Groundwater flow and mixing in a wetland–stream system: Field study and numerical modeling

Sachin Karan a,*, Peter Engesgaard a, Majken C. Looms a, Troels Laier b, Jolanta Kazmierczak a

aDepartment of Geosciences and Natural Resource Management, Section of Geology, University of Copenhagen, Øster Voldgade 10, 1350 Copenhagen K, Denmark
bGeological Survey of Denmark and Greenland (GEUS), Øster Voldgade 10, 1350 Copenhagen K, Denmark

corresponding author. Tel.: +45 27219333.
E-mail address: sak@geo.ku.dk (S. Karan).

1. Introduction

Stream water quality is highly dependent on the nutrient concentrations in the discharging groundwater converging near the streambed (Winter et al., 1998). Before reaching the stream the fate of nutrients is controlled by the geological and geochemical characteristics of the subsurface (Toth, 1999). The individual flow paths therefore play a significant role in controlling the ecological state of streams (Tesoriero et al., 2000). Thus, to meet the requirements of the Water Framework Directive (WFD) in Europe that focuses on groundwater–surface water quality it is important to understand the controlling factors for nutrient loading in wetlands and streams, as they are the recipients of groundwater discharging to stream valleys. Dahl et al. (2007) proposed a general classification of flow path types through riparian zones near streams in Denmark and evaluated riparian zone denitrification capacities.

However, a typology cannot always account for the complexity in exchange of water and nutrients between a riparian zone and a stream. For example, Cey et al. (1999) showed how nitrate loading to a stream was reduced due to denitrification as a result of transport beneath the adjacent anoxic riparian zone. This investigation and several others (e.g., Brüsch and Nilsson, 1993; Böhlke et al., 2002; Hedin et al., 1998; Kellogg et al., 2005; Vidon and Hill, 2004) have shown that the riparian zones adjacent to streams can act as a zone for nitrate reduction. They also exemplify the need for geological information to understand and predict the impacts of these buffer zones on the nutrient loading to streams. Recent studies have shown that while groundwater fluxes across the streambed may be largest at the center opposite patterns of low nitrogen (N) fluxes at the center can be observed as well (Kennedy et al., 2009). This shows that natural riparian environments may differ substantially from common theoretical perceptions in terms of the complexities in the flow paths through a wetland to the stream, and how mixing and geochemical processes affect nutrient loading.
Typically, discharging groundwater originates from both near and far from the stream (Modica et al., 1998) and the stream acts as a convergence zone for mixing groundwater from different sources with potentially differing age and geochemical composition. Preferential flow paths to streams play an important role since groundwater originating from aerobic surficial aquifers are likely to carry nitrate directly to the wetland–stream system (depending on the land use), while groundwater originating from deeper aquifers tends to be nitrate-free (Dahl et al., 2007). Hedin et al. (1998) observed that the change in the chemical composition of discharging groundwater to a stream was linked to different flow paths. The chemical composition of sampled groundwater in proximity to a stream had elevated NO
\textsubscript{3} concentrations and differed distinctly compared to samples free of nitrate originating from a different source (Hedin et al., 1998). Hence, understanding the subsurface architecture of the riparian zone of wetland is vital in understanding groundwater–stream exchange. There is a need for combining several investigations to more accurately determine groundwater–stream exchanges in order to in the end manage surface water quality.

Electrical resistivity tomography (ERT) is a recognized tool for assessing geological information and has been applied for several years (e.g., Bowling et al., 2007; Dahlin, 2001; Dahlin and Løke, 1998; Griffiths and Barker, 1983; Swarzenski et al., 2007). In recent years some investigations have focused on delineating the geology and heterogeneity within surface-near aquifers and have also showed the applicability of ERT to characterize the subsurface directly underneath streams (Crock et al., 2008; Nyquist et al., 2008). Moreover ERT measurements have also been used for detecting spatio-temporal changes of localized groundwater recharge (Musgrave and Binley, 2011). While ERT, combined with other geophysical methods, have proved its usefulness for interpreting groundwater surface water exchange patterns these have not, to our knowledge, been included directly in numerical groundwater flow models and tested in terms of the applicability of the inferred hydrogeology to simulate observed data.

The main objective of this study is to understand the effects of spatial heterogeneity on groundwater–stream exchange in a wetland system. An approach of combining results from land- (transverse) and stream- (longitudinal) based ERT with synthetic field measurements in a numerical groundwater flow model is evaluated; (1) in a quantitatively manner by guiding the conceptual modeling of the hydrogeology from measured hydraulic conductivities combined with ERT and then calibrating and testing the numerical model against the mean of observed hydraulic heads along the transect with a pressure transducer in place near the screen using the falling head method (Fitts, 2002). Data was analyzed with the Hvorslev method for unconfined aquifers (Hvorslev, 1951). A minimum of three consecutive slugs tests were carried out in piezometers with fast dropping hydraulic heads.

Water samples for major-ion (K\textsuperscript{+}, Na\textsuperscript{+}, Ca\textsuperscript{2+}, Mg\textsuperscript{2+}, Cl\textsuperscript{−}, SO\textsubscript{4}\textsuperscript{2−}, NO\textsubscript{3}, NO\textsubscript{2}) and CFC analysis were collected in multiple depths below the streambed at the northern bank (S1) in November 2010 and in multiple depths in the wetland in June 2011 (S2–S5). A piezometer made of 2.8-cm-diameter galvanized steel pipe with a 9 cm steel screen was pushed down using a pneumatic hammer to the desired depth. Then a peristaltic pump was connected to the piezometer and the system was flushed at a rate of 0.2–0.3 l/min. Water samples were collected when field measurements indicated constant values for pH, dissolved oxygen and EC. Water was collected in 100 ml glass bottles, which were filled from the bottom up. The bottle was placed in a one liter container during sampling and was firmly capped under water when the container had been filled, leaving no gas bubbles in the flask. Subsequently, the piezometer was pushed down and the procedure repeated. Water samples were analyzed for major ions using ion chromatography with a conductivity detector (Dionex LD20/CD20) at the Geological Survey of Denmark and Greenland (GEUS). Water samples for analysis of CFC\textsubscript{13} (CFC-11) and CFC\textsubscript{12} (CFC-12) were also taken to the GEUS laboratory within a week and analyzed according to the procedure described by Busenberg and Plummer (1992) but substituting flame-sealed ampoules with screw cap bottles (see modified method at http://water.usgs.gov/lab/chlorofluoro carbons/sampling/bottles/). The use of screw cap bottles was more convenient during field work. This sampling method has been applied elsewhere (Kidmose et al., 2011), with good results by just keeping the storage time at a minimum. Ages were based on a comparison of CFC equilibrium partial pressures corrected for recharge temperatures with a chronology of atmospheric partial pressures (Busenberg and Plummer, 1992). A range of CFC ages was calculated for each sample assuming recharge temperature from 6 to 10 °C, annual mean temperature being 8 °C.

