simulated fluxes are better constrained to the real values.

3.4. Water Flux across the River-Groundwater Interface and Hyporheic Exchange

In Figure 6, the simulated hydraulic head distribution, the resulting flow field (a-c), the rivergroundwater exchange flux (d-f) and hyporheic flow path lengths and residence times (g-i) are presented. The selected snapshots are representative for low (0.25 m$^3$s$^{-1}$), moderate (1.28 m$^3$s$^{-1}$) and high (3.65 m$^3$s$^{-1}$) river discharges and include a discharge-dependent shape of the river shoreline.

For low river discharges, hyporheic flow cells formed across the in-stream gravel bar, the point bar, and around the riffle (Fig. 6a). Median flow path length was 9.4 m, maximum depth was 4.2 m and the median residence time was 37.7 h. Flow path lengths and residence times were lognormally distributed (Fig. 6g) and highly correlated (correlation coefficient [CC] = 0.85). Total hyporheic exchange was 2.8 × 10$^{-3}$ m$^3$s$^{-1}$. The volume of hyporheic exchange was two orders of magnitude smaller than the river discharge over the river length of about 250 m and 1.12 % of the river discharge penetrated through the saturated sediment. Dominant river water infiltration, with infiltration rates up to 4.5 × 10$^{-5}$ m s$^{-1}$, occurred upstream of the geomorphological structures extending across the whole river channel at the upstream side of the riffle, whereas exfiltration mainly occurred at the downstream sides of the
geomorphological structures (Fig. 6d). At $3.5 \times 10^{-5}$ m s$^{-1}$, the highest exfiltration rates appeared at the deep pool downstream of the point bar and the riffle, where the effect of hyporheic exfiltration was increased by the lateral inflow of riparian groundwater.

For moderate river discharges, the direction of the hyporheic flow paths substantially changed in comparison to the low river discharges (Fig. 6b). Median flow path length was 8.3 m, maximum depth was 2.9 m and the median residence time was 7.8 h. Residence times of the hyporheic flow paths were bimodally distributed, with two comparably shaped modes peaking at 1.5 h and at 5.1 h (Fig. 6h). The shift from the lognormal towards the bimodal distribution reflected the changing relation between hyporheic exchange flux at the in-stream gravel bar and around the point bar. Flow path lengths and residence times were highly correlated (CC = 0.9). Total hyporheic exchange was $2.4 \times 10^{-3}$ m$^3$ s$^{-1}$, thus slightly decreased in comparison to low river discharges (- 14%). Hyporheic exchange comprised only a fraction of moderate river discharges (0.19 %) over the river length of about 250 m. Dominant river water infiltration occurred upstream of the in-stream gravel bar and the point bar, whereas infiltration upstream of the riffle diminished for moderate river discharges (Fig. 6e).

For high river discharges, the dominant riverbed structures submerged, and the hyporheic exchange almost disappeared. Total hyporheic exchange was $9.12 \times 10^{-5}$ m$^3$ s$^{-1}$, several orders of magnitude smaller than the river discharge, but larger than for smaller river discharges. Median flow path length was 7.3 m, maximum depth was 1.6 m and the median residence time was 4.8 h. Flow path lengths and residence times were highly correlated (CC = 0.8). With increasing river discharge, the bulk of losing flow paths at both riverbanks increased (Fig. 6c). At $1.8 \times 10^{-5}$ m s$^{-1}$, the highest infiltration rates towards the riparian groundwater occurred at the southern riverbank. The central parts of the river channel were characterized by comparably low infiltration rates (Fig. 6f).

Most of the infiltrated river water outside the characteristic hyporheic flow cells left the domain via the lateral boundaries of the model domain (Fig. 6a-c), which is the result of the river reach characterized by losing conditions. The overall losing flux increased with increasing river discharge, from $0.8 \times 10^{-3}$ m$^3$ s$^{-1}$ (low) to $3.3 \times 10^{-3}$ m$^3$ s$^{-1}$ (high) (+ 312 %). Simultaneously, the hyporheic exchange flux decreased. In combination the net water exchange remained relatively constant, ranging between $3.6 \times 10^{-6}$ m$^3$ s$^{-1}$ and $3.3 \times 10^{-6}$ m$^3$ s$^{-1}$ (- 8 %) for the river discharges presented.
Simulated median hyporheic residence times at the river reach varied inversely with the river discharge. The hyporheic exchange developed independently of the vertical hydraulic gradient (losing condition) and remained persistent over time until the geomorphological structure became completely submerged (details about the relation between river discharge, hydraulic gradient and hyporheic residence time are provided as supporting information S3).

3.5. Heat Transport and Subsurface Temperature Patterns

Advective and conductive heat flux across the river groundwater interface follow distinct daily and seasonal cycles (Fig. 8). Advective heat transport into the streambed is dominated by the river temperatures, highest in summer and lowest in winter. The conductive heat flux was dominated by the temperature gradient with depth, positive (into the bed) when river temperature was higher than groundwater temperature and slightly negative when river temperature was lower than groundwater temperature. Generally the magnitude of advective heat flux was higher than the conductive one (Fig. 8). Subsurface temperature pattern in the shallow streambed are influenced by the advective heat fluxes along the hyporheic flow path and conductive heat fluxes in vertical direction (for details about the spatial distribution of the ratio between advection and conduction are provided in the supporting information S4).

On May 20th, characteristic for the spring-summer period, the riverbed at hyporheic and losing flow paths experienced larger daily temperature amplitudes and higher average daily temperatures originating from infiltrating river water than the gaining locations (Fig. 7a, b). This yields an overall, high spatial variability of the daily amplitudes and average daily temperatures in the river bed (Fig. 7d, e). Below a depth of ~1 m, temperatures were almost independent of the diurnal fluctuations at the land surface, and the amplitudes diminished.

