groundwater®
Exhibits • Workshops • Summit
Las Vegas, Nevada • December 3-5, 2019

GROUNDWATER WEEK 2019
PROMOTES SCIENCE!

Hear the latest in groundwater research during presentations from scores of scientists and practitioners at this year’s Groundwater Week.

Session topics range from emerging contaminants and managed aquifer recharge to groundwater remediation and sustainability — plus everything in between.

Register: www.GroundwaterWeek.com
Explicit Modeling of Radon-222 in HydroGeoSphere During Steady State and Dynamic Transient Storage

by B.S. Gilfedder1,2, I. Cartwright3, H. Hofmann4, and S. Frei5

Abstract

Transient storage zones (TSZs) are located at the interface of rivers and their abutting aquifers and play an important role in hydrological and biogeochemical functioning of rivers. The natural radioactive tracer 222Rn is a particularly well-suited tracer for studying TSZ water exchange and age. Although 222Rn measurement techniques have developed rapidly, there has been less progress in modeling 222Rn activities. Here, we combine field measurements with the numerical model HydroGeoSphere (HGS) to simulate 222Rn emanation, decay and transport during steady state (riffle-pool sequence) and transient (bank storage) conditions. Comparing the HGS mean water ages with the conventional 222Rn apparent ages during steady state showed a systemic underestimation of apparent age with increasing dispersion and especially where large concentration gradients exist within the subsurface. A large underestimation of apparent water age was also observed at the advective front during bank storage where regional high 222Rn groundwater mixes with newly infiltrated surface water. The explicit modeling of radiogenic tracers such as 222Rn offers a physical interpretation of this data as well as a useful way to test simplified apparent age models.

Introduction

Subsurface transient storage zones (TSZs) are areas of active water exchange at the interface between streams, rivers, and the aquifer system. They are an important component of river corridor science and influence the chemistry, hydrology, biology, and physics of water ways (Harvey and Gooseff 2015). Examples of subsurface transient storage include hyporheic flow, paraflluvial flow, bank storage, and flood recharge (Boano et al. 2014). Dispersive mixing between groundwater and surface water in TSZs is responsible for an array of chemical and biological processes that are ecologically significant (Cardenas 2016). For example, the hyporheic zone often facilitates a redox gradient important for both the production and reduction of nitrate, for carbon turnover and oxygen consumption due to microbial respiration (Zarnetske et al. 2011b; Kessler et al. 2012). Hyporheic areas in particular provide an important spawning ground for fish species and a vital habitat for stream invertebrates (Boulton 1993; Boulton et al. 1998; Marzadri et al. 2012; Cardenas et al. 2016). The larger TSZs (e.g., bank storage) can also be responsible for buffering and delaying storm flow, supplying rivers with water during base flow and regulating stream temperatures (Todd 1955; Arrigoni et al. 2008; McCallum et al. 2010).

A fundamental distinction can be made between TSZs that are active under both steady and nonsteady streamflow (e.g., hyporheic exchange) and those that are only activated during transient hydrological events such as storm flow (e.g., bank storage). Applying this distinction to TSZ biogeochemistry has led to the concepts of “hot spots” which are usually associated with TSZs in the streambed that are active at all times (Zarnetske et al. 2011a) and “hot moments” that are initiated by transient...
hydrological events (e.g., high river stage) and are only sporadically present (Gu et al. 2012; Frei and Peiffer 2016).

Both hydrological quantities (e.g., storage volume, water flux) and chemistry (e.g., kinetic redox processes) of TZSs are dependent on surface water residence times, or “age” ($\tau$) of water since infiltration (Boano et al. 2014). However, deriving water age from field data is nontrivial (Wörman and Wachtel 2007), especially when time-series measurements during transient hydrological events are needed (Ward et al. 2013). Natural tracers are a useful series of measurements during transient hydrological events (Worman and Wachniew 2007), especially when time-series measurements during transient hydrological events (e.g., high river stage) and are only sporadically present (Malcolm et al. 2013). This effort is due in part to the low cost of temperature sensors, but also to the complexity of the heat transport equation when used for hyporheic sediments with variable boundary conditions (one- to three-dimensional (1-3D) flow, variable upper boundary, variably saturated flow (Cuthbert and Mackay 2013; Rau et al. 2017). In contrast, the quantification of $\tau$ using $^{222}$Rn has not significantly evolved since Hoehn and von Gunten (1989), a model which is conceptually similar to many “apparent age” models used in the groundwater age dating literature (Ekwurzel et al. 1994; Sanford 2010). The Hoehn and von Gunten (1989) model was originally developed to quantify the steady-state flow velocity away from a losing river (i.e., riverine ecosystems). This is especially important during dynamic hydrological events such as bank return flow where water changes direction, over-bank recharge, and subsequent return flow, or mixing of flow lines by dispersion in complex 2D and 3D flow fields. There is also still very little information on the uncertainty that system complexity creates in $\tau_{app}$ when using the Hoehn and von Gunten (1989) model, especially the dispersive mixing of water with different ages. The development of novel continuous $^{222}$Rn measurement techniques in TSZs (e.g., Gilfedder et al. 2013) as well as the increasing use of $^{222}$Rn as a residence time tracer in TSZs (Bourke et al. 2014; Cranswick et al. 2014) requires further development of robust $^{222}$Rn modeling techniques.