CFC-ages express the time since water was isolated from the unsaturated zone atmosphere, therefore these are referred to as apparent CFC-ages and regarded as minimum ages (Busenberg and Plummer, 1992). CFCs are often regarded as conservative tracers after reaching the water table, but concentrations can be affected by biodegradation and sorption especially in anoxic environments (Bratton et al., 2004; Cook et al., 1995; Hinsby et al., 2007).
3.2. Land- and stream based ERT

Both land- and in-stream electrical resistivity surveys were conducted with a Wenner array configuration using a SYSCAL Pro resistivity meter (IRIS Instruments, Orleans, France). The land-based ERT survey transverse to the stream was carried out in May 2008 with a 1 m electrode spacing for a distance of 81 m resulting in a maximum penetration depth of ~13 m. Elevation of the electrodes were measured with a Trimble R8 GPS. The longitudinal in-stream ERT was carried out in December 2010 with a 0.5 m spacing for a distance of 35 m resulting in a penetration depth of ~6 m. Data was processed with RES2DINV by Geotomo Software using a least-squares inversion technique to produce a 2D resistivity image. The noise level in the land-based data set was estimated by comparing the original measurements with their corresponding reciprocal measurements (LaBrecque et al., 1996). Unfortunately, no reciprocal measurements were conducted in the in-stream ERT.

3.3. Temperature analysis

Three temperature probes were installed across and below the streambed in 2008 with sensors distributed over ~2.5 m into the streambed (see Jensen and Engesgaard, 2011, for details). Temperatures were monitored every 20 min. from April 2008 to February 2011. Monthly averages were applied in the subsequent analysis. Some data were omitted due to drifting (Stonestrom and Contantz, 2003) caused by water condensation in the connector box and battery failure. The vertical (upward) water fluxes were calculated on the basis of a solution to the one-dimensional convection–dispersion equation of Stallman (1965):

$$\frac{\partial T}{\partial t} = \kappa_e \frac{\partial^2 T}{\partial z^2} - \frac{n \eta f}{\gamma} \frac{\partial T}{\partial z},$$

(1)

where $T$ is temperature (°C), $t$ is time (s), $z$ is depth (m), $\gamma = \rho c / \rho_f c_f$, the ratio of volumetric heat capacity of the streambed ($\rho_c$) to the fluid ($\rho_f$) both in ($J \cdot m^{-3} \cdot °C^{-1}$), $n$ is porosity, $\eta_f$ is average linear velocity ($m \cdot s^{-1}$), for which a negative value indicates upward flow, and $\kappa_e$ is effective thermal diffusivity of the saturated sediment ($m^2 \cdot s^{-1}$) defined as:

$$\kappa_e = \frac{\lambda}{\rho c},$$

(2)

where $\lambda$ is thermal conductivity of the saturated sediment ($J \cdot s^{-1} \cdot m^{-1} \cdot °C^{-1}$). Eq. (1) is solved analytically using the steady-state solution of Bredehoft and Papadopoulos (1965) with a fixed boundary conditions $T = T_0$ for $z = 0$ (temperature at streambed surface) and $T_L$ for $z = L$ where $L$ (m) is the vertical extent of the domain.
where \( q_z \) is vertical Darcy flux (m s\(^{-1}\)) = \( n \cdot v_z \). Darcy fluxes were calculated by fitting Eq. (3) to data in MATLAB with the “fitoption” function. Streambed thermal properties used in the analysis are given in Table 1.

\[
\frac{(T(z) - T_0)}{(T_h - T_0)} = \exp\left(\frac{q_z \rho C_0}{-L} z\right) - 1
\]

\[
\exp\left(\frac{q_z \rho C_1}{-L} L\right) - 1
\]

where \( q_z \) is vertical Darcy flux (m s\(^{-1}\)) = \( n \cdot v_z \). Darcy fluxes were calculated by fitting Eq. (3) to data in MATLAB with the “fitoption” function. Streambed thermal properties used in the analysis are given in Table 1.

3.4. Numerical model

A 2D steady-state MODFLOW (Harbaugh, 2005) model was set up with a cell size of 0.5 by 0.25 m (in the horizontal and vertical direction, respectively). The hydraulic conductivity (K) distribution was based on the geology as interpreted from the ERT and the estimated hydraulic conductivities with the following procedure; (1) The ERT and K data were used to delineate zones with similar properties, (2) for each zone, the average K from slug tests carried out in piezometers screened in this zone was assigned (see Section 4.2), and (3) the unrecorded zones at the flanks of the ERT profile were assigned a hydraulic conductivity equivalent to the mean of the relative distribution of the hydraulic conductivity from the geological layers above:

\[
K = \frac{\sum_K A_i}{\sum A_i}
\]