The mean temperatures in the river (15 °C) were higher than the mean groundwater temperatures (9.5 °C). With increasing depth, the sediment temperature decreased to 13.1 °C at 1 m below the riverbed and thus was almost 2 °C lower than the river temperature (Fig. 7e). The depth of the 13°C isotherm substantially varied in space (Fig. 7b), relative to the geomorphological structures (around 4.2 m at the river infiltration sites, 2.7 m at the instream gravel bar and only 0.3 m at the discharge location at the point bar/riffle). The subsurface temperature variance increased with increasing depth from 0.05 °C at 0.2 m to 1.3 °C at 1 m below the riverbed (Fig. 7e).
On November 22nd, characteristic for the autumn-winter period, the river temperature was 3.5 °C only (Fig. 7c). Generally, the hyporheic and losing zones of the river channel experienced lower average daily temperatures than the gaining parts of the river channel. With increasing depth, the sediment temperature significantly increased to 9.2 °C at 1 m below the riverbed and thus was about 5.7 °C higher than the river temperature (Fig. 7f). The subsurface temperature variance increased with increasing depth, from 0.2 °C at 0.2 m to 3.2 °C at 1 m below the riverbed (Fig. 7f).

Driven by the seasonal and daily temperature dynamics in the river, the average hyporheic temperature varied on seasonal and daily time scales (Fig. 8). Generally, the average temperature of the hyporheic zone strongly depends on the river temperatures (CC = 0.99), whereas the variance in hyporheic flow path temperatures was uncorrelated with daily temperature amplitudes in the river (CC = 0.14, p-value = 0.5). The maximum annual temperature difference was 19.2 °C; with a maximum average hyporheic temperature of 19.4 °C (1 July 2012) and a minimum average hyporheic temperature of 0.16 °C (26 February 2013).

We analysed model results in order to find a relationship between the hyporheic residence times, depth and average hyporheic flow path temperature (Fig. 9). In autumn-winter, the average temperature of slow and long flow paths tends to be higher compared to quick and short flow paths (Fig. 9a), especially for the very long residence times. On opposite in spring-summer, the average temperature of long flow paths was lower compared to short ones (Fig. 9b). At flow paths with long residence times, temperatures tend to have higher standard deviations than those with short residence times. The standard deviations ranged between 0 and ~2 °C for residence times of 1 hour and increased to 0 and ~3 °C for residence times of 100 hours (Fig. S5). The distribution between flow path temperatures (\( T_{\text{flow path}} \)) and residence times (\( t_{\text{res}} \)) follow a power law relationship with exponent of 0.37 (\( R^2 = 0.56, t_{\text{res}} > 0.37 \) h) (Fig. 9c):

\[
\log(T_{\text{norm}}) = -1.06 + 0.3684 \times \log(t_{\text{res}}) \quad (4a)
\]

\[
T_{\text{flow path}} = T_{\text{river}} + 0.0864 \times t_{\text{res}}^{0.3684} \times (T_{\text{GW}} - T_{\text{river}}) \quad (4b)
\]

Including hyporheic flow paths with residence times smaller than 0.37 h would decrease the accuracy (\( R^2 \)) of the derived empirical model; the model is not valid for shallow flow paths where flow path temperature is additionally affected by the atmospheric energy input. Here the atmospheric heat input additionally influences the shallow hyporheic flow path...
temperatures by heat conduction through the unsaturated streambed structures. Consequently the shallow hyporheic temperatures around non-submerged geomorphological streambed structures were cooled in winter (Fig. 9a) and warmed in summer (Fig. 9b) as a result of advective and conductive heat transport processes at the interface between atmosphere, streambed and groundwater. For such flow paths $T_{\text{norm}}$ becomes smaller than zero. Details about the linear relation between penetration depth of hyporheic flow path and $T_{\text{norm}}$ are provided in the supporting information (Fig. S6). The linear regression model between normalized flow path temperature and flow path depth ($R^2 = 0.48$) is less accurate than the nonlinear regression between normalized flow path temperature and hyporheic residence times ($R^2 = 0.56$) and thus describes less of the observed variance in subsurface temperature pattern compared to hyporheic residence times.

To further include the depth effect on estimates of hyporheic flow path temperature we extended our analyses to a multiple nonlinear regression between flow path residence time, penetration depth and flow path temperature (equation 5). The $R^2$ of the regression model further increases to 0.69.

$$T_{\text{norm}} = 0.04 \times t_{\text{res}}^{0.31} + 0.097 \times \text{depth}$$

(5)

3.6. Implications of Hyporheic Residence Time and Temperature for Biogeochemical Processes in the Streambed

Advective water and associated heat flux within the shallow riverbed created unique environments with potential implications for the spatial and temporal dynamics of biogeochemical processes and related reaction kinetics. This will be reflected in the oxygen concentrations along the flow paths. Taking the river temperature as proxy for the usually unknown hyporheic flow path temperature will imply some error. We can evaluate that error, based on the corrected hyporheic flow path temperature according to the empirical equation 4b. Hyporheic residence time was found to describe most of the observed variance in subsurface temperature pattern and was therefore used for temperature correction for further analysis. The difference between empirical oxygen consumption calculated via river temperature and via empirical flow path temperature is presented in figure 10. In general, longer residence times lead to higher absolute oxygen consumption (Zarnetske et al. 2011; Marzadri et al. 2012). Our results demonstrate that oxygen consumption was substantially
increased in autumn-winter (flow paths with long residence times are potentially warmer than short ones) and decreased in spring-summer (flow paths with long residence times are potentially cooler than short ones), when the effect of hyporheic flow path temperature on solute transformation is regarded (Fig. 10). For example, oxygen consumption in flow paths with residence times of about 37.7 h (median residence time at the Selke river reach for low discharges) and river temperatures of 2.6 °C was found to be 29 % higher when accounting for residence time-dependent temperature relationships.

