Hydrogeology Journal
https://doi.org/10.1007/s10040-021-02435-8

PAPER

Estimating groundwater mean transit time from SF_6 in stream water: field example and planning metrics for a reach mass-balance approach

Craig R. Jensen[1] · David P. Genereux[1] · Troy E. Gilmore[2,3] · D. Kip Solomon[4] · Aaron R. Mittelstet[3] · C. Eric Humphrey[4] · Markus R. MacNamara[1] · Caner Zeyrek[2] · Vitaly A. Zlotnik[5]

Received: 26 March 2021 / Accepted: 29 November 2021

Abstract

Stream reach mass-balance can be used to estimate groundwater age at scales larger than individual wells. Atmospherically derived SF_6 was used as an age-dating tracer in reach mass-balance analysis of six stream reaches in Nebraska, USA. The goal was to estimate the flow-weighted mean SF_6 concentration in the groundwater discharge to each reach (C_{gw}) and the groundwater mean transit time (MTT) based on C_{gw}. The C_{gw} in one reach (1.48 fmol/kg) was consistent with a MTT of 27 years. High uncertainty in C_{gw} prevented plausible MTT estimates in the other reaches. Monte Carlo analyses indicated that low rates of groundwater inflow to the stream reaches (Q_{gw}) led to high uncertainty in C_{gw}. This issue can affect the success of any reach mass-balance analysis, for age dating tracers or other solutes. To improve the likelihood of success in future applications of reach mass-balance, two new planning metrics are proposed for use, separately or together: the ratio of groundwater recharge rate (R) to stream density (D) within a watershed, and the long-term average Q_{gw}/Q_d in a stream reach (where Q_d is stream discharge at the downstream end of the reach). In comparison to the Nebraska study site ($R/D = 107$ m^2/year), more accurate estimation of C_{gw} was achieved in a previous SF_6 study in North Carolina (USA) with greater R/D (231 m^2/year). Larger Q_{gw}/Q_d is associated with more accurate outcomes in reach mass-balance, and long-term average Q_{gw}/Q_d in stream reaches may be estimated by GIS analysis of watershed areas and/or stream lengths.

Keywords USA · Groundwater/surface-water interaction · Age-dating tracers · Field techniques · Sulfur hexafluoride · High Plains Aquifer

Introduction

Groundwater age is a fundamental variable in hydrogeology that can inform understanding of groundwater flow rates, recharge rates, and contaminant transport. In surficial (unconfined) aquifers, groundwater age has been estimated with a variety of atmospherically derived age-dating tracers suitable for dating groundwater back to roughly the mid-twentieth century: chlorofluorocarbons (Busenberg and Plummer 1992), ^{3}H/^{3}He (Poreda et al. 1988; Schlosser et al. 1988; Ekwurzel et al. 1994; Cook and Solomon 1997), and sulfur hexafluoride (Busenberg and Plummer 2000; Zoellmann et al. 2001; Katz 2004; Alvarado et al. 2005; Delin et al. 2007; Sanford et al. 2015; Lauffenburger et al. 2018). Typically, these tracer methods are applied to estimate the ages of groundwater samples collected at wells, giving ages at individual points in a groundwater system. Vertical profiles of such ages have been used to estimate recharge rates or groundwater mean transit time (MTT; Böhlke and Denver 1995; Alvarado et al. 2005; Chacha et al. 2018), provided that a groundwater transit time distribution is assumed (transit time refers to the groundwater travel time from recharge at the water table to discharge at a stream bed or other discharge face).

✉ Craig R. Jensen
crjense2@ncsu.edu

[1] Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC, USA

[2] Conservation and Survey Division, School of Natural Resources, University of Nebraska, Lincoln, NE, USA

[3] Biological Systems Engineering Department, University of Nebraska, Lincoln, NE, USA

[4] Department of Geology and Geophysics, University of Utah, Salt Lake City, UT, USA

[5] Department of Earth and Atmospheric Sciences, University of Nebraska, NE, Lincoln, USA

Published online: 06 January 2022

Stolp et al. (2010) and Solomon et al. (2015) introduced a new approach to applying these atmospherically derived age-dating tracers at a larger scale, in streams rather than wells. This "reach mass-balance" approach involves formulating a 'conservation of mass' equation for a tracer in a reach (a section) of a stream channel, and measuring all variables in the equation except one, which is then computed: the flow-weighted mean tracer concentration in the groundwater discharge to the stream reach (C_{gw}). Mean groundwater age is then estimated from C_{gw}. Measurements required include the tracer concentrations and stream flow at the upstream and downstream ends of the stream reach, and the rate of gas exchange between the stream water and overlying air (for a gaseous tracer).