In this paper, we use a combination of point and continuous $^{222}$Rn measurements and explicit modeling of $^{222}$Rn via advection-dispersion-reaction in a fully coupled numerical computer code (HydroGeoSphere, HGS). The aim is to (1) demonstrate how HGS can be used to model $^{222}$Rn tracer data during steady and dynamic flow, (2) understand how processes such as dispersion affecting $^{222}$Rn transport and apparent ages in 2D TSZs, and (3) show $^{222}$Rn can be used to calibrate complex process-based models at short residence times. While it is based on specific field examples, the methods and results presented in this study have broad general applicability and will improve the understanding of processes in TSZs in riverine ecosystems.

### Research Methods

**$^{222}$Rn as a Residence Time and Water Flux Tracer in TSZs**

$^{222}$Rn is a radioactive ($t_{0.5} = 3.81$ days) noble gas in the $^{238}$U decay chain. It is produced by the decay of $^{226}$Ra in rocks and sediments and accumulates in aquifer pore water, with activities in excess of 100,000 Bq/m$^3$ locally recorded (Cecil and Green 2000). $^{222}$Rn activities at secular equilibrium depend on aquifer mineralogy, with the highest activities recorded in granitic sediments and the lowest in beach sands. Secular equilibrium (steady state between decay and emanation) is reached after five times the half-life which is approximately 21 days. Surface water generally has very low $^{222}$Rn activities (less than 1000 Bq/m$^3$) due to loss to the atmosphere and radioactive decay. This contrast in activities is commonly exploited to map and quantify the influx of groundwater to lakes and rivers as well as calculate $\tau_{app}$ in the subsurface (Cook 2012; Kluge et al. 2012; Gilfedder et al. 2015). As surface water infiltrates into a TSZ and flows along a hypothetical idealized stream tube the $^{222}$Rn activities increase with time due to emanation from the aquifer matrix (called ingrowth) according to:

$$\frac{d^{222}\text{Rn}[Bq/L^3]}{d\tau} = \frac{\gamma}{\theta} - \lambda^{222}\text{Rn}[Bq/L^3]$$

where $^{222}\text{Rn}[Bq/L^3]$ is the $^{222}$Rn activity in TSZ pore water at apparent age $\tau_{app}$ [T], $\gamma$ is the emanation rate from the sediments [Bq/L$^3$/T], $\lambda$ [per T] is the decay constant ($0.18$ per day), and $\theta$ [-] is the sediment porosity. $\gamma$ is related to $^{222}$Rn at secular equilibrium by $^{222}\text{Rn}_{\gamma \rightarrow \infty} = d\text{Rn}/dt = 0 = \gamma/\theta \lambda$. Equation 1 may be solved analytically for $\tau_{app}$. (Hoehn and von Gunten 2012; Kluge et al. 2012; Gilfedder et al. 2015). This is especially important during dynamic hydrological events such as bank return flow where water changes direction, over-bank recharge, and subsequent return flow, or mixing of flow lines by dispersion in complex 2D and 3D flow fields. There is also still very little information on the uncertainty that system complexity creates in $\tau_{app}$ when using the Hoehn and von Gunten (1989) model, especially the dispersive mixing of water with different ages. The development of novel continuous $^{222}$Rn measurement techniques in TSZs (e.g., Gilfedder et al. 2013) as well as the increasing use of $^{222}$Rn as a residence time tracer in TSZs (Bourke et al. 2014; Cranswick et al. 2014) requires further development of robust $^{222}$Rn modeling techniques.
1989). Including the river water $^{222}\text{Rn}$ concentration (Cranswick et al. 2014; Pittroff et al. 2017) results in:

$$\tau_{\text{app}} = \frac{1}{\lambda} \ln \left[ \frac{^{222}\text{Rn}_\infty - ^{222}\text{Rn}_R}{^{222}\text{Rn}_\infty - ^{222}\text{Rn}_\tau} \right] \quad (2)$$

where $^{222}\text{Rn}_R$ [Bq/L$^3$] is the $^{222}\text{Rn}$ activity in the river water endmember. Apparent flow velocities through the TSZ are then calculated by dividing the flowpath length by $\tau_{\text{app}}$ (Hoehn and von Gunten 1989; Bertin and Bourg 1994). The maximum $\tau_{\text{app}}$ measurable is about 20 days (depending on detector precision) after which secular equilibrium is reached with the TSZ matrix. The primary assumptions associated with this approach are (1) that transport is dominated by advective flow, (2) no mixing between streamlines or with the regional groundwater occur, (3) emanation is a constant in space and time, and (4) the stream water endmember is time invariant.

**Field Site 1: Steady-State Hyporheic Exchange Roter Main River Germany**

The field site 1 is used here as an example of estimating $\tau_{\text{app}}$ during steady-state hyporheic flow. As part of a previous study on hyporheic residence times $^{222}\text{Rn}$ (as well as other chemical parameters) was measured in samples from 5 to 45 cm sediment depth at the beginning (downwelling zone) of a riffle section in the Roter Main River, NE Bavaria, Germany (Pittroff et al. 2017). The Roter Main River is a forth order river with a mean discharge of approximately 4 m$^3$/s and base flow events. The Pearces Lane site is located in the middle section of the Avon River, approximately 30 km upstream from where it enters Lake Wellington and 20 km from the township of Stratford. Pearces Lane is typical for the majority of the lowland sections of the river with extensive highly conductive (K$_{gw} \sim 10^{-3}$ m/s) gravel beds either side of the main channel. A transect of 5 bores were drilled perpendicular to the river at distances between 3 (bore 5) and 55 (bores 1 and 2) meters from the river (Figure 1B). The continuous $^{222}\text{Rn}$ measurements presented here were conducted in bore 5, a 5.5 m deep bore located directly on the river bank (Figure 1B). The continuous $^{222}\text{Rn}$ measurements were conducted using the OctoRad instrument described by Gilfedder et al. (2013). An OctoRad comprises eight 1.5 m long gas-permeable silicone membrane tubes strung vertically between two hollow end pieces which allow air to cycle through the tubes. It is placed in the screened section of a bore and connected in closed circuit with a $^{222}\text{Rn}$ detector situated at the top of the bore using PVC tubing. $^{222}\text{Rn}$ in the groundwater diffuses through the gas-permeable silicone tubes into the air loop and is then measured by a $^{222}\text{Rn}$ detector at the top of the bore. We employed a Geiger counter system (AWARE Electronics Corp, Wilmington, Delaware.) for $^{222}\text{Rn}$ detection as discussed in Gilfedder et al. (2013). $^{222}\text{Rn}$ activities in the water were calculated from the $^{222}\text{Rn}$ activities in the air loop using the Ostwald solubility coefficient $\alpha$ ($R_{GW} = R_{air} * \alpha$). Alpha is calculated by $\alpha = 0.105 + 0.405 e^{-0.0502t}$, where $t$ is the water temperature in °C (Meyer and Schweidler 1916; Weigel 1978). Errors are on the order of less than 20%, except when approaching the detection limit of 1000 Bq/m$^3$, where they can be larger. Measurements were made from October 29 to November 11, 2012. In
addition to $^{222}\text{Rn}$ we also measured river stage, EC and groundwater level using an Aqua Troll 200 (In-situ Inc., Fort Collins, Colorado) in bore 5 and groundwater level in bore 2 (Rugged Troll, In-situ Inc.).