4. Discussion

4.1. Hydraulic Conductivity and Thermal Diffusivity

Previous modelling studies have revealed the need for a heterogeneous representation of riverbed hydraulic conductivities to accurately simulate river-groundwater exchange, hyporheic flow path and residence times [Naranjo et al., 2012; Karan et al., 2014]. Especially processes within the hyporheic zone are controlled by residence times of river water within the sediment, which strongly depends on heterogeneity in the riverbed hydraulic conductivity [Salehin et al., 2004; Sawyer and Cardenas, 2009; Tonina et al., 2016]. We found that the riverbed hydraulic conductivity depends on the relative position around the geomorphological structures. The lower hydraulic conductivity at the head of the gravel and point bar could be caused by continuous river water infiltration into the riverbed, transporting organic matter and fine sediments. These materials potentially accumulate in the sediment interstices within the river water infiltration zones (clogging), decreasing the hydraulic conductivity of the riverbed [Brunke and Gonser, 1997; Packman and Mackay, 2003]. The calibrated $K_h$ was 1.6 times larger than the measured values, demonstrating the relevance of parameter calibration even if there is a profound experimental data base. Furthermore, although the Hvorslev falling head slug test is considered a robust method for measuring $K_h$ in shallow riverbed sediments [Landon et al., 2001], there could be a slight bias for layered media, e.g. by a vertical flow component in the slug tests.

The observed hydraulic conductivities of the streambed sediments slightly decreased with depth (Table 3) which has been reported by other studies [Landon et al., 2001; Song et al., 2007]. The hydraulic conductivity of unconsolidated sediment tends to decrease with depth as a result of sediment compaction caused by increasing overburden pressures [Helm, 1976].
Decreasing hydraulic conductivity within the alluvial deposits below the streambed may limit the penetration depth of river-groundwater exchange fluxes.

The anisotropy ratio of 17 for the horizontal to vertical hydraulic conductivity found in this study is rather high, but plausible and within the range of values reported by Trauth et al. [2015] for a smaller-scale numerical model representing the in-stream gravel bar. The anisotropy within the riverbed can be explained by the strong preferential orientation of ellipsoidal, flat gravels and cobble stones in the direction of flow (imbrication), which has been observed at the streambed surface and in freeze cores of the streambed sediments. Comparable anisotropy ratios have been observed for alluvial aquifers [e.g. Chen and Chen, 2003].

The thermal diffusivity determined locally from diurnal temperature signals were at the high end of the expected range presented by Lapham [1989] for saturated streambed sediments ($2.3 \times 10^{-7}$ m$^2$ s$^{-1}$ to $1.5 \times 10^{-6}$ m$^2$ s$^{-1}$). However, the estimated values were in line with experimentally derived streambed thermal diffusivities in sandy and gravelly sediments [Munz et al., 2011a; Rau et al., 2012; Rosenberry et al., 2017].

4.2. Model Setup and Model Validation

The presented model setup goes beyond existing modelling efforts of river reaches, as the heat transport is driven by the use of atmospheric input routines allowing the simulation of river temperatures. Often, the surface heat boundary in models has been defined as fixed temperatures at the top nodes of the modelling domain [e.g. Bartsch et al., 2014; Karan et al., 2014]. Such settings do not account for the potential influence of groundwater discharge on the streambed interface and stream temperatures, which might be reasonable for a river under losing conditions but not for gaining conditions, especially when the temperature gradient between river and groundwater is high.

The calibrated model fit the observed hydraulic heads with the exception of a single observation point, ML_5, where the simulated hydraulic head was systematically overestimated (MAE = 0.2). The hydraulic head at this location was sensitive to changes in riverbed friction and hydraulic conductivity. That indicates that the mismatch was caused by strong heterogeneities in the riverbed structure or in the hydrogeology of the south-eastern part of the river banks, which seems not to have been resolved with sufficient detail.
The highest deviations in maximum and minimum annual temperatures were found in groundwater where simulated temperature peaks were generally lower than the observed ones. The mismatch in temperature variations could be caused by an underestimation of river-groundwater exchange at the river bank (bank filtration), limiting the advective heat exchange along the losing flow paths. On the other hand, the missing temperature dynamics could also be related to the definition of the lateral temperature boundaries in the model, as the deep subsurface flow field was characterized by groundwater entering the domain via the lateral boundary. An underestimation in temperature dynamics at the lateral boundaries would therefore result in an underestimation of the groundwater temperature dynamics at the observation wells.

In the region of interest, the shallow zone of the saturated sediment where hyporheic exchange dominates, the simulated temperatures matched the observed temperature very well, highlighting that the modelled magnitude and direction of exchange fluxes as well as the temperature dynamics accurately represents the condition within the river reach. Thus, the coupled water flow and heat transport model for this Selke river reach provides a sound basis for quantitatively investigating hyporheic exchange properties.

4.3. Implications of Coupled Water Flow and Heat Transport Simulation for River-Groundwater Exchange

For reach-scale applications, three-dimensional numerical modelling in conjunction with point measurements has been used earlier to quantify the exchange flux across the river-sediment interface [e.g. Storey et al., 2003; Wondzell et al., 2009; Munz et al., 2011b]. While the exchange flux is the main variable of interest (besides the state variables itself), model calibration is based on the simulated hydraulic heads. Previous studies have shown that model calibration and validation based on observed hydraulic heads was not sufficient to constrain the model parameters sufficiently [Bravo et al., 2002; Naranjo et al., 2012]. Thus, temperature data have been used to better constrain the estimation of subsurface parameters in steady-state groundwater models and to support calibration of hydraulic conductivities [Doussan et al., 1994; Karan et al., 2014]. The assimilation of temperature data was also reported to lead to a better characterization of the spatial distribution of leakage parameters [Kurtz et al., 2014].
Going beyond steady-state, in the coupled water flow and heat transport simulations that progressed to transient conditions, the hydraulic head turned out to be insensitive to parameter variations in mean, variance and range of the saturated hydraulic conductivity (Fig. 2). The hydraulic gradient between groundwater and river is highly restricted by the lateral model boundary and the surface inflow and outflow boundaries in combination with the riverbed roughness determining the hydraulic head in the surface domain. Due to this assumed physical restriction on changes of heads the changes in hydraulic conductivity should directly alter the river-groundwater exchange flux; and this is indeed what the simulations yield, including that the hydraulic head patterns are actually not changing and river-groundwater exchange fluxes directly respond to variations in hydraulic conductivity (Fig. 2). Simulated hydraulic head was sensitive to changes in the coupling length. Thus aiming toward a correct representation of simulated hydraulic head requires a calibration of $l_c$. In turn it would not be sufficient to solely define $l_c$ for the initial model targeting a constant exchange flux between river and groundwater.