where \( K \) and \( K_i \) are the hydraulic conductivities of the unrecorded zone and the geological layer \( i \), respectively, and \( A_i \) is the area of geological layer \( i \). Other approaches of assigning a K value to the two zones at the flanks could have been used (e.g. assuming that the different zones in the interior of the model extended out to the flanks). However, since the contrasts in K between the different zones at the flanks are small (a factor close to one), the different approaches would yield approximately the same \( K \) value. An aquifer thickness of 15 m was assumed based on nearby bore hole data (available from the national data base, JUPITER). The bottom of the aquifer was defined as a no-flow boundary and the northern and southern model boundaries together with stream cells were assigned a constant head equal to the average head in P1, and average water level in the creek and stream, respectively (Fig. 1). The simulation/calibration was not sensitive to the anisotropy ratio \( K_h/K_v \) so a value of 3 was applied (Kidmose et al., 2011; Hinsby et al., 2007). The porosity was also unknown and a value of 0.3 representing sand was assumed. Recharge was uniformly distributed at the uppermost cells of the model with a rate of 420 mm yr\(^{-1}\) as precipitation for the catchment was 888 mm yr\(^{-1}\) (Kidmose et al., 2011), while actual evapotranspiration for forest and wetlands in the catchment is 460 and 480 mm yr\(^{-1}\) (Ringgaard et al., 2011). Hydraulic head measurements were used as calibration targets and the model was tested against temperature-based streambed fluxes. MODPATH (Pollock, 1994) simulations were used to investi-

gate the flow paths to the stream by backtracking particles released within the stream and the sampling points.

4. Results and discussion

4.1. Transect measurements

Mean hydraulic heads are shown in Table 2. The gradient from the northern and southern boundary were 0.01 and 0.03, respectively, showing decreasing heads towards the stream. The observed range in hydraulic heads was 0.18–0.9 m with an average of 0.33 m showing some areas within the system to be more susceptible to dynamical effects. Slug tests were performed in 15 piezometers giving hydraulic conductivities ranging between 0.2 and 19.6 m d\(^{-1}\) (Table 2) indicating a sediment composition ranging from silty sand/peat to medium sand.

Age-dating by the CFC method rely on the steady increase in groundwater CFC resulting from equilibrium with increasing CFC in the atmosphere since 1940. After the agreement on the Montreal Protocol in the 1989, however, a reversal has been observed for CFC-11 (1994) and CFC-12 (2002), implicating a less precise age dating by the CFC method. The age of the youngest groundwater having the highest concentration of CFC could therefore only be approximated by \(<10 \text{ yr}\) or \(<15 \text{ yr}\) (Table 3). Age-dating of groundwater samples from 10 of the 15 levels at location S1 and S2 had to be based on CFC-11 alone as the CFC-12 concentrations exceeded the maximum possible for air–water equilibrium possibly due to contamination. The source of the contaminant is discussed in Section 4.4. As CFC-11 is susceptible to degradation under anoxic conditions groundwater ages based on CFC-11 are considered to reflect maximum ages under such conditions (Table 3). For some samples of relative young groundwater very high concentrations of nitrous oxide, most likely from denitrification, partly obscured the CFC-12 peak in the analysis. This problem was circumvented by prolonged purging of the cold trap by which most N\(_2\)O was removed. The coefficient of variation for duplicate CFC-11 and CFC-12 concentrations ranged from 0% to 41% (average 6%) and from 0% to 5% (average 2%), respectively.

At S1, multi-level sampling was conducted from 1 m below the streambed down to 4 m. It was not possible to sample at 0.5 m below the streambed due to the presence of silty deposits failing to conduct enough water to the screen during pumping. Results from the major-ion analysis (K\(^+\), NO\(_3\)\(^{-}\), SO\(_4\)\(^{2-}\)) together with dissolved oxygen (DO) concentrations and apparent CFC-ages are shown in Table 3 and Fig. 2. Note that the sampling depths under the streambed are shown relative to terrain. DO decreased from over 7 mg L\(^{-1}\) in the upper 2.5 m to less than 2 mg L\(^{-1}\) below 3.5 m. A lab check of the oxygen sensor afterwards indicated that the low values (below 2 mg L\(^{-1}\)) were probably even lower, possibly zero. A similar pattern was evident for NO\(_3\) with an average concentration of 58.2 mg L\(^{-1}\) in the top part of the streambed, while below 3.5 m, the concentrations decreased significantly to non-detectable limits (<0.05 mg L\(^{-1}\)). A shift from an oxidized to a reducing environment could explain the abrupt change indicating the presence of a redoxcline (transition zone from oxic to anoxic environment) at a depth of \(<3.5\text{ m}\). The decrease in NO\(_3\) concentrations could be controlled by denitrification processes and the generally inverse patterns of SO\(_4\)\(^{2-}\) compared to NO\(_3\) and DO indicate that nitrate reduction could be partly controlled by pyrite oxidation; elevated SO\(_4\)\(^{2-}\) concentrations in the deeper reduced environment suggest that groundwater from 3 to 5 m depths undergo pyrite oxidation generating excessive SO\(_4\)\(^{2-}\) (Postma et al., 1991; Tesoriero et al., 2000). Postma et al. (1991) also found that pyrite was the main electron donor in a sandy outwash aquifer 50 km north of our field site even though the content of organic material was high in the

<table>
<thead>
<tr>
<th>Thermal parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volumetric heat capacity of the water, ( \rho C_0 )</td>
<td>( 4.19 \times 10^6 \text{ J m}^{-3} \text{ K}^{-1} )</td>
</tr>
<tr>
<td>Thermal conductivity of the saturated sediment, ( \lambda )</td>
<td>( 1.8 \text{ J s}^{-1} \text{ m}^{-1} \text{ K}^{-1} )</td>
</tr>
<tr>
<td>Porosity, ( n )</td>
<td>0.3</td>
</tr>
<tr>
<td>Vertical extent of the domain, ( L )</td>
<td>4 m</td>
</tr>
<tr>
<td>Groundwater temperature at, ( T_z )</td>
<td>8 °C</td>
</tr>
<tr>
<td>Stream temperature at, ( T_s )</td>
<td>Varies(^a)</td>
</tr>
</tbody>
</table>

\(^a\) Assumed.
\(^b\) The streambed temperature is based on a monthly average and considered to represent steady-state condition for the particular month.
sediments. We have no direct evidence of the presence of pyrite, but the high SO$_2^-$ concentrations indicate that pyrite oxidation is a possible mechanism also at our site. Furthermore, elevated K$^+$ concentrations (>6 mg L$^{-1}$) down to 3 m suggests that the upper samples could be affected by manure from livestock in the wetland or adjacent fields.