Stolp et al. (2010) used the reach mass-balance approach for tritiogenic ^{3}He, in order to apply the ^{3}H/^{3}He age dating method at reach scale, and Solomon et al. (2015) used it with SF_6. In both cases the calculated tracer concentration was used to determine the mean age of the groundwater discharging into a stream reach. Because the mean age calculated from reach mass-balance inherently flow-weights all groundwater discharge to the stream reach, this age is the MTT of the groundwater discharge to the reach (Stolp et al. 2010; Solomon et al. 2015).

Calculation of C_{gw} by reach mass-balance has been applied to solutes other than groundwater age-dating tracers such as ^{222}Rn (Genereux and Hemond 1990; Genereux et al. 1993), toluene and other volatile organic compounds (VOCs) (Kim et al. 1995; Kim and Hemond 1998), nitrate (Gilmore et al. 2016a), and heavy metals (Kimball et al. 2001). The general concept has also been used to estimate solute mass loading to stream reaches (mass/time) without explicit calculation of C_{gw} (Holtzman et al. 2005; Banks and Palumbo-Roe 2010). These efforts to determine C_{gw} or mass loading required knowledge of Q_{gw}, the rate of groundwater discharge into a stream reach. Cook (2013) provides a thorough review of the opposite problem: determining Q_{gw} from C_{gw}. The two problems, finding C_{gw} from Q_{gw} and Q_{gw} from C_{gw}, share a common conceptual and mathematical basis and the goal of detecting a groundwater "signal" in a stream.

The reach mass-balance approach to estimation of MTT has different strengths compared to sampling at wells (Gilmore et al. 2016a). This includes sampling over a larger support volume to obtain a MTT value at km scale as opposed to the point scale of an individual well, and lower analytical costs because MTT is determined from sampling at only two points in the stream (the ends of the reach) rather than numerous wells. These advantages are based on the integrative nature of groundwater discharge to streams and make reach mass-balance a potentially powerful tool that merits development.

While reach mass-balance analysis of age-dating tracers to determine groundwater MTT holds promise for understanding groundwater transit times at larger scales, few studies have utilized this method, namely Stolp et al. (2010) using ^{3}H/^{3}He in Austria and Solomon et al. (2015) using SF_6 in North Carolina, USA. The study reported here is an additional evaluation of the approach in the semiarid Sand Hills area of western Nebraska, USA. Sulfur hexafluoride was used as the age-dating tracer. The goals were to quantify C_{gw} and groundwater MTT to several stream reaches, compare results with prior similar efforts, and analyze sources of uncertainty. In addition, new concepts and tools are proposed for planning successful reach mass-balance work at new field sites.

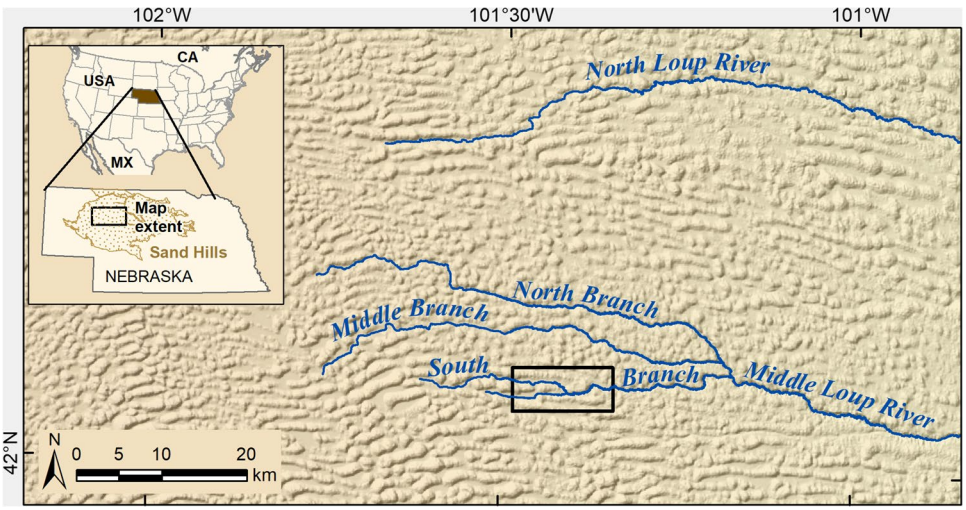

Fig. 1 Location of the study area in the Nebraska Sand Hills, USA. The black rectangle indicates the domain of the following figure, in the watershed of the South Branch Middle Loup (SBML) River