Simulated temperatures are highly sensitive to variations in mean, variance, and range of the saturated hydraulic conductivity. The change in subsurface temperature corresponds directly to changes in advective heat transport (river-groundwater exchange flux) caused by variations in average hydraulic conductivity or its spatial distribution. These results highlight that heat is an additional quantity which helps to constrain highly parameterized, fully-integrated models commonly affected by parameter non-uniqueness. Besides the improvements in model calibration (reduction in head and temperature residuals), we aimed to quantify the benefits of temperature on simulated water flux across the river-groundwater interface. A substantial change in the mean of the simulated exchange flux was achieved (~46 %) by including the temperature in addition to the hydraulic head information in order to constrain the estimation of the subsurface parameters of transient, fully-integrated numerical models.

4.4. Water Flux across the River-Groundwater Interface and Hyporheic Exchange

Gravel bars and fluvial islands are common in-stream geomorphic feature in a large range of natural and regulated river channels. The fully-integrated 3-D modelling approach reveals how a typical lowland river reach and groundwater are dynamically connected. The ratio of hyporheic exchange to river discharge decreased for the considered flow conditions. With increasing discharge, water level differences (hydraulic head gradients) across the
geomorphological structures decrease resulting in lower hyporheic exchange. The head difference across the non-submerged geomorphological structures was highest for low discharges because the low discharges forced river heads to follow riverbed morphology more closely. Thus, hyporheic exchange at the reach-scale is highly variable in space and time and strongly depends on river discharge. In contrast, the net water flux across the river-groundwater interface remained relatively constant. The portion of hyporheic versus losing flow paths changed in dependence on river discharge. Also, hyporheic residence times and penetration depths of the partly submerged geomorphological structures varied inversely with the river discharge, as observed also in other studies [Shope et al., 2012; Trauth et al., 2015]. The knowledge of hyporheic residence time distributions and its moments is vital information for assessing hyporheic solute turn-over of dissolved substances in the hyporheic zone [Marzadri et al., 2011; Briggs et al., 2013].

In our simulations the subsurface residence times were short relative to the transition of discharge. Temporal variation of low and medium river discharges is limited to only 3-8 % during the corresponding time period of mean residence times. Only for high discharges the average hyporheic residence times were particularly long compared to transient changes in river discharge. Thus, the approximation of taking a steady flow field for particle tracking calculation introduces some ambiguity for high discharge situations only. An extension of the model domain towards the floodplain and the groundwater would yield some longer flow paths with longer residence times, though with a low overall contribution. The longest flow paths already included were almost at the groundwater average temperature of 9 °C. A further increase in flow path length (i.e. also residence time) could not alter much further that flow path temperature. Here a model would reach a limit where flow path temperature becomes independent of the flow path residence time (Fig. 9). This relation does not reflect shallow hyporheic flow path through low permeability zones caused by streambed heterogeneities creating potentially important localized pockets of anomalously long residence time (Briggs et al., 2015). These flow paths should be strongly impacted by the surface water temperature (through conduction) independent of the flow path residence time. However, such local heterogeneities were not represented by our empirical approach which was derived from hyporheic exchange through highly permeable streambed sediments.

4.5. Heat Transport and Subsurface Temperature Patterns
A strong variability in daily temperature patterns within riverbed sediments arises due to hyporheic exchange beneath dune-like triangular bedforms [Cardenas et al., 2007; Norman and Cardenas, 2014], riffles [Storey et al., 2003; Marzadri et al., 2013], as well as beneath weirs and large woody debris [Hester et al., 2009; Sawyer et al., 2012]. Our results show how hyporheic up- and downwelling induced by non-submerged geomorphological structures drive substantial thermal heterogeneity within riverbed sediments at the reach-scale. Heat was moving from the river into the saturated sediment along the hyporheic and losing flow paths, driving the temperature dynamics in the shallow riverbed sediments.

The average temperature in the hyporheic zone follows the temperature in the river (characterized by distinct annual and daily cycles), but along individual hyporheic flow paths, temperatures vary substantially around the average hyporheic temperature. The average hyporheic flow path temperature ranges between the atmospheric temperature and the yearly average groundwater temperature and strongly depends on the flow path residence time (flow path length) and penetration depth. Average river temperature is a good predictor for the average hyporheic temperature and is proposed to be suitable to predict solute turn-over of dissolved substances in the hyporheic zone at the morphological unit (Marzadri et al., 2013). However complex hyporheic temperature patterns contribute to highly variable biogeochemically active zones. We found that hyporheic flow path temperature could be estimated more precisely by the presented regression models derived from the detailed numerical simulations (Equations 4, 5) that accounts for hyporheic residence time, penetration depth, river and groundwater temperature.