At S2 the observed pattern of decreasing DO and NO$_3^-$ concentrations is also evident, but not as pronounced as in S1. Below 4.5 m NO$_3^-$ concentrations were <1 mg L$^{-1}$ (Fig. 2A). Also, the pattern of increasing followed by decreasing SO$_4^{2-}$ concentrations with depth was observed, but with a broader peak compared to S1. The peak in SO$_4^{2-}$ concentrations at S2 is shifted 2 m deeper relative to S1 and SO$_4^{2-}$ concentrations reach the same level as observed in S1 between ~3.5 and 4.5 m below terrain. Likewise, the inferred redoxcline shifted to a depth of ~4.5 m compared to S1. The decreasing NO$_3^-$ concentrations can again be explained by pyrite oxidation. Also, samples possibly affected by manure are observed down to 4.5 m. The apparent CFC-ages were similar in range for S1 and S2; in the top the apparent age range is $\leq$20 yr increasing to $\leq$30 yr with depth. The apparent CFC-ages are inversely correlated to NO$_3^-$ concentrations with young groundwater having the highest NO$_3^-$ concentrations and vice versa. High NO$_3^-$ concentrations in younger groundwater have been attributed to increased use of fertilizer (Böhlke and Denver, 1995; Modica et al., 1998; Kennedy et al., 2009). However, since more anoxic conditions are found in the deeper parts of the system the lower nitrate concentrations here are more likely due to denitrification rather than less use of fertilizer. In this zone of denitrification, both CFC-11 and CFC-12 could undergo degradation meaning that the apparent ages are likely younger (Table 3).

At S3, S4, and S5 the sampling could not continue for more than 3–6 m below terrain due to difficulties in penetrating deeper into the aquifer. NO$_3^-$ and SO$_4^{2-}$ concentrations were high in all samples together with elevated concentrations of DO (>2.0 mg L$^{-1}$) and with apparent CFC-age range <20 yr (Fig. 2B). One exception is the sample collected in S4 at a depth of 5.25 m, where the DO concentration <2.0 mg L$^{-1}$. In S4 and S5 there is a peak in NO$_3^-$ concentrations at a depth of 4–5 m increasing from ~20 to ~40 mg L$^{-1}$. Thus, in summary, the wetland–stream system is stratified with

### Table 2

<table>
<thead>
<tr>
<th>ID</th>
<th>Hydr. cond. K (m d$^{-1}$)</th>
<th>Std. dev. (m d$^{-1}$)</th>
<th>Mean, h (m)</th>
<th>Obs. range (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0.2</td>
<td>–</td>
<td>51.33</td>
<td>0.20</td>
</tr>
<tr>
<td>P2</td>
<td>–</td>
<td>–</td>
<td>51.47</td>
<td>0.20</td>
</tr>
<tr>
<td>P3</td>
<td>4.8</td>
<td>0.2</td>
<td>51.24</td>
<td>0.18</td>
</tr>
<tr>
<td>P4</td>
<td>4.3</td>
<td>0.8</td>
<td>51.16</td>
<td>0.25</td>
</tr>
<tr>
<td>P5</td>
<td>2.4</td>
<td>0.1</td>
<td>51.04</td>
<td>0.27</td>
</tr>
<tr>
<td>P6</td>
<td>–</td>
<td>–</td>
<td>51.21</td>
<td>0.90</td>
</tr>
<tr>
<td>P7</td>
<td>–</td>
<td>–</td>
<td>50.91</td>
<td>0.21</td>
</tr>
<tr>
<td>P8</td>
<td>1.5</td>
<td>1.7</td>
<td>51.00</td>
<td>0.30</td>
</tr>
<tr>
<td>P9</td>
<td>5.3</td>
<td>0.2</td>
<td>50.98</td>
<td>0.30</td>
</tr>
<tr>
<td>P10</td>
<td>7.7</td>
<td>0.6</td>
<td>50.96</td>
<td>0.31</td>
</tr>
<tr>
<td>P11</td>
<td>6.5</td>
<td>0.5</td>
<td>51.02</td>
<td>0.45</td>
</tr>
<tr>
<td>P12</td>
<td>16.7</td>
<td>0.4</td>
<td>50.86</td>
<td>0.26</td>
</tr>
<tr>
<td>P13</td>
<td>15.3</td>
<td>0.5</td>
<td>50.80</td>
<td>0.33</td>
</tr>
<tr>
<td>P14</td>
<td>–</td>
<td>–</td>
<td>51.13</td>
<td>0.63</td>
</tr>
<tr>
<td>P15</td>
<td>19.6</td>
<td>0.4</td>
<td>50.76</td>
<td>0.21</td>
</tr>
<tr>
<td>P16</td>
<td>0.5</td>
<td>–</td>
<td>50.89</td>
<td>–</td>
</tr>
<tr>
<td>P17</td>
<td>0.2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>P18</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>P19</td>
<td>0.2</td>
<td>–</td>
<td>51.44</td>
<td>–</td>
</tr>
<tr>
<td>P20</td>
<td>–</td>
<td>–</td>
<td>51.11</td>
<td>–</td>
</tr>
</tbody>
</table>

The measured hydraulic conductivities within the piezometers with the standard deviation and the mean of the hydraulic head measurements with the observed range. It is noted that no range is reported for P16–P20 due to the late installation of these piezometers.