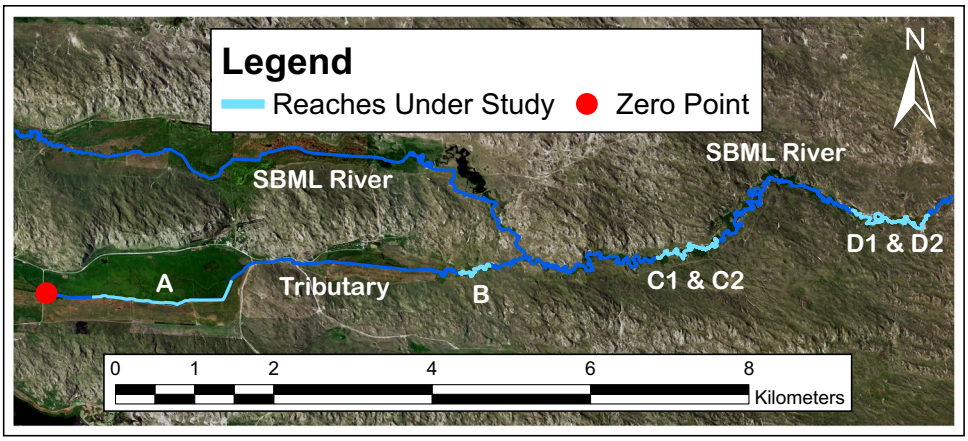

Fig. 2 The study site included the South Branch Middle Loup (SBML) River and an unnamed tributary of the river. A culvert in the tributary was assigned as an arbitrary zero-point for the measurement of distance along the channel. Flow is from west to east

Study site

The study site is the watershed of the South Branch Middle Loup (SBML) River. The 271 km^2 watershed lies in the Nebraska Sand Hills, an aeolian dune field covering approximately 58,000 km^2 (Fig. 1; Qi 2009). The dunes reach up to 100 m above the surrounding valley floors and are responsible for complex groundwater flow patterns (Gosselin et al. 1999). Within the historical record, the Sand Hills have been stabilized by grass; however, the geologic record shows multiple instances of vegetation loss in the last 15,000 years that correlated with destabilization and migration of the dunes (Loope and Swinehart 2000; Billesbach and Arkebauer 2012).

Average annual precipitation and evapotranspiration in the Sand Hills are 44.8 and 41.6 cm, respectively (Chen et al. 2019). Evapotranspiration is greatest in the "wet meadows", which are drained wetland valleys with significant groundwater discharge to streams and/or uptake by plant roots (Gosselin et al. 1999). The dunes and dry valleys are areas of groundwater recharge (Gosselin et al. 1999). Some of this recharge replenishes the regional flow system in the High Plains aquifer, which underlies the Sand Hills (Szilagyi and Jozsa 2012). Regional groundwater flow is west to east (Miller and Appel 1997).

The work was carried out at Gudmundsen Sandhills Laboratory (GSL), a cattle ranch owned and operated by the University of Nebraska-Lincoln. Sampling and measurements were made at two locations on the SBML River and two locations on an unnamed tributary (Fig. 2). Near a culvert designated as the zero-point for measuring distance along the channel (Fig. 2), the tributary was straightened and deepened into a drainage ditch in the 1930s to drain the wet meadow at GSL for hay production (Gosselin et al. 1999). This ditch transitions to a natural meandering channel roughly 2.7 km upstream of the confluence with the SBML River. Previous research by Gosselin et al. (2006) suggests groundwater flows to the tributary under the dune on the south side of the valley (i.e., this dune is a topographic divide but not a groundwater divide), while the northern dune acts as a groundwater divide between the tributary and the SBML River. The SBML River meanders strongly as it winds through a valley between the dunes, and oxbows (cutoff meanders) are common.

Reach mass-balance SF_6 age-dating was performed in six reaches in four areas of the channel system (Fig. 2). Reaches A, B, C1, and D1 were studied in May 2019, while reaches C2 and D2 were studied in August 2019. Reaches C2 and D2 are located in roughly the same locations as reaches C1 and D1, respectively (Table 1).