Virtual Experiments: Explicit Simulation of $^{222}\text{Rn}$ During Steady and Transient Conditions

The process-based hydrological model HGS was used to model $^{222}\text{Rn}$ transport and behavior at the two field sites. HGS (Aquanty Inc. 2015) provides a fully integrated 3D solution for variably saturated subsurface flow (Richards equation) and a 2D depth-averaged solution for surface flow based on the diffusive wave approximation to the St. Venant equations (Aquanty Inc. 2015). HGS is being increasingly used for the simulation of coupled surface-subsurface hydrologic systems (e.g., Brookfield et al., 2009; Jones et al., 2006). The two HGS models include (1) a 2D cross section located along the thalweg, representing a typical riffle structure for the Roter Main River (Figure 1A) and (2) a 2D cross section of the Avon River bank based on the field site described above (Figure 1B). HGS was used in order to represent the hydrological conditions for the field sites and the migration of $^{222}\text{Rn}$ in porous media which involves advective and dispersive transport, $^{222}\text{Rn}$-emanation and -decay according to Equation 3:

$$-\nabla \left( q^{222}\text{Rn} - \Theta_s D^{222}\text{Rn} \right) - \Theta_s \lambda^{222}\text{Rn} + \gamma = \frac{\partial \left( \Theta_s^{222}\text{Rn} \right)}{\partial t}$$

(3)

In Equation 3, $q$ [L/T] represents the subsurface fluid flux derived from solving 3D Richards and surface flow equations, $\Theta_s$ [-] the saturated water content, $D$ [L$^2$/T] the hydrodynamic dispersion tensor and $\gamma$ [Bq/L$^3$/T] a zero-order source term representing emanation of $^{222}\text{Rn}$ from the sediments. Both models are rather generic in nature, falling into the category of virtual experiments (Weiler and McDonnell 2004), and are not meant to actually represent the full complexity of the two field sites. For both models we use simplified boundary conditions and a loose calibration to $^{222}\text{Rn}$ field data with the purpose of capturing a realistic range of $^{222}\text{Rn}$ activities. Both models were calibrated using the automated calibration software PEST (Doherty and Hunt 2010) based on measured steady state $^{222}\text{Rn}$ depth profiles (Roter Main River) and a 12 day time-series of continuous $^{222}\text{Rn}$ data measured in bore 5 on the Avon River. Parameters that were subject to calibration are longitudinal $\alpha_L$ [L] and transverse $\alpha_T$ [L] dispersivities which are part of the hydrodynamic dispersion tensor $D$ and values for the saturated hydraulic conductivities $K_{sat}$ [L/T]. Model settings and final calibrated parameters are listed in Table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>HGS Parameters for the Avon River and Roter Main Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avon River Model</td>
</tr>
<tr>
<td></td>
<td>Roter Main Model</td>
</tr>
<tr>
<td>$\gamma$ [Bq/m$^3$/day]</td>
<td>540</td>
</tr>
<tr>
<td>$\lambda$ (per day)</td>
<td>0.18</td>
</tr>
<tr>
<td>$C_{\text{stream}}$ (Bq/m$^3$)</td>
<td>100</td>
</tr>
<tr>
<td>$K_{sat}$ (m/s) (calibrated)</td>
<td>$8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\alpha_L$ (m) (calibrated)</td>
<td>$1.78 \times 10^{-2}$</td>
</tr>
<tr>
<td>$\alpha_T$ (m) (calibrated)</td>
<td>$1.25 \times 10^{-2}$</td>
</tr>
<tr>
<td>$\Theta$</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>0.34</td>
</tr>
</tbody>
</table>
The steady state Roter Main River TSZ model is 20 m long and has an average thickness of around 0.6 m, and represents the hyporheic area below the river (Figure 1A). The vertical and horizontal grid resolution of the model domain varies between 2 and 10 cm. Although vertical groundwater inflow is common for the Roter Main River, we applied no flow boundaries for the lateral and lower boundaries of the model (Figure 1A). As our main intention is to highlight the potential of combining in-situ Rn measurements and numerical modeling and not to represent a specific field site as realistically as possible, we use no flow boundaries as simplifying assumptions. The upper model boundary represents the streambed interface with a spatially varying bed form topography that was measured using a theodolite during the field survey. Variations of surface heads along the streambed interface, which are an important control on hyporheic exchange processes, are explicitly simulated in HGS by solving the diffusive wave approximation to the St. Venant equations. The representation of surface heads in the model only accounts for hydrostatic pressure and gravitation but not for dynamical pressure components which would require a separate computational fluid dynamics (CFD) simulation. Dynamic pressure components were reported to be important especially for small-scale in-stream bedforms at the centimeter-scale such as rippled bedforms (Kessler et al. 2012; Azizian et al. 2015) but only play a subordinate role for structures at larger scales with a less pronounced bedform amplitude (Tonina and Buffington 2007) such as the presented pool-riffle section of the Roter Main. The exchange of water between the surface and subsurface flow domain utilizes the conductance concept that is implemented within HGS (Aquanty Inc. 2015). Under fully saturated conditions the exchange between the surface and subsurface flow domain \( \Gamma_o \) [L/T] is represented according to Equation 4. In Equation 4, \( K_z \) [L/T] is the vertical saturated hydraulic conductivity, \( h - h_0 \) [L] is the head difference between subsurface and surface and \( l_{exch} \) [L] is the user defined coupling length.