### Table 3

<table>
<thead>
<tr>
<th>ID</th>
<th>Depth (m)</th>
<th>CFC-11 (pg kg$^{-1}$)</th>
<th>CFC-12 (pg kg$^{-1}$)</th>
<th>App. CFC-11 (yr)</th>
<th>App. CFC-12 (yr)</th>
<th>K$^+$ (mg L$^{-1}$)</th>
<th>NO$_3^-$ (mg L$^{-1}$)</th>
<th>SO$_4^{2-}$ (mg L$^{-1}$)</th>
<th>O$_2$ (mg L$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.4</td>
<td>762</td>
<td>389</td>
<td>$\leq$15</td>
<td>$\leq$10</td>
<td>6.4</td>
<td>33.1</td>
<td>21.6</td>
<td>9.6</td>
</tr>
<tr>
<td>M</td>
<td>0.4</td>
<td>28</td>
<td>12</td>
<td>$\leq$50</td>
<td>$\leq$60</td>
<td>1.3</td>
<td>0.1</td>
<td>23.9</td>
<td>2.2</td>
</tr>
<tr>
<td>S</td>
<td>0.4</td>
<td>19</td>
<td>1485</td>
<td>$\leq$50</td>
<td>$\leq$50</td>
<td>1.7</td>
<td>0.1</td>
<td>56.5</td>
<td>3.2</td>
</tr>
<tr>
<td>S1</td>
<td>2.0</td>
<td>818</td>
<td>293</td>
<td>$\leq$25</td>
<td>$\leq$25</td>
<td>9.0</td>
<td>58.1</td>
<td>42.6</td>
<td>7.0</td>
</tr>
<tr>
<td>S2</td>
<td>1.5</td>
<td>808</td>
<td>397*$^*$</td>
<td>$\leq$20</td>
<td>$\leq$10</td>
<td>5.5</td>
<td>46.9</td>
<td>9.8</td>
<td>2.5</td>
</tr>
<tr>
<td>S3</td>
<td>3.0</td>
<td>1127</td>
<td>415*$^*$</td>
<td>$\leq$15</td>
<td>$\leq$10</td>
<td>5.5</td>
<td>38.8</td>
<td>24.5</td>
<td>8.4</td>
</tr>
<tr>
<td>S4</td>
<td>3.5</td>
<td>690</td>
<td>780*$^*$</td>
<td>$\leq$25</td>
<td>$\leq$25</td>
<td>6.9</td>
<td>37.9</td>
<td>60.1</td>
<td>2.8</td>
</tr>
<tr>
<td>S5</td>
<td>4.0</td>
<td>54</td>
<td>2360</td>
<td>$\leq$50</td>
<td>$\leq$50</td>
<td>2.3</td>
<td>0.1</td>
<td>69.7</td>
<td>1.7</td>
</tr>
</tbody>
</table>

$^*$ CFC-12 peak strongly affected due to very high N$_2$O in water sample. Prolonged purging of the cold trap removed most NO$_2$ in the second sample analyzed.

$^b$ Excessive concentrations exceeding that possible for air–water equilibrium. ID: N, M, and S are groundwater samples from the stream at the northern bank, in the middle and southern bank.
a shallow 3–4 m deep zone with relatively young, oxic, and nitrate-rich groundwater followed by a deeper reduced zone with older groundwater free of nitrate.

4.2. ERT

The modeled resistivity values of the land-based ERT had an RMS difference of 3.0% between the measured and the calculated electrode voltages. A data set with reciprocal errors <5% was used in the inversion leaving 883 of the collected 891 measurements. The land-based ERT profile (Fig. 3A) shows two general characteristic differences with relatively low and high resistivity materials underneath the wetland on the northern side of the stream (<500 m) and in the streambed and on the southern bank of the stream (≥500 m), respectively. The stream-based ERT had an RMS of 4.7% between the measured and calculated electrode voltages (Fig. 3B). A layer with low resistivity (<150 m) covers the top of the profile. Furthermore, three wedges of higher resistivity values (>300 m) were found at the flanks and in the middle of the profile. In between the wedges, resistivity values are from 150 to 300 m. The horizontal 18 m mark in the in-stream ERT crosses the land-based ERT at its 58 m mark. The wedge with high resistivity values (>550 m) found below the stream in the land-based ERT is also detected in the stream-based ERT profile. Temperature

Fig. 2. Water quality and apparent CFC-age range from sampled locations and flow path simulation for case 1. The apparent CFC-age ranges are maximum apparent age ranges calculated from the CFCs. Colors in cross sections are described in Fig. 4. (A) Sampling results and flow path simulation for S1 and S2. (B) Sampling results and flow path simulation for S3–S5. (C) General flow path simulation for the wetland–stream system.
effects in surface-near ERT can give an effect of 1.8–2.2% change in bulk EC per °C (Hayley et al., 2007). The largest temperature contrasts at our field site is ~8 °C between groundwater and stream. Therefore, the changes in bulk EC due to temperature effects are considered negligible relative to the modeled range in resistivity. Higher ion concentrations were observed at some depths (Table 3) and could also impact the ERT measurements. However, the entire ion balance is needed to estimate the impacts and this was not available. But the affected depths do not show conspicuously low resistivity values in the ERT profile and therefore the ion concentration effects on the ERT measurements were regarded as negligible.