Methods

Reach mass-balance

At baseflow, the SF_6 concentration in stream water (after correction for gas exchange with the air) represents the flow-weighted mean concentration of SF_6 in the groundwater discharging to the stream (C_{gw}). Further, the MTT of groundwater discharge to the stream can be estimated from C_{gw} (Solomon et al. 2015). C_{gw} was computed using a reach mass-balance approach:

$$C_{gw} = \frac{Q_d C_d - Q_u C_u - \tau k_{SF6} \overline{Q}(\overline{C}_{eq} - \overline{C})}{Q_{gw}} \quad (1)$$

where Q is the volumetric flow rate of the stream, Q_{gw} is equal to $Q_d - Q_u$ (subscripts "d" and "u" indicate the downstream and upstream ends of the stream reach, respectively), C is the SF_6 concentration in the stream water, C_{eq} is the SF_6 concentration the stream water would have if it were at equilibrium with the atmosphere at the temperature of the stream water, k_{SF6} is the first-order gas exchange rate constant for SF_6 in the stream reach ($time^{-1}$), the overbar

Table 1 Information on the six study reaches, including field measurements and calculated values used in Eq. (1)

Parameter	Reach					
	A	B	C1	C2	D1[b]	D2[b]
Study date in 2019	19 May	31 May	27 May	6 Aug	26 May	8 Aug
Reach location (km)[a]	0.6–2.6	6.0–7.0	12.4–14.0	13.3–14.2	19.4–21.8	20.5–21.9
Q_u: stream flow at upstream SF_6 station (L/s)	27	203	1,060	171	1,567	184
Q_d: stream flow at downstream SF_6 station (L/s)	61	222	1,135	185	1,899	199
Q_{gw}: groundwater discharge (L/s)	34	19	75	14	332	15
Q_{gw}/Q_d	57%	8%	7%	7%	17%	7%
C_u: upstream SF_6 concentration (fmol/kg)	3.06	2.10	2.29	1.91	2.31	1.60
C_d: downstream SF_6 concentration (fmol/kg)	2.43	2.11	2.39	1.72	2.34	1.60
C_{eq}: equilibrium SF_6 concentration (fmol/kg)	3.60	2.72	2.46	2.29	2.61	2.30
τ: travel time (min.)	765.4	76.3	65.0	56.3	69.0	55.8
Mean stream depth (m)	0.34	0.53	0.63	0.22	0.87	0.24
k_{C3H8} (day^{-1})	12.1	12.1	5.3	12.1	21.3	27.2
v_{C3H8} (m/day)	4.1	6.4	3.3	2.7	18.7	6.6
k_{SF6} (day^{-1})	8.2	8.4	3.7	8.5	14.6	19.1
v_{SF6} (m/day)	2.8	4.4	2.3	1.9	12.7	4.6
α	5.4	5.1	2.4	4.3	3.6	9.5

[a]Location in km downstream from the culvert at the zero-point (Fig. 2)

[b]There is a 140-m oxbow between 20.96 and 21.10 km downstream from the zero-point that is cut off from the SBML River. This oxbow is not included in the total reach length

Table 2 Uncertainty in the input variables for calculating C_{gw} with Eqs. (1) and (3). Error is reported as a percent of the best estimate. Variables are defined in the text below Eqs. (1) and (3). Details are in section S2 of the ESM

Variable	Uncertainty average	Uncertainty range
C	5.0%	5.0%
C_{eq}	2.4%	1.1–5.5%
D_{C3H8}	3%	3%
D_{SF6}	3.6%	3.6%
k_{C3H8}	19.2%	5.9–53.2%
k_{SF6}	21.8%	10.2–53.9%
n	21.4%	21.4%
τ	0.5%	0.1–0.6%
Q	2.7%	0.5–8.4%
Q_{gw}	31.8%	14.8–55.7%

indicates a reach average value equal to the average of the values at the upstream and downstream ends of the reach, e.g., $\overline{Q} = (Q_u + Q_d)/2$, and τ is the stream water travel time through the reach. This equation was derived from the general steady-state, one-dimensional (1D) transport equation for a volatile (gaseous) dissolved solute in a stream (Genereux and Hemond 1990; Cook 2013; Solomon et al. 2015).

Data collection

Each study reach was defined by upstream and downstream measurement stations at the ends of the reach (Table 1). Solving for C_{gw} with Eq. (1) required determination of four variables at each measurement station: Q, C, travel time to the station from a tracer injection location upstream of the reach, and the concentration of an injected gas needed to determine k_{SF6}. The main methods are summarized in the following, along with the uncertainties of the variables needed to apply Eq. (1) (Table 2). The electronic supplementary material (ESM) has additional details on measurements (section S1 of the ESM) and uncertainty (section S2 of the ESM).