\[
\Gamma_o = K_z \frac{(h - h_0)}{l_{exch}}
\]

For the surface flow domain, lateral boundaries in the Roter Main model are implemented using constant head boundaries with a head difference of 0.1 m producing flow from left to right in the model (Figures 1A and 3A). For all computational nodes representing the streambed interface, \(^{222}\)Rn concentrations were fixed to the measured \(^{222}\)Rn in-stream activity (~200 Bq/m²). Initially, \(^{222}\)Rn activities for all nodes in the subsurface were set to the corresponding \(^{222}\)Rn activities in the river and simulations were carried out using a homogenous \(^{222}\)Rn emanation term \( \gamma \) (see Table 1). For the given set of transport and flow boundary conditions, the model was run until steady state was reached.

The model of the Avon River banks is 1000 m long and represents the transition from the stream channel (40 m) to the floodplain (Figure 1B). The vertical and horizontal grid resolution of the model domain varies between 25 cm close to the streambed interface and 4 m for the hinterland areas (Figure 1). In the model, the boundary located at \( x = 1000 \) m is a constant head boundary with an assigned constant head of 15.41 m (Figure 1B). This represents the long-term head gradient at the site. The lower no flow boundary condition corresponds to a basal clay layer that was found 24 m below the surface (Figure 1B). The effect of a variably gaining and loosing stream (bank storage) was represented in the model by applying a transient head boundary for the stream nodes (Figure 1B) based on a measured 12-day head record. To represent low \(^{222}\)Rn activities of infiltrating stream water, the river nodes were fixed to a constant value of 100 Bq/m³ during the entire simulation. This is similar to the values measured in the Avon river by Cartwright and Hofmann (2016). In the model, \(^{222}\)Rn emanation \( \gamma \) in the subsurface is constant (see Table 1) and initial conditions for head potentials and \(^{222}\)Rn concentrations result from a spin-up simulation that was run until the model reached steady-state base flow conditions.

For both models the mean water age distribution of subsurface water was simulated by using the travel time probability method implemented within HGS. This method estimates the mean age since infiltration for each element from the first moment form of an corresponding Advection-Dispersion-type equation (Aquanty Inc. 2015):

\[
- \nabla q \ A + \nabla \theta D \nabla A - q_o \ A + \theta = 0
\]

Here, \( (A) [T] \) represents the mean age of water since infiltration, \( q [L/T] \) and \( q_o [L/T] \) is the subsurface fluid flux and the exchange flux with the surface flow domain and \( \Theta [\cdot] \) is the sediment porosity. The mean water age is relative to a starting location, where the age signature is zero. The starting location here is where surface water infiltrates into the sediments which in the case of the Roter Main takes place across the streambed interface (Figure 1A) and for the Avon along the wetter perimeter of the river (Figure 1B). The mean age simulations for both models were performed based on the calibrated values listed in Table 1.

**Results**

**Steady-State TSZs: Hyporheic Flow Roter Main River**

The Roter Main River discharge at the time of measurement was 1 m³/s, which is below the yearly average, but about equal to the average base flow discharge. Flow had been steady for 3 weeks prior to sampling, with the last high discharge event on the 11th of June. \(^{222}\)Rn activities increased from approximately 200 Bq/m³ in the river water to 8700 Bq/m³ at 45 cm depth in the hyporheic zone (Figure 2). Insufficient sample could be removed from 50 cm depth for \(^{222}\)Rn measurements. The groundwater endmember assumed to be representative at 1 m below the streambed, had a \(^{222}\)Rn activity of 24,000 Bq/m³ and
is assumed to be in secular equilibrium with the hyporheic sediments.