Results from slug tests (Table 2) are included in the analysis relating these with typical resistivity values for Danish sediments given by Ahrentzen et al. (1987) to assign zone-based hydraulic conductivities, \( K \), in units defined by the modeled resistivity ranges. The locations of the piezometer screens in the ERT profile are shown in Fig. 3A. The interpreted hydrogeological units within the ERT profile are shown in Fig. 4 and consists of seven units, M1–M7, with increasing \( K \). \( K \) in M1 was based on slug tests in P1, P17 and P19, which were screened at or in close proximity to areas with relatively low resistivity values ranging from 50 to 250 \( \Omega \cdot m \). \( K \) in M2 was based on the average of \( K \) determined in P5 and P8 screened in resistivity layers ranging between 250 and 300 \( \Omega \cdot m \). \( K \) in M3 was the average of slug tests from P3 and P4 screened in resistivity layers ranging primarily from 350 to 450 \( \Omega \cdot m \), but also enclosing resistivity layers up to 550 \( \Omega \cdot m \). \( K \) in M4 was based on results from P8, P10, and P11. The layers in which P9 and P10 were screened included a wide range of resistivity values (300–800 \( \Omega \cdot m \)). But, due to the minimized extent of the high resistivity layers located near the two screens and the similarity in measured \( K \), the layers were considered as one unit (Fig. 3A). While P11 was screened between layers having resistivity values of 300–350 \( \Omega \cdot m \), as also seen in M3, the resistivity values surrounding P11 were included in M4 due to the similarity of \( K \) with values in P9–P10. \( K \) in M5 was based on results from P12 and P13 although they both seem to be screened in resistivity layers ranging between 50 and 300 \( \Omega \cdot m \). With such low resistivity values \( K \) would be expected to be lower reflecting finer sediment. But as the electrodes were placed in the streambed it is likely that the low resistivity values measured immediately beneath the stream are an artifact caused by the current running through the stream water (~36.4 \( \Omega \cdot m \)) instead of through the underlying streamed. Moreover, previous results from 40 slug tests performed in October 2010 in screens 0.5 m below the streambed covering a bank to bank area of 40 m² (average \( K = 15.9 \) m d⁻¹), are in agreement with the \( K \) assigned to M5. An M1 zone is assigned to the uppermost part of the northern bank based on the fact that visual inspection revealed fine-grained sediments and the water sampling failed 0.5 m below the streambed. \( K \) in M6 was based on results from P15. Here, the screen may have been situated in layers with lower resistivity values compared to the high \( K \). But as the resistivity layers (400–500 \( \Omega \cdot m \)) include a wedge with high resistivity values in both ERT profiles, all resistivity layers are considered as one highly conductive zone. M7 included the highest measured resistivity values (>800 \( \Omega \cdot m \)) that are typical for sand and gravel above the water (Ahrentzen et al., 1987). Hence, \( K \) in M7 was assumed equal to M6.

Although the measured resistivity values were overall comparable with typical resistivity values for Danish sediments it is clear that resistivity values cannot stand alone in order to determine the hydraulic conductivity of a given geological unit. The measured hydraulic conductivities within a defined unit, thus also varies. Secondly, the areal extent of the zones may be difficult to determine due to the inherent smooth result of the least square ERT inversion, and is therefore not always coherent with the changes in hydraulic conductivity. Thirdly, the hydraulic conductivity and resistivity values also overlap preventing an unambiguous determination of the sediment composition. Hence, some degree of subjectivity was required in the interpretation of the ERT and \( K \) data. In Section 4.4 the deduced heterogeneous conceptual hydrogeological model (Fig. 4) is compared to a homogeneous model with

![Fig. 3](image-url)
hydraulic conductivity equivalent to the mean of the slug test measurements.

### 4.3. Temperature analysis

Estimates of discharge based on monthly-averaged temperature profiles are shown in Table 4. T1, T2, and T3 are located in the northern, middle, and southern part of the stream, respectively. Examples of best and worst fits are shown in Fig. 5 and illustrate the varying temperature distribution across the streambed. All three locations have temperatures of groundwater (around 8°C) in the bottom of the profile but show distinct differences in temperature in the upper part of the profile; predominantly, in T1 where stream temperature effects are observed down to ~1 m. The fluxes vary considerably across the stream with up to one order of magnitude. The lowest mean flux was observed at the northern bank, while the highest mean flux was observed in the middle of the stream. The fluxes at T1 were better defined in terms of the coefficient of determination, \( r^2 \), than for T2 and T3. Lateral- or hyporheic flow may explain the lower \( r^2 \) due to the violation of the 1D assumption resulting in a poorer fit. A similar conclusion was made by Jensen and Engesgaard (2011) and the issue of lateral and hyporheic flow is discussed further below.

### 4.4. Numerical modeling

The groundwater flow model was calibrated against the average measured hydraulic heads. To test the application of a conceptual model based on the combined ERT and slug test measurements three cases were studied; (1) a distributed \( K \) based on the ERT-\( K \) zones described in Sections 3.4 and 4.2, (2) an automated parameter estimation using PEST (Doherty, 2004) calibrating \( K \) in the most sensitive ERT-\( K \) zones from (1), and, finally, (3) a model with a homogeneous \( K \) equal to the mean of the measured hydraulic conductivities. In Fig. 2C the flow path pattern in case 1 show an asymmetrical convergence of flow lines from the boundaries to the stream. Flow lines from the southern boundary terminate across most of the stream bed opposed to flow lines from the northern boundary that are restricted to the northern stream bank. The flow paths are clearly governed by the geology, for example, the flow lines from the upper southern boundary are diverted around the low conductivity zone close to stream. Lateral flow to the banks is also occurring. In case 1, observed vs. simulated heads gave \( r^2 \) equal to 0.73. For case 2, where \( K \) in the most sensitive zones were calibrated, \( r^2 \) was slightly improved to 0.77. For case 3, \( r^2 \) was lower, 0.66. Even though case 2 improved model performance slightly, the optimal \( K \) values from the PEST calibration were still uncertain, i.e., having a large 95% confidence interval. The optimal values are therefore not unique and, moreover, in four of the six zones calibrated values were different by more than one order of magnitude from the measured values. The unrealistic parameterization is probably due to hydraulic heads failing to constrain the calibration. But as discussed further below, case 1 is interpreted to satisfactorily represent the flow within the wetland–stream system and the observed vs. simulated heads are shown in Fig. 6. There seems to be a constant difference (±0.2 m) between simulated and observed heads, which is likely related to the hydraulic head measurements not capturing the dynamics entirely and thus producing less correct averages used to represent steady-state conditions.

The simulated stream fluxes are plotted in Fig. 7 and show large variation across the streambed with differences up to one order of magnitude. The lowest fluxes were simulated at the northern bank (except for case 3) then increasing towards the middle followed by a slight decrease and then a large peak right at the southern bank. Although case 2 performs better than case 1 in matching the simulated and observed heads, the overall simulated fluxes in case 1 are closer to the three temperature-based fluxes. Compared to the temperature-based fluxes the simulated flux at T1 in case 1

---

Table 4: Temperature based fluxes.