Q was determined by standard salt-dilution stream gauging (e.g., Kilpatrick et al. 1989; Oviedo-Vargas et al. 2015); 2–6 kg of high purity NaCl were dissolved in 7–30 L of stream water in 19-L buckets and instantaneously released over the width of the stream channel at the tracer injection location. The conductivity breakthrough curve at each measurement station was measured with a HOBO© conductivity data logger. The travel time τ through the reach was determined as the difference between the times to peak conductivity on the breakthrough curves at the upstream and downstream stations. All conductivity data were temperature corrected to 25 °C using established methods (US Geological Survey 2019).

Stream water samples for SF_6 analysis were collected at the upstream and downstream ends of each study reach in 1-L plastic-coated amber glass bottles (US Geological Survey 2018; Busenberg and Plummer 2000; Solomon et al. 2015). A submersible pump with outlet tubing at the bottom of the bottle was used to flush and overflow each bottle multiple times before it was capped.

The value of k_{SF6} was determined with widely used gas exchange methods (Kilpatrick et al. 1989; Genereux and Hemond 1992; Parker and DeSimone 1992; Choi et al 1998; Jones and Mulholland 1998; Hope et al 2001; Wallin et al. 2011; Oviedo-Vargas et al. 2015; Natchimuthu et al. 2017). At the tracer injection site upstream of each reach, high purity 99.0% propane (Airgas Inc.) was continuously bubbled into stream water from a standard 20 lb (9 kg) liquified propane tank using fine-pore diffusers at 165–179 kPa. In each reach, the gas exchange rate constant for propane (k_{C3H8}) was calculated as:

$$k_{C3H8} = \frac{1}{\tau}\ln\left(\frac{[C_3H_8]_u Q_u}{[C_3H_8]_d Q_d}\right) \quad (2)$$

where $[C_3H_8]$ is the concentration of propane in the stream water (Wallin et al. 2011; Oviedo-Vargas et al. 2015; Natchimuthu et al. 2017).

It was assumed that reaching steady-state propane concentrations at each measurement station would require steady propane injection for a period equal to at least four times the travel time between the injection site and the measurement station (Kilpatrick et al. 1989). However, in cases where large eddies were present in the stream, propane delivery was extended up to 14 times the travel time to better ensure steady state. Stream water propane analyses were conducted on-site by gas chromatography (see ESM).

Values of k_{C3H8} were converted to k_{SF6} using a well-known relationship (e.g., Jähne et al. 1987; Cook et al. 2003):

$$\frac{k_{SF6}}{k_{C3H8}} = \left(\frac{D_{SF6}}{D_{C3H8}}\right)^n \quad (3)$$

where D is the aqueous diffusion coefficient and n is a constant between 0.5 (turbulent conditions) and 1 (non-turbulent conditions). The value of n was 0.7, the mean of published experimentally-determined values (see ESM). D_{SF6} and D_{C3H8} are temperature dependent and were calculated based on the measured stream-water temperature and empirical equations from King and Saltzman (1995) and Wise and Houghton (1966), respectively. The ratio of k_{SF6} to k_{C3H8} ranged from 0.65 to 0.71 in May 2019 and from 0.60 to 0.62 in August 2019.

C_{eq} for use in Eq. (1) was computed from data on stream-water temperature from the HOBO data loggers at the ends of each reach and data on the SF_6 atmospheric mixing ratio of modern air from a monitoring station (at Niwot Ridge, Colorado) of the Earth System Research Laboratory (ESRL) at the US National Oceanic and Atmospheric Administration (NOAA 2020). Niwot Ridge, which has data from December 1994 to the present, is the closest SF_6 monitoring station to GSL and is generally upwind. For each measurement period, 5 months of NOAA data on SF_6 atmospheric mixing ratios were averaged to reduce the effect of any outliers: March–July 2019 data for the May 2019 reach mass-balance work, June–October 2019 for the August 2019 work. A modified Henry's Law was used to convert the mixing ratio to the C_{eq} value for each stream reach (Weiss and Price 1980):

$$C_{eq} = K_H x^0_{SF6}\left(P_{atm} - p_{H2O}\right) \quad (4)$$

where K_H is the Henry's Law constant for SF_6 (mol kg^{-1} atm^{-1}), P_{atm} is the atmospheric pressure at the field site (atm), p_{H2O} is the water vapor pressure (atm), and x^0_{SF6} is the atmospheric mixing ratio for SF_6, for the May 2019 or August 2019 reach mass-balance work (10.086 or 10.126 parts per trillion by volume, respectively). K_H and p_{H2O} are temperature dependent and were calculated using the temperature data from the HOBO data loggers and empirical equations from Bullister et al. (2002) and Weiss and Price (1980), respectively. P_{atm} was estimated based on elevation (List 1949; see ESM).