Changes in surface water temperature, large and rapid particularly in small streams, penetrate into the saturated sediment by advection and conduction. Thereby the temperature signal is attenuated and delayed along the flow path in the streambed sediments and the thermal front velocity is retarded along hyporheic flow path compared to the pore water velocity. Our analyses of hyporheic flow path residence times (flow path length) and penetration depth reveal that the residence time strongly controls hyporheic flow path temperatures. They demonstrate that substantial hyporheic temperature differences in small streams also occur vertically and not only horizontally due to relatively high advective heat fluxes.

### 4.6. Implications of Hyporheic Residence Time and Temperature for Biogeochemical Processes in the Streambed

This article is protected by copyright. All rights reserved.
Microbially mediated solute transformation in the hyporheic zone largely contribute to the whole stream metabolism, particularly in small streams as the one in this study; e.g., hyporheic biofilms contribute to total stream respiration of 76–96 % [Naegeli and Uehlinger, 1997] or 40–93 % [Fellows et al., 2001] Biogeochemical reactions associated with stream oxygen and nitrogen cycling, such as nitrification and denitrification, can be strongly controlled by water and solute residence times and by daily and seasonal temperature dynamics in the hyporheic zone [Kaplan and Bott, 1989; Haliday et al., 2013], since the higher the temperatures, the higher the rates for microbially mediated solute transformation [Thamdrup and Fleischer, 1998; Acuna et al., 2008]. Temperature variations have a stronger impact in small and steep streams compared to large low gradient streams because of the shorter residence times [Marzadri et al., 2012].

Changes in flow path residence time in addition to changes in flow path temperature strongly affect the extent of the oxygenated flow path in the streambed sediments. Discharge events at the Selke drastically reduced the median residence time of 37.7 hours to less than 4.8 hours affecting the ratio of characteristic hydrologic transport time scale to oxygen uptake time scale. Consequently, short term changes in the oxygen availability in the streambed sediments are attributed to fluctuations in residence time (fluctuations in river discharge), while more gradual changes in oxygen availability are due to seasonal variations of the river water temperature due to a temperature dependent first-order respiration rate coefficient.

For example, low temperatures during winter associated with moderate residence times, favour residual concentrations of dissolved oxygen in hyporheic sediments. Low temperatures decrease the reaction rate coefficient and consequently the fraction of oxygen consumed along the flow path, increasing the extent of the oxygenated flow path in the streambed sediments. To accurately account for the seasonal variations in oxygen consumption along hyporheic flow path the average hyporheic flow path temperature needs to be corrected for hyporheic residence time and potentially also for hyporheic penetration depth. Here, we focus the discussion on the correction for hyporheic residence times as they are directly linked to the fraction of consumed oxygen along hyporheic flow paths (see equation 2). Furthermore, Harvey et al., [2013] found that effective penetration depth contribute insignificantly to reach-scale solute turn-over of dissolved substances in the hyporheic zone.

The effect of temperature correction is most obvious in winter when hyporheic temperatures are low and the extent of the oxygenated zone is comparatively high because of the prevailing anaerobic conditions along the streamlines. For such circumstances (e.g. $T = 2.6 \, ^\circ C$ and $t_{res} =$
37.7 h) oxygen consumption with corrected temperature effect is about 29 % higher compared to the oxygen consumption without temperature correction. If temperature is assumed to be constant a comparable difference in oxygen consumption of about 29 % could be achieved by an increase in the residence times from 37.7 h to 59 hours (Figure 10b). With increasing temperatures the effect of the temperature correction diminishes as the transport time scale becomes the dominant factor in respect to biological oxygen uptake time. The infiltrated oxygen is completely consumed quickly and thus the extent of the oxygenated zone is independent to corrections in hyporheic temperatures. Figure 10a illustrates the range of flow path residence times and temperatures where a temperature correction has a substantial effect on oxygen consumption and thus also on the extent of the oxygenated flow path in the streambed sediments.

Despite a continuously increasing interest in solute turn-over of dissolved substances in the hyporheic zone a detailed analysis of the temporal biogeochemical dynamics, linking dynamic flow processes with the effects of varying temperatures, reaction kinetics and the resulting overall biogeochemical turn-over rates at the river groundwater interface is still missing. In this study we derived an empirical relation which provides a strong basis to better predict hyporheic temperature dynamics in relation to hyporheic residence time, river and groundwater temperatures. We furthermore highlight the relevance to accurately estimate hyporheic temperatures as they have a substantial effect on solute turn-over of dissolved substances. Along hyporheic flow paths of infiltrating river water the terminal electron acceptors, such as dissolved oxygen and nitrate, are sequentially consumed. Thus, variations in oxygen consumption influence the development of distinct redox zones and thus also the potential of hyporheic zones to promote net denitrification in the hyporheic zone and turn-over of other chemical species.

5. Summary and Conclusions

The goal of this study was to investigate the factors dominating water and heat fluxes across the river-groundwater interface and hyporheic temperature patterns along a natural lowland river reach and to aim for a better quantitative basis to account for their effect on temperature-sensitive microbial processes. We used the fully-integrated, surface-subsurface flow and heat transport model HydroGeoSphere in combination with particle tracking techniques in order to simulate these processes, based on continuous measurements of hydraulic heads and
temperatures at different depths in the river bank and within the river over a two-year period from May 2011 until June 2013. Based on the field investigations and the simulation results, the following conclusions can be drawn:

1. Horizontal hydraulic conductivity depends on the location relative to the geomorphological structures, and is lowest at the head and increasing towards the center and tail.

2. The most sensitive parameters affecting the groundwater and hyporheic temperatures—in addition to the properties characterizing the distribution of the hydraulic conductivity (mean, variance and anisotropy) are the parameters related to the heat transport itself, i.e. thermal conductivity, bulk density and heat capacity.

3. Including the temperature in addition to the hydraulic head information constrains the estimation of subsurface parameters and decreases uncertainty in simulated exchange flux substantially. This leads to more accurate estimates of the flux between stream and groundwater systems.