<table>
<thead>
<tr>
<th></th>
<th>T1 (m d(^{-1})) ( r^2 )</th>
<th>T2 (m d(^{-1})) ( r^2 )</th>
<th>T3 (m d(^{-1})) ( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008 April</td>
<td>0.15</td>
<td>0.96</td>
<td>0.33</td>
</tr>
<tr>
<td>May</td>
<td>0.11</td>
<td>0.99</td>
<td>0.33</td>
</tr>
<tr>
<td>June</td>
<td>0.10</td>
<td>1.0</td>
<td>0.46</td>
</tr>
<tr>
<td>July</td>
<td>–</td>
<td>–</td>
<td>0.71</td>
</tr>
<tr>
<td>August</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>September</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>October</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2009 January</td>
<td>0.10</td>
<td>0.98</td>
<td>1.34</td>
</tr>
<tr>
<td>February</td>
<td>0.09</td>
<td>0.99</td>
<td>1.31</td>
</tr>
<tr>
<td>March</td>
<td>0.12</td>
<td>0.89</td>
<td>1.24</td>
</tr>
<tr>
<td>April</td>
<td>–</td>
<td>–</td>
<td>0.98</td>
</tr>
<tr>
<td>May</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>June</td>
<td>–</td>
<td>–</td>
<td>0.84</td>
</tr>
<tr>
<td>July</td>
<td>–</td>
<td>–</td>
<td>0.95</td>
</tr>
<tr>
<td>2010 April</td>
<td>–</td>
<td>–</td>
<td>1.34</td>
</tr>
<tr>
<td>May</td>
<td>0.17</td>
<td>0.95</td>
<td>1.23</td>
</tr>
<tr>
<td>June</td>
<td>0.11</td>
<td>0.99</td>
<td>1.31</td>
</tr>
<tr>
<td>July</td>
<td>0.10</td>
<td>0.99</td>
<td>1.06</td>
</tr>
<tr>
<td>Mean</td>
<td>0.11</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>Std. dev</td>
<td>0.02</td>
<td>0.03</td>
<td>0.37</td>
</tr>
</tbody>
</table>

1D flux ranges, \( q_z \), for the groundwater discharge deduced from monthly averaged temperature measurements. \( r^2 \) is the coefficient of determination. The mean and the standard deviation are shown for all the measurements for each probe. The maximum and minimum flux is found in the middle (T2) and in the northern part of the stream (T1) and overall the fluxes deduced for T1 are well determined while \( r^2 \) for T2 and T3 are noticeably lower.
was underestimated 81%. The subtle changes in the simulated fluxes at the northern bank, however, are generally very close to that observed at T1. For example, model cell 1, right next to the location of T1, simulated a flux with only a −17% difference, i.e. close to the mean and within the range of the temperature-based fluxes (Fig. 7). Therefore a finer model representation of the streambed geometry could lead to a better correlation between the temperature-based and simulated flux at T1. The simulated flux at T2 was very close to the mean (within 3%) and within the high range of the temperature-based fluxes, thus suggesting that the model is able to capture the discharge pattern from the northern bank to the middle. The simulated flux for T3 were lower than the temperature based mean (−39%) and just outside the lower range. Here, it has to be recalled that the temperature-based fluxes at T3 are uncertain (Table 4) likely because lateral flow is occurring at the southern bank resulting in the poorer fits of the 1D analytical heat solution to the temperature profiles. Simulated flow paths also showed horizontal flow to the model cell (Fig. 2C) at the southern boundary supporting lateral flow. Horizontal flow is also simulated at the northern bank, but the temperature fluxes at T1 are still well determined indicating that the lower lateral flow at the bank shifts to near-vertical flow at T1. Still, the model captures the higher discharge at the southern bank compared with that of the northern bank. Simulated fluxes in case 3 are largest at the banks and do generally not vary across the stream. Neither are the simulated fluxes within the temperature based ranges. Thus,
with the well simulated hydraulic heads and fluxes case 1 was believed to represent groundwater flow conditions within the domain.

The simulated flow paths (case 1) are depicted in Fig. 2A and B where released particles at S1 were tracked to both sides of the stream. Particles released 2–3 m below terrain were tracked back to the northern boundary. Flow paths from these depths also have the youngest apparent CFC-age range. Of these, samples from 2 to 2.5 m below terrain originate from the top part of the aquifer having the shortest travel path. The flow path for the sample 3 m below terrain originate deeper within the aquifer, which is also reflected in an increasing apparent CFC-age range (Fig. 2A). Particles released from 3.5 to 5 m below the streambed originate on the other hand from the southern boundary. Flow paths for the four bottom samples originate from the bottom of the southern boundary traveling in the high conductivity zones and bringing in water with older apparent CFC-ages. The apparent increasing CFC-age range suggests that the water may come from further away from the simulated boundary. This means that the groundwater discharging to the stream through this high conductivity zone may necessarily originate from far away to generate the relative older ages. The simulated residence time for the samples originating from the southern boundary is ~0.3 yr.

In S2, all the released particles were tracked back to the northern boundary, where particles released down to 4.5 m originate deeper and deeper within the aquifer according to the sampling depths. Thus, the flow paths are longer matching the general increase in the apparent CFC-age range with depth (Fig. 2B). For the deepest samples (5.5–8.5 m below terrain), the flow paths originate from the bottom of the northern boundary with all particles forced to travel through the low conductivity layer, M1, and thus bringing in older water. The simulated residence time for the samples in S1 and S2 traveling in the upper oxic part of the aquifer is ~1.7 yr. In contrast, the samples from the deeper part of the aquifer originating from the northern boundary have a simulated residence time of ~7.4 yr.

Flow paths for S3–S5 were tracked back to the upper part of the northern boundary and followed the flow patterns simulated for the uppermost samples in S2 all originating from the upper part of the aquifer. The simulated residence time for these samples decreased from ~1.2 yr to ~0.3 yr from S3 to S5, respectively.