In three reaches, groundwater samples were collected beneath the streambed for analysis of SF_6: reaches B (28 May), C1 (27 May), and D1 (19 May). The groundwater samples were not necessary for the reach mass-balance analysis (Eq. 1), but measured SF_6 concentrations in the samples provided a point of comparison for C_{gw} values calculated for those three reaches with Eq. (1). The groundwater samples were collected with streambed piezometers inserted 50 cm into the streambed using methods outlined by Gilmore et al. (2016b). The groundwater seepage rate (specific discharge) was measured at each groundwater sampling point using automated tube seepage meters (Solomon et al. 2020), allowing calculation of the flow-weighted mean groundwater SF_6 concentration for the 4–6 groundwater samples collected within a given reach (the SF_6 concentration and associated age at each streambed sampling point were weighted by the specific discharge at the same point).

Age dating: estimating MTT from C_{gw}

Henry's Law was used to relate C_{gw} to x_{SF6} (Busenberg and Plummer 2000):

$$x_{SF6} = \frac{C_{gw}}{K_H\left(P_{atm} - p_{H2O}\right)} \quad (5)$$

where x_{SF6} is the atmospheric SF_6 mixing ratio that would be at equilibrium with the computed C_{gw} value. K_H and p_{H20} were calculated using the average groundwater recharge temperature at the study site, 11.4 °C. Average recharge temperature at the site was estimated by applying noble gas thermometry to groundwater samples collected during the same time period as the reach mass-balance work reported here (unpublished data). P_{atm} was estimated based on elevation (List 1949; see ESM).

For each x_{SF6} value from Eq. (5), the "forecasting" function in the USGS Excel workbook TracerLPM (Jurgens et al. 2012) was used to find the groundwater MTT that would give that observed x_{SF6} value in the decimal year 2019.5 (the approximate time of field sampling in 2019), for the exponential mixing model (EMM). The use of EMM seems reasonable and physically plausible for a sandy unconfined aquifer without impervious surface to impede recharge, given the integrative nature of reach mass-balance with all groundwater flowpaths contributing to a stream reach in proportion to their flow rates. The data used for the SF_6 atmospheric curve in TracerLPM came from two sources: SF_6 data from the NOAA ESRL Niwot Ridge station (1995–2020) and the Maiss and Brenninkmeijer (1998) values preloaded in TracerLPM (1950–1994).

Reach mass-balance results

Groundwater input to the six study reaches ranged from 7 to 57% of Q_d and the gas exchange velocity v_{SF6} (the product of k_{SF6} and mean stream depth in the reach) ranged from 1.9 m/day to 12.7 m/day (Table 1). Among five of the reaches (A, B, C1, C2, and D2) v_{SF6} had a smaller range of 1.9–4.6 m/day which is within the range in the literature (see ESM). The high v_{SF6} in reach D1 (12.7 m/day) occurred during high stream flow on 26 May 2019, when large eddies along with active destabilization and erosion of stream banks were observed. It seems possible that these conditions promoted greater gas exchange between the stream water and atmosphere.

Successful application of the reach mass-balance approach to estimating groundwater MTT requires the stream water and overlying air not be in equilibrium with regard to SF_6, i.e., that $C < C_{eq}$, indicating at least partial retention of the "groundwater age signal" in the stream water (Solomon et al. 2015). This criterion was met in all cases (Table 1), an encouraging sign for reach mass-balance estimation of groundwater MTT.

The SF_6 concentrations in stream water samples were converted to ages using Eq. (5) and TracerLPM to give the minimum apparent age of each sample, 2–24 years based on the EMM (Fig. 3). Conceptually, this minimum groundwater age computed for stream-water samples reflects the flow-weighted mean age of all the groundwater discharged into the SBML River and its tributaries upstream of the point of stream sample collection. It is a minimum because it does not account for gas exchange between the stream water and air. Such gas exchange would decrease the apparent groundwater age; water fully equilibrated with the atmosphere with respect to SF_6 concentration would have zero age.