4. Non-submerged streambed structures drive substantial thermal heterogeneity within riverbed sediments at the reach-scale. Non-submerged streambed structures influence shallow hyporheic flow path temperatures as a result of advective and conductive heat transport processes at the interface between atmosphere, streambed and groundwater. The average hyporheic temperature is controlled by the river temperature. Individual mean hyporheic flow path temperature strongly depends on the flow path residence time, and river and groundwater temperature. In autumn-winter, the average temperature of long flow paths is potentially higher compared to the ones associated with short flow paths, whereas in spring-summer the average temperature of longer flow paths is lower compared to shorter ones. The distribution between flow path temperatures and residence times follow a power law relationship, derived by reach-scale simulations calibrated to field data.

5. Reach-scale hyporheic oxygen consumption can be adjusted for hyporheic residence times and average flow path temperature.

Our analysis highlights the hyporheic temperature variations in space and time and demonstrate how they relate to river temperature, groundwater temperature and hyporheic residence times. Understanding these links provides the basis from which to assess hyporheic temperature conditions in river reaches. Since biogeochemical processes depend on solute
residence time and temperature, an understanding of flow and temperature regimes in the riverbed sediments is essential to quantify the reactive potential of hyporheic zones.

The 3-D, transient water flow and heat transport simulation applying the HydroGeoSphere code is furthermore a sound basis from which to investigate water and heat fluxes across the river-groundwater interface and hyporheic temperature patterns, only relying on measurements of discharge, groundwater heads and climate conditions and could be particularly useful for extreme discharges when direct measurements at the river-groundwater interface are difficult to conduct.

Acknowledgements

The first author acknowledges the Helmholtz graduate research school GeoSim for funding the PhD studentship. We thank Martin Briggs and two anonymous reviewers for their constructive feedback on the manuscript. Special thanks to: N. Trauth, M. Vieweg, T. Keller for supporting the field work and G. Baroni for his helpful suggestions on parameter sensitivity analysis. This study was supported by TERENO (Terrestrial Environmental Observatories). The data of the field site and the numerical model can be found in the supporting information.
References


Chen, X. H., and X. Chen (2003), Effects of aquifer anisotropy on the migration of infiltrated stream water to a pumping well, J Hydrol Eng, 8(5), 287-293.


Marzadri, A., D. Tonina, and A. Bellin (2012), Morphodynamic controls on redox conditions and on nitrogen dynamics within the hyporheic zone: Application to gravel bed rivers with alternate-bar morphology, J Geophys Res-Biogeo, 117.


Munz, M., S. E. Oswald, and C. Schmidt (2011a), Sand box experiments to evaluate the influence of subsurface temperature probe design on temperature based water flux calculation, Hydrol Earth Syst Sc, 15(11), 3495-3510.

Munz, M., S. E. Oswald, and C. Schmidt (2016), Analysis of riverbed temperatures to determine the geometry of subsurface water flow around in-stream geomorphological structures, J Hydrol, 539, 74-87.


Nutzmann, G., C. Levers, and J. Lewandowski (2014), Coupled groundwater flow and heat transport simulation for estimating transient aquifer-stream exchange at the lowland River Spree (Germany), Hydrol Process, 28(13), 4078-4090.


Table 1. Flow and transport parameters used for HydroGeoSphere simulations. The min and max values define the parameter space used for global sensitivity analysis and model calibration.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Calibrated</th>
<th>Min</th>
<th>Max</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hydraulic/Hydrodynamic – Porous Media</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porosity</td>
<td>θ</td>
<td>0.3</td>
<td>0.2</td>
<td>0.4</td>
<td>(-)</td>
</tr>
<tr>
<td>Hydraulic conductivity (mean)</td>
<td>K&lt;sub&gt;H&lt;/sub&gt; mean</td>
<td>2.30 × 10&lt;sup&gt;-3&lt;/sup&gt;</td>
<td>0.1&lt;sup&gt;1&lt;/sup&gt;</td>
<td>10&lt;sup&gt;2&lt;/sup&gt;</td>
<td>m s&lt;sup&gt;-1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hydraulic conductivity (var)</td>
<td>K&lt;sub&gt;H&lt;/sub&gt; variance</td>
<td>3.00 × 10&lt;sup&gt;-6&lt;/sup&gt;</td>
<td>0.5&lt;sup&gt;2&lt;/sup&gt;</td>
<td>2&lt;sup&gt;2&lt;/sup&gt;</td>
<td>m s&lt;sup&gt;-1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hydraulic conductivity (range x/y direction)</td>
<td>K&lt;sub&gt;H&lt;/sub&gt; Range</td>
<td>23.8 / 5.8</td>
<td>1 / 1</td>
<td>50 / 15</td>
<td>m</td>
</tr>
<tr>
<td>Anisotropy</td>
<td>K&lt;sub&gt;vv&lt;/sub&gt;/K&lt;sub&gt;vv&lt;/sub&gt;</td>
<td>17</td>
<td>1</td>
<td>100</td>
<td>(-)</td>
</tr>
<tr>
<td>Specific Storage</td>
<td>S&lt;sub&gt;S&lt;/sub&gt;</td>
<td>10&lt;sup&gt;-4&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>m&lt;sup&gt;-1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Longitudinal/transversal dispersivity</td>
<td>α&lt;sub&gt;L&lt;/sub&gt;</td>
<td>0.1 / 0.01</td>
<td>-</td>
<td>-</td>
<td>m&lt;sup&gt;2&lt;/sup&gt; s&lt;sup&gt;-1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Effective diffusion coefficient</td>
<td>D</td>
<td>10&lt;sup&gt;-10&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>m&lt;sup&gt;2&lt;/sup&gt; s&lt;sup&gt;-1&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Overland flow (Channel/Floodplain)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friction (roughness)</td>
<td>N</td>
<td>0.022</td>
<td>0.01</td>
<td>0.05</td>
<td>s m&lt;sup&gt;1/3&lt;/sup&gt;</td>
</tr>
<tr>
<td>Coupling length</td>
<td>l&lt;sub&gt;c&lt;/sub&gt;</td>
<td>3.5 × 10&lt;sup&gt;-4&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-5&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-2&lt;/sup&gt;</td>
<td>(-)</td>
</tr>
<tr>
<td><strong>Thermal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal conductivity of the sediment</td>
<td>λ&lt;sub&gt;s&lt;/sub&gt;</td>
<td>1.2</td>
<td>1</td>
<td>5</td>
<td>W m&lt;sup&gt;-1&lt;/sup&gt; °C&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Thermal conductivity of water</td>
<td>λ&lt;sub&gt;W&lt;/sub&gt;</td>
<td>0.7</td>
<td>-</td>
<td>-</td>
<td>W m&lt;sup&gt;-1&lt;/sup&gt; °C&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Heat capacity of the sediment</td>
<td>C&lt;sub&gt;s&lt;/sub&gt;</td>
<td>1330</td>
<td>1000</td>
<td>2500</td>
<td>J kg&lt;sup&gt;-1&lt;/sup&gt; °C&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Heat capacity of water</td>
<td>C&lt;sub&gt;W&lt;/sub&gt;</td>
<td>4185</td>
<td>-</td>
<td>-</td>
<td>J kg&lt;sup&gt;-1&lt;/sup&gt; °C&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bulk density</td>
<td>ρ&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1350</td>
<td>1200</td>
<td>2400</td>
<td>kg m&lt;sup&gt;-3&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ground surface albedo (dry/saturated)</td>
<td>α&lt;sub&gt;g&lt;/sub&gt;</td>
<td>0.2 / 0.07</td>
<td>-</td>
<td>-</td>
<td>(-)</td>
</tr>
</tbody>
</table>