From the flow path simulation and the major-ion analysis it was evident that NO3 concentrations were only detected in groundwater originating from the northern boundary, i.e. from areas with agricultural land use. The simulated residence time for the upper samples compared to the lower samples from the northern boundary were also more than seven times lower suggesting relatively limited time for reduction processes. In S1, NO3 concentrations were only observed in the top three samples (2–3 m below terrain) agreeing with particles originating from the northern boundary. However, flow path for the sample at 3 m below terrain travel within the deeper part of the aquifer inferred to be anoxic and thus not expected to carry NO3 concentrations (Section 4.1). The sample from 3 m below terrain has NO3 concentration ~6 times lower than the samples above in the upper oxidized part of the aquifer, suggesting mixing of deeper nitrate-free groundwater with shallow nitrate-containing groundwater. The four bottom samples without NO3 concentrations originating from the southern boundary could reflect a different source and geochemical environment. Previous sampling in September 2009 at three locations across the stream (at the northern and southern bank and in the middle) at 0.4 m below the streambed showed contaminated CFC-12 concentrations exceeding that possible for air–water equilibrium at the southern bank only. It is suggested that the contaminant CFC-12 source is most likely related to illegal dumping of waste in an old gravel pit recognized 60–100 m south of location S1.

In S2, NO3 concentrations were observed down to 4.5 m agreeing with particles originating from the top of the aquifer. At depths >4.5 m, where the NO3 concentrations decrease markedly, particles originate from the bottom of the aquifer traveling within the inferred anoxic zone. Still, four of the eight samples in S2 have contaminated CFC-12 concentrations as seen in samples originating from the southern boundary. Therefore, mixing of groundwater with differing origin is likely to take place in between S1 and S2 coinciding with the convergence of flow lines from the northern and southern boundary (Fig. 2B and C). The decreasing SO4 concentrations below the peak in both S1 and S2 were not readily elucidated from the flow path simulation. We cannot outline the possibility of upward leaking older groundwater. Thus, a possible explanation could be that the area in between S1 and S2 was affected by mixing of older groundwater from deeper parts of the aquifer carrying another chemical composition with lower SO4 concentrations causing the depletion. For example, the chemical composition from deep groundwater sampled in 2010 from ~126 m below terrain at the local water work (~4.5 km east of the field site) showed SO4 concentration of 4 mg L~1.

All samples in S3–S5 have NO3 concentrations supporting the above findings that the discharging groundwater from the northern boundary may not be subjected to denitrification during its passage through the wetland adjacent to the stream.

4.5. Implications

Typologies have been proposed to help understand and predict the type of groundwater–surface water interactions (Dahl et al., 2007) and they may be adequate at larger scales, but not to predict the specific flow regime in a wetland–stream system. Thus, when assessing the fate of nutrients in a wetland–stream system, typologies becomes insufficient and better and more detailed conceptualizations of the effects and impacts of the variation in the subsurface hydrogeology on groundwater–surface water systems are needed. Recognizing the increased need to manage groundwater and surface water quality (e.g. according to the WFD in Europe) this study combined hydrogeological, hydrogeophysical, and hydrochemical field methods with numerical modeling to improve the characterization and conceptualization of the subsurface architecture in a wetland–stream system. Our study shows the following; (i) a homogeneous model does not predict the correct distribution of discharge to the stream as compared to a model based on ERT and a hydrogeological characterization. (ii) Groundwater head data were not enough to constrain the model, however auxiliary data such as estimated discharge to the stream (from temperature profiles), CFC-age dating and hydrochemical profiles were helpful in providing confidence in the model conceptualization. Thus, both measured hydraulic heads and streambed fluxes were in good agreement with simulated values. Furthermore, the flow path simulation underscored the importance of the subsurface architecture in controlling the flow paths to the stream, which could not have been determined from the hydraulic head data and/or the geochemical sampling alone. The initial ERT-K model was, however, able to simulate this variability without any calibration. This study did not attempt to address nitrate attenuation as such although we did find evidence of near-stream nitrate reduction processes. Our data on water and sediment quality does not allow us to be more quantitative about nitrate reduction processes and capacities. Future studies addressing the fate of nutrients should therefore include a characterization of the nitrate reduction capacity of the wetland system (organic matter, pyrite) and continuously monitoring/sampling of geochemical components. This information can then be used to define appropriate reactive processes in a transport model and through simulations come up with, e.g. nitrate reduction parameters. In the design of future studies in
similar settings the applied and integrated combination of hydrogeological, hydrogeophysical, and hydrochemical methods together with monitoring would enhance the understanding of the groundwater–surface interaction with respect to both water and contaminant fluxes. This would also lead to better strategies for managing or using wetland systems for attenuation of contaminants.

5. Conclusion

We show that: (1) by combining ERT with results from slug tests into a numerical model we are able to simulate hydraulic heads and streamed fluxes in correspondence with field observations and (2) that the properties of the subsurface cannot be based exclusively on modeled resistivity values from ERT or hydraulic conductivity measurements. Flow path simulations combined with major-ion analysis of groundwater show that the origin of the discharging groundwater to the stream is markedly different and therefore carry a chemical composition reflecting differing sources. Moreover, the upper part of the wetland receiving most of the nitrogen input from nearby crop fields does not act as a buffer for the NO3− contamination as reported in many studies.

The 2D simulation with a hydrogeological model based on ERT combined with hydraulic conductivities generates fluxes and path lines revealing lateral groundwater discharge and different origin of the sampled groundwater that could not have been predicted from the collected data alone. ERT in combination with synoptic measurements and numerical modeling offers a strong tool to delineate groundwater flow patterns.

Acknowledgements

We would like to thank Louise A. Andreassen, Mie Andreassen, Astrid M. Bisgaard, Carlos Duque, Valentina Marconi, Sascha Müller, Helene H.M. Nielsen, Eva Sebok and the students who have been helping over the years with collecting field data. We also thank an anonymous reviewer and the associate editor. This work is part of the Hydrology Observatory, HOBE (http://www.hobecenter.dk) funded by the Villum Foundation.

References


Hvorslev, M.J., 1951. Time lag and soil permeability in groundwater observations. Waterways Experimental Station Bulletin 36 United States Army Corps of Engineers, MS, USA.