The observed partial retention of the groundwater age signal in the stream water is consistent with the values of α (Table 1), defined as v_{SF6}/v_{gw}, where v_{gw} is the specific discharge of groundwater into the stream (Cook 2013; Sanford et al. 2015; Solomon et al. 2015). For a wide shallow channel, v_{gw} is approximately Q_{gw} divided by the length and width of the stream reach. At greater α, more of the groundwater age signal is lost by gas exchange. Solomon et al. (2015) advised that α should be less than 10 so that at least 10% of the groundwater signal

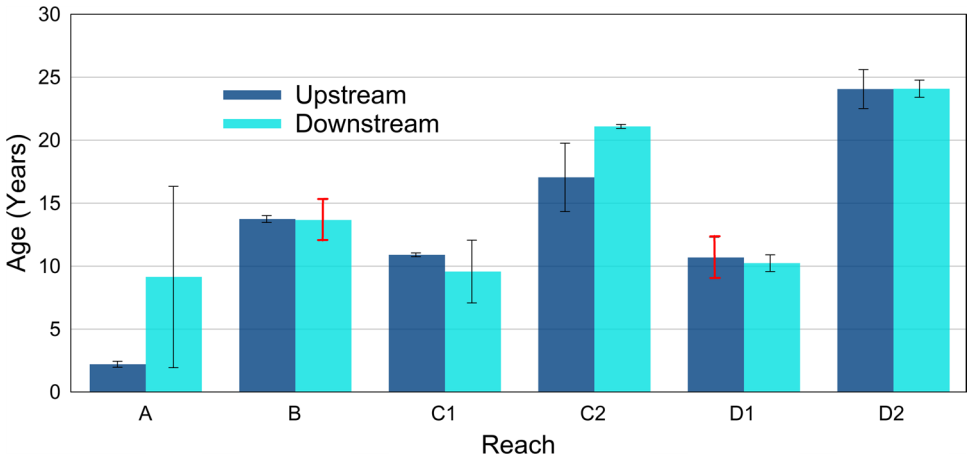

Fig. 3 The minimum apparent groundwater age based on stream-water samples analyzed for SF_6. Each bar represents the average of two samples collected at the site. The error bars represent the sample standard error of the mean. In two cases, the downstream end of reach B and the upstream end of reach D1, only one stream water sample was available for SF_6 analysis and ±1.61 years is shown as a red error bar (the overall average sample standard error of the mean was 1.61 years)

is retained by the stream, a criterion met by the study reaches (Table 1).

C_u and C_d less than C_{eq} (Table 1), nonzero minimum ages (Fig. 3), and $\alpha < 10$ (Table 1) were three positive indicators for the potential success of reach mass-balance. However, the C_{gw} and MTT results from reach mass-balance in the six reaches included more implausible than plausible results. The estimated C_{gw} in reach D1 was 1.48 fmol/kg, corresponding to a plausible MTT of 27 years. It was not practical to account for groundwater excess air in this calculation, and the estimated MTT could be low by a few years if typical amounts of excess air were present (1–3 cc/kg at STP). The estimated C_{gw} of reach C1 was 3.50 fmol/kg, which is greater than the expected C_{eq}, an implausible result that would suggest age in the near future. Estimated C_{gw} values for the remaining reaches A, B, C2, and D2 were negative, which is not a physically meaningful result; thus, in spite of the three positive indicators previously listed, the reach mass-balance analysis (Eq. 1) led to unrealistic results in most cases. This motivated an in-depth analysis of uncertainty in the C_{gw} values calculated through reach mass-balance.

Monte Carlo analysis

Uncertainty in C_{gw}

A Monte Carlo analysis was performed to determine the approximate "one-sigma" confidence interval for the C_{gw} of each reach, meaning the limits that would contain the central 68% of the area of each Monte Carlo histogram (Fig. 4). For each reach, each variable in Eqs. (1)–(3) was varied randomly over 15,000 iterations, producing 15,000 simulated C_{gw} values that were then binned and plotted in a histogram. The lowest and highest 2400 C_{gw} values (lowest and highest 16% of the 15,000) defined the boundaries on the middle 10,200 values, and thus the lower and upper 68% confidence limits of C_{gw}. The NORM.INV function in Microsoft Excel was used to randomly vary each variable around its mean (the best estimate of the variable), within