<sup>1</sup> maximum parameter space, partitioned into 5 equidistant levels for the sensitivity analyses
<sup>2</sup> factor multiplied by the initial value

Table 2. Analyses of riverbed hydraulic conductivity variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Groups (degree of freedom)</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>2011, 2012, 2013</td>
<td>3.68</td>
<td>1.84</td>
<td>0.071</td>
</tr>
<tr>
<td>Depth (m)</td>
<td>0.0-0.2, &gt;0.2 - 0.4, &gt;0.4 - 0.8</td>
<td>0.97</td>
<td>0.24</td>
<td>0.841</td>
</tr>
<tr>
<td>Location (reach)</td>
<td>In-stream gravel bar, Point bar, River</td>
<td>1.86</td>
<td>1.86</td>
<td>0.101</td>
</tr>
<tr>
<td>Location (structure)</td>
<td>Head, Crest, Tail</td>
<td>6.57</td>
<td>3.28</td>
<td>0.009*</td>
</tr>
<tr>
<td>Total</td>
<td>231</td>
<td>191.39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>*</sup> significant at the level p < 0.05

Table 3. Measured riverbed and aquifer hydraulic conductivities

<table>
<thead>
<tr>
<th>Location (depth below surface)</th>
<th>Average hydraulic conductivity</th>
<th>Model Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riverbed (&lt; 0.80 m)</td>
<td>1.40 × 10&lt;sup&gt;-5&lt;/sup&gt; m s&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>1-14</td>
</tr>
<tr>
<td>Shallow (1.70 m)</td>
<td>1.06 × 10&lt;sup&gt;-3&lt;/sup&gt; m s&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>12,13&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
<tr>
<td>Middle (2.70 m)</td>
<td>6.93 × 10&lt;sup&gt;-4&lt;/sup&gt; m s&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>14-15&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
<tr>
<td>Deep (3.70 m)</td>
<td>2.30 × 10&lt;sup&gt;-4&lt;/sup&gt; m s&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>16 - 27&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bottom (9.00 m)</td>
<td>2.30 × 10&lt;sup&gt;-8&lt;/sup&gt; m s&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>28&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>1</sup> small hydraulic conductivity at the last model layer represents impermeable underlying sediment (value not measured)
<sup>12</sup> homogeneous for the entire layer
Figure 1. (a) Study site and model domain, showing experimental infrastructure, surface elevation, relevant geomorphological structures, and the finite element mesh (white-framed triangles). The black line indicates the river shape for discharges, here taken of 0.25 m$^3$s$^{-1}$. (b) Cross-section through the model domain (A to B). The lower 6 meters of the model domain are not shown, because they are fully saturated and no observation boreholes extended to that depth range.

Figure 2. Morris plots of $\mu^*$ versus $\sigma$ of the elementary effect of each parameter with respect to (a) river-groundwater exchange flux, (b) hydraulic head, (c) groundwater temperatures, and (d) hyporheic temperatures. See Table 1 for a description of parameter symbols.

Figure 3. (a) Relation between hydraulic head bias and temperature bias, with river-groundwater exchange flux indicated in color, for all sensitivity simulations. (b) Distribution of simulated river-groundwater exchange flux considering the best 10% of all sensitivity simulations regarding the hydraulic head bias, hydraulic head and groundwater temperature bias, and the hydraulic head and groundwater/hyporheic temperature bias.

Figure 4. (a) Observed and simulated hydraulic heads, and (b) observed and simulated temperatures of the calibrated model. Measurement locations of the river (161.14 m a.s.l.) and ML_2 (157.9 m a.s.l.) are indicated in Fig. 1 a.

Figure 5. Measured vs. simulated hydraulic heads (first row), groundwater temperatures (second row) and riverbed temperatures (third row) for the calibrated model. Hydraulic head and temperature time series of the river and ML_2 are shown in Fig. 5. All measurement locations are indicated in Fig. 1 a.