The HGS model of the riffle-pool sequence was calibrated using PEST to the measured $^{222}$Rn profile using only the first five sampling points (to a depth of 25 cm) as we have found previously (Pittroff et al. 2017) that the deeper samples are influenced by upwelling regional groundwater which is neglected in our simulations. This can be seen by the rapid increase in $^{222}$Rn activities below 30 cm depth, as well as in other chemical parameters such as Cl$^-$. The optimized hydraulic conductivity ($K_{sat}$) was $3.4 \times 10^{-3}$ m/s and longitudinal and transverse dispersivities $2.13 \times 10^{-2}$ and $1.14 \times 10^{-2}$ m, respectively, which is on the lower side but consistent with a sandy aquifer with a length scales of 0.5–1 m (Table 1) (Schulze-Makuch 2005).

The HGS model reproduced the upper part of the measured $^{222}$Rn profile well ($r = 0.91$), with both the general trend of increasing $^{222}$Rn with depth and the absolute concentrations in the sediments being captured in the simulation (Figure 2). The modeled $^{222}$Rn activities in the 2D riffle-pool sequence show clearly the down- and upwelling zones within the hyporheic zone (Figure 3B), with young low $^{222}$Rn water in the downwelling areas and old, $^{222}$Rn rich water where upwelling dominates (Figure 3C). This is also reflected in the mean age calculated by HGS, with a range 0–5 days (Figure 3D). There is considerable heterogeneity in modeled $^{222}$Rn activities and water ages over the 20 m sequence, reflecting the complex flowpath response to the surface heads (Figure 3A). It also can be seen that the 2D simulation has areas with large contrasting activities due to the local up- and downwelling flow cells caused by the spatial variation of surface heads along the streamed interface. Such concentration gradients are unlikely to arise in a 1D model, but are an important part of dispersive mixing in 2D and 3D flow in the hyporheic zone (Hester et al. 2017). The modeled sequences also show features such as the “chimney effect” (Azizian et al. 2015) where deep old upwelling hyporheic water enters the stream with little interaction with shallow flow cells (at e.g., ~5 and 10 m). Chimney features are usually associated with pressure variations on the surface of the streamed interface, causing deeper water in the hyporheic zone to flow upwards and exfiltrate into the stream. They were reported for small-scale in-stream ripples (cm scale) Kessler et al. 2012 and (Azizian et al. 2015) as well as for riffle-pool structures at the meter scale (Frei et al. 2018).

HGS calculates the mean water age based on the advection-dispersion equation (Equation 5), while the $^{222}$Rn $\tau_{app.}$ model (Equation 2) is conceptually similar to the piston flow model. If the mean water age is treated as the reference age, we can test how well the simple $\tau_{app.}$ ingrowth model approximates water ages at steady state and systemically investigate how processes such as dispersion influence $\tau_{app.}$. Figures 3 and 4 show that (1) there is considerable scatter between the mean water ages calculated by HGS and $^{222}$Rn $\tau_{app.}$ both above and below the 1:1 line, and (2) the $\tau_{app.}$ tend to be shorter (below the 1:1 line) than the reference mean water ages. The ratio between the $\tau_{app.}$ and the mean water age of the base case scenario ranged between 0.1 and 8.4, with a mean of $0.76 \pm 0.1$ and median of 0.78. Deviations lying above the 1:1 line (Figure 4) are located close to the upper and right model boundaries (red and blanked areas in Figure 3). Values greater than one are physically impossible and can be considered as model artifacts that arise from the fact that different boundary conditions for the $^{222}$Rn (fixed concentration) and mean age (flux boundary) simulation were used. Two extra runs were conducted with identical conditions except for one run with $5 \times$ the base case dispersion ($5 \times D$) and the other with half the base case value ($0.5 \times D$). This shows clearly that the higher dispersion values increase the off-set from the 1:1 line to lower $\tau_{app.}$, an observation that has also been made using other apparent age tracers (Weissmann et al. 2002). What was unexpected is that the lower dispersion values ($0.5 \times D$) lead to more scatter between HGS ages and $\tau_{app.}$ than the highest dispersion values (Figure 4). In this case it appears that dispersive mixing reduces the concentration gradients in the TSZ and thus the scatter in the data.

**Transient Conditions: Bank Storage at the Avon River**

During the deployment of the OctoRad there was a single rain event of 30 mm in the Avon River catchment. This caused the river stage at Pearces Lane to rise by almost 1 m (Figure 5). The rise in stage caused a clear increase in groundwater heads close to the river. The head in bore 5 rapidly increased on the rising limb and decreased on the falling limb while the heads in bore 2, located 55 m from the river, rise less rapidly and fall more slowly than in bore 5 as expected for bank storage (Gerecht et al. 2011; Fovet et al. 2015). $^{222}$Rn activities decreased from around 5000 Bq/m$^3$ prior to the event to below the instrument detection limit over about 12 h.
As head levels declined, the $^{222}\text{Rn}$ activities increased rapidly over approximately 3 days and then increased more slowly for the remainder of the measurement period. $^{222}\text{Rn}$ activities had only reached 4000 Bq/m$^3$ when the measurements were ceased on eighth of November, 9 days after the start of the event.

**Modeling Bank Storage**

The HGS model of bank storage on the Avon was calibrated using the $^{222}\text{Rn}$ activities as the advective front of the infiltrating flood wave passed through bore 5 (Figure 1B). The correlation coefficient for the best fit model was $r = 0.89$, with $K_{\text{sat.}} = 8.4 \times 10^{-3}$ m/s, $\alpha_L = 1.78 \times 10^{-2}$ m and $\alpha_T = 1.25 \times 10^{-2}$ m. This is similar to the $K_{\text{sat.}}$ values measured by slug tests at the site ($0.5–5 \times 10^{-3}$ m/s). The model was able to capture both the decrease in $^{222}\text{Rn}$ activities as the advective front passed through the bore, and the increase of $^{222}\text{Rn}$ activities to near-background activities on the receding limb of the event as water flowed back toward the river.

There were some differences between modeled and measured $^{222}\text{Rn}$ activities during bank return flow, with the measured values being generally higher than modeled (Figure 5). It was also not possible to capture the flattening of measured $^{222}\text{Rn}$ activities 4 days into the receding limb. This may be due to modeled dispersion being too low so that mixing with the high $^{222}\text{Rn}$ regional groundwater was underestimated, heterogeneity in $K_{\text{sat.}}$ in the bank, or perhaps the most likely scenario is that the three dimensional flow fields in the gravel beds were not fully captured in the two dimensional simulation. Heterogeneity in the emanation rate within the gravel beds may also have contributed to the more rapid increase in measured $^{222}\text{Rn}$ values during bank return flow compared to modeled values. This is suggested by $^{222}\text{Rn}$ measurements from bores 2 and 3 being up to 15,000 Bq/m$^3$. It would possible to construct a 3D
Comparing the HGS mean water ages with $\tau_{app}$ from the $^{222}$Rn ingrowth model gives a more complex picture than the Roter Main steady state simulation (Figure 6). The most obvious effect is the systematic underestimation of $\tau_{app}$ compared to the HGS mean ages. This occurs along the advective front (Figure 7B and 7C) where high $^{222}$Rn regional groundwater and young low $^{222}$Rn river water mix within the bank (points far off the 1:1 line in Figure 7A), increasing activities above ingrowth alone. This effect has been well described for other age tracers such as $^{14}$C where very old water from aquitards mixes with aquifer water (Bethke and Johnson 2002). The difference between linear mixing and exponential decay or ingrowth processes leads to a large deviation of apparent ages from mean water age. Away from the advective front toward the stream, there is a very good agreement between the water ages estimated from the two models, with values lying very close to the 1:1 line (Figure 7A). Within the mixing zone between stream and groundwater (Figure 7B and 7C) there are some spatially limited areas located close to the model boundaries that show a ratio between $\tau_{app}$ and the HGS mean age above 1 (red areas in Figure 7). In Figure 7A, these points lie above the 1:1 line and as described previously for the Roter Main model can be explained by model artifacts that arise due to the different boundary conditions applied for the $^{222}$Rn and mean age simulations.