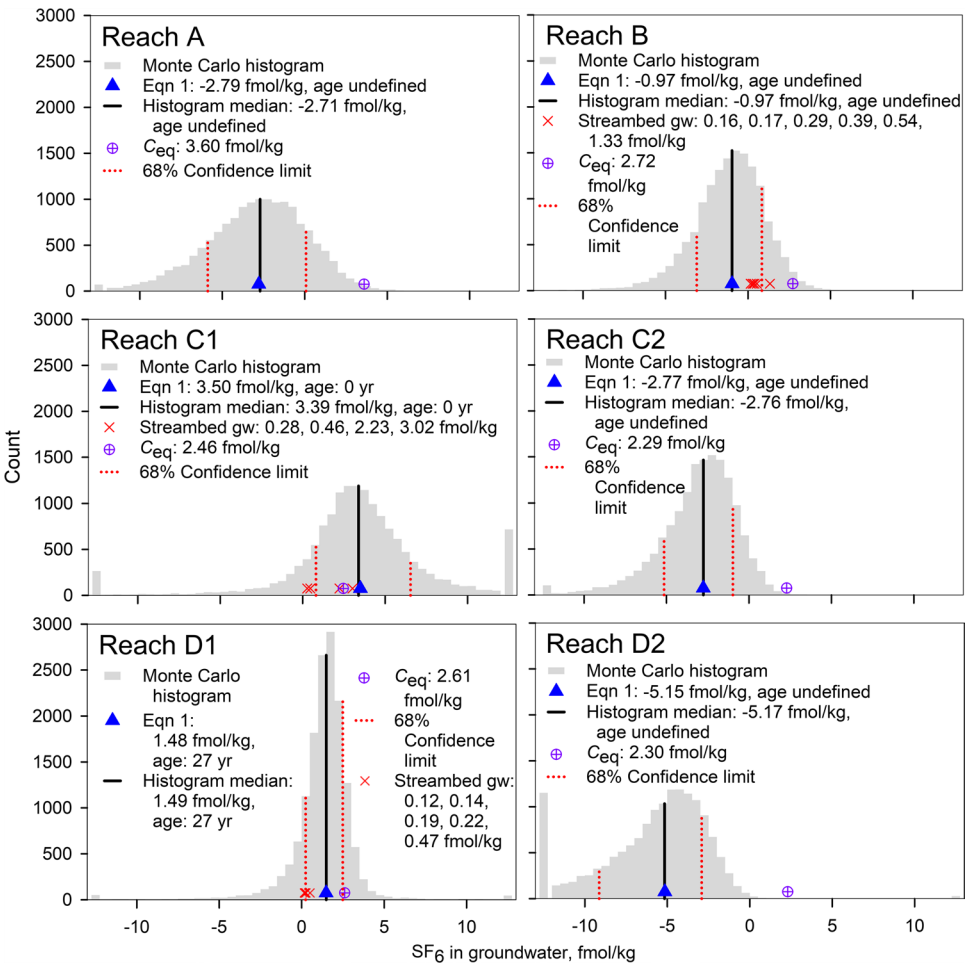

Fig. 4 Histograms of Monte Carlo results (15,000 iterations) for the SF_6 concentration in groundwater discharge to each stream reach, based on reach mass-balance. C_{gw} calculated from Eq. (1) using best estimates of all input variables is shown along with the median simulated C_{gw} from the Monte Carlo analysis. The 68% confidence interval from Monte Carlo analysis is shown. C_{eq} is the stream SF_6 concentration that would have been in equilibrium with the atmosphere at the stream water temperature in the reach. For five of the histograms (all except for reach B), simulated C_{gw} values that extended beyond the lower and/or upper limit of the SF_6 axis were accumulated into "overflow bins" on the left and/or right sides of the plot (at concentrations of 12.5 and –12.5 fmol/kg)

a normal distribution defined by the mean and an estimated standard deviation (the uncertainty in Table 2).

On average for the six reaches, the median simulated C_{gw} from Monte Carlo analysis differed by only 0.04 fmol/kg from the C_{gw} value estimated by direct application of Eq. (1) with best estimates of the input parameter values, suggesting that 15,000 iterations were sufficient in the Monte Carlo analysis. However, 68% confidence intervals for C_{gw} were all quite wide (Fig. 4), indicating relatively high uncertainty in C_{gw}. The 68% confidence interval for the Monte Carlo estimate of C_{gw} included physically impossible negative values for every reach except reach D1, and implausibly high values above C_{eq} for one reach (C1).

A meaningful groundwater age can be calculated if C_{gw} falls between 0 and C_{eq}. For a given Monte Carlo analysis, the percentage of simulated C_{gw} values that fall in this range represents the likelihood that MTT could be estimated for a given reach. Reach D1, with plausible C_{gw} and MTT values (1.48 fmol/kg and 27 years, respectively