Figure 6. (a-c) Hydraulic heads of the streambed and groundwater streamlines for low, moderate and high river discharges under losing conditions. (d-f) River-groundwater exchange flux (Negative values indicate surface water infiltration to groundwater); the black solid line shows the actual river shape. (g-i) Flow path length (upper) and residence time distributions for hyporheic flow path. The x-axis is scaled logarithmically.

Figure 7. Cross-sections along transect A-B showing (a) daily temperature amplitudes for low discharges ($q = 0.25$ m$^3$/s) and moderate amplitudes ($A = 2$ °C), (b) daily average temperatures for spring-summer (negative temperature gradient between river and groundwater), and (c) autumn-winter (positive temperature gradient between river and groundwater). (d-f) Boxplot of average daily temperature for each streambed node at 0, 0.2 m, 0.4 m 0.6 m, 0.8 m and 1 m depth below the river-groundwater interface.
Figure 8. (a) Distribution of hyporheic flow path temperatures every 6 days from July 2011 until March 2013 compared to simulated average daily temperature for river and groundwater. (b) Distribution of hyporheic flow path temperatures on 18 August 2011 every hour (+ 4 hours of the previous and following day). (c) corresponding conductive and advective heat flux.

Figure 9. Relation between residence time of hyporheic flow paths ($t_{res}$) versus average flow path temperature ($T_{flow\ path}$) for (a) an autumn day (22 November 2011) and (b) a spring day (20 May 2012). The empirical model is presented in equation 4b. (c) Relation between residence time of hyporheic flow paths versus normalized flow path temperature ($T_{norm}$). The fitted linear model is presented in equation 4a. (d) Multiple nonlinear regression model for evaluation of combined effect of flow path residence time and penetration depth on $T_{norm}$. The fitted model is presented in equation 5.

Figure 10. (a) Difference in oxygen consumption ($\Delta O_2$) calculated via river temperature and via empirical flow path temperature (Equation 4) for naturally occurring flow path residence time and river temperatures and (b) Changes in $\Delta O_2$ for increasing residence times from $t_{res} = 37.7 \ h$.  

This article is protected by copyright. All rights reserved.
Figure 2.
a) River-GW flux

b) Hydraulic head

c) Groundwater temperature

d) Hyporheic temperature

This article is protected by copyright. All rights reserved.
Figure 4.
Simulated hydraulic head (m a.s.l)

Water level (2)

MAE = 0.03
NSE = 0.92

ML_2
MAE = 0.06
NSE = 0.86

ML_5
MAE = 0.20

ML_1
MAE = 0.02
NSE = 0.95

ML_4
MAE = 0.03
NSE = 0.93

ML_6
MAE = 0.08
NSE = 0.63

ML_3
MAE = 0.07
NSE = 0.69

ML_1
MAE = 0.02
NSE = 0.95

Simulated temperature (°C)

MAE = 0.36
NSE = 0.99

ML_2
MAE = 0.3
NSE = 1.0

ML_5
MAE = 0.64
NSE = 0.92

ML_1
MAE = 0.53
NSE = 0.97

ML_4
MAE = 0.5
NSE = 0.98

ML_6
MAE = 0.64
NSE = 0.93

ML_3
MAE = 0.57
NSE = 0.92

ML_1
MAE = 0.36
NSE = 0.99

MAE = 0.57
NSE = 0.92

MAE = 0.53
NSE = 0.97

MAE = 0.5
NSE = 0.98

MAE = 0.36
NSE = 0.99

Measuring hydraulic head (m a.s.l.)

Measured temperature (°C)

18 (River)

MAE = 1.75
NSE = 0.73

ML_6
MAE = 0.89
NSE = 0.95

ML_3
MAE = 1.75
NSE = 0.73

ML_5
MAE = 0.3
NSE = 1.0

ML_2
MAE = 0.81
NSE = 0.83

ML_4
MAE = 1.33
NSE = 0.75

ML_1
MAE = 1.01
NSE = 0.70

3 (-0.6 m)

MAE = 0.64
NSE = 0.93

ML_6
MAE = 1.2
NSE = 0.92

ML_3
MAE = 0.57
NSE = 0.92

ML_5
MAE = 0.53
NSE = 0.97

ML_2
MAE = 0.5
NSE = 0.98

ML_4
MAE = 0.36
NSE = 0.99

ML_1
MAE = 0.47
NSE = 0.99
Figure 6.

Accepted Article
Figure 7.
a) Daily amplitude

b) Average daily temperature
c) Average daily temperature

d) Amplitude (°C)

e) Temperature (°C)
f) Temperature (°C)
Figure 8.
Distribution of hyporheic flow path temperatures

Atmospheric River (water level probe 2)

Groundwater (ML_2)

Heat fluxes (W m$^{-2}$)

Advective

Conductive

This article is protected by copyright. All rights reserved.
T = 2.5 °C

T = 9.03 °C

GW

-4 -3 -2 -1 0 1 2 3 4
Residence time (log(hours))

-2 -1 0 1 2 3 4
Flow path temperature (°C)

T = 16.8 °C

T = 2.5 °C

-4 -3 -2 -1 0 1 2 3 4
Residence time (log(hours))

-4 -3 -2 -1 0 1 2 3 4
Residence time (log(hours))

0.8
0.6
0.4
0.2
0
0

Penetration depth (m)

T (°C)

norm

-2 -1 0 1 2 3 4
Residence time (log(hours))

-2 -1 0 1 2 3 4
Residence time (log(hours))

-2 -1 0 1 2 3 4
Residence time (log(hours))

-2 -1 0 1 2 3 4
Residence time (log(hours))

This article is protected by copyright. All rights reserved.
Figure 10.
Anual mean groundwater temperature

$t = 7.8 \ h$

$\tau_{res} = 37.7 \ h$

Relative increase in residence time

Difference in oxygen consumption (%)

T = 0°C

T = 3°C

T = 5°C

T = 10°C

Flow Path Residence Time (log(hours))

Relative increase in residence time