Discussion

For the two test cases, Roter Main and Avon River, we have demonstrated how the combination of in-situ $^{222}$Rn measurements and numerical modeling can be used to mechanistically understand TSZ’s under steady state and transient flow conditions. In the steady state example from the hyporheic zone we noticed a systematic underestimation of water ages using the apparent age model. This was related to dispersive effects and increased from $0.5 \times D$ to $5 \times D$. Similar effects on radioactive tracer transport and dating have also been observed for $^3$H in simple 1D systems, but over much larger time scales due to $^3$H’s longer half-life. One important difference between $^{222}$Rn and these previous studies is that the input of $^{222}$Rn is constant whereas more conventional tracers have defined input shapes such as the bomb pulse for $^3$H or exponential increase for SF$_6$ and CFCs (Schlosser et al. 1989; Weissmann et al. 2002). The effect of dispersion will be dependent on the length scale as well as the concentration gradient, and thus it is conceivable that the error incurred by using the apparent age model increases with flowpath length. The error incurred by the apparent age model will also be exacerbated where water of highly different concentrations are in close proximity. This is
likely the reason that the deviation between HGS mean water ages and the apparent age model ages are largest at the edges of the upwelling zones. Such effects would be increased during paraflood flow or bank storage where flow paths tend to be tens of meters in length and where TSZ water abuts regional groundwater with high 222Rn activities. This can be seen during bank storage on the Avon River where the HGS model computed a systematic underestimation of apparent ages, with the underestimation at the advective front being particularly large. The incorporation of 222Rn into HGS allows for complex flow processes to be captured and evaluated and thus is a useful tool for interpreting 222Rn tracer data. For example, it is possible extend the model to 3D flow fields, or to model saturated-unsaturated flow processes, which are important during over bank and return flow processes.

Galser et al. (2016) suggests that we still lack approaches that include spatially distributed data in the evaluation of physically based models that are supposed to represent spatiotemporal processes ranging from a few centimeters up to the scale of entire catchments. Although both the cases here only represent generic cases of real world conditions, including 222Rn measurements into model calibration shows a high potential to further constrain the nonuniqueness of hydrological models and to get an improved mechanistic understanding of hydrological systems on time scales of hours to days. 222Rn is unique compared to other tracers as it is emitted from the aquifer matrix, and as a radioactive isotope, has an explicit and well-defined time scale for emanation and decay. However, it is also clear from previous studies that calibration to a single field measurement type leads to significant uncertainty in model parameter uniqueness. The more field parameters that can be incorporated into the calibration and validation process the lower the uncertainty in the model results and more unique the individual parameters will be (Sanford 2010). 222Rn offers a new and powerful tracer that can be used explicitly within the model calibration process, but this should not be to the exclusion of other calibration data that can be measured in the field, such as hydraulic conductivity, pressure head, and other chemical and radioactive tracers.

Representation of 222Rn in subsurface hydrological systems requires a model that is capable of representing advective and dispersive transport processes as well as 222Rn-decay and -emanation in the subsurface. Many

![Figure 6. HGS simulation of mean water ages and 222Rn activities in the Avon floodplain. (A-E) Time series of water age and 222Rn activities during a bank storage event on the Avon River. Red numbers indicate contour lines for mean water age.](image-url)
numerical hydrological models are capable of representing reactive solute transport with simple first order reaction kinetics. However, the inclusion of a constant source term that can be used to account for $^{222}\text{Rn}$-emanation is not often the case. Schilling et al. 2017 have recently used HGS to understand and incorporate geochemical data including $^{222}\text{Rn}$ into the model calibration processes when simulating bank filtration. While a novel approach, these authors have not modeled the tracers explicitly (i.e., via the advection-dispersion-reaction equation), instead using the dynamic mixing cell approach and then (in the case of $^{222}\text{Rn}$) the apparent age method to quantify water ages. This will likely have led to errors where dispersion is an important part of the mixing process, especially if there is a component of regional groundwater in the samples or where concentration gradients exist as water of various ages converge on the pumping well. The zero-order source term, that is implemented in HGS and used to account for $^{222}\text{Rn}$-emanation, could also be extended to other radioactive tracers that are emitted from the aquifer matrix such as $^4\text{He}$ and $^{37}\text{Ar}$ (Schilling et al. 2017). In evaluating hydrological models that are capable of solving surface and subsurface flow processes simultaneously a major challenge is still to extend $^{222}\text{Rn}$-transport to the surface flow domain. In surface flow $^{222}\text{Rn}$ is lost not only by radioactive decay, but additionally is subject to degassing to the atmosphere which represents a major sink in surface water bodies. Contrary to emanation, degassing cannot be represented by a simple constant sink term as it highly depends on the characteristic of surface flow such as the degree of turbulence. This point remains to be addressed in future model developments.

**Acknowledgments**

This work was funded using the internal funds of the Limnological Research Station, University of Bayreuth as well as the National Centre for Groundwater Research and Training (NCGRT). Three reviewers and the editor are thanked for improving previous versions of the manuscript.

**Authors’ Note**

The authors do not have any conflicts of interest or financial disclosures to report.

**References**


Boulton, A.J. 1993. Stream ecology and surface-hyporheic hydrologic exchange: Implications, techniques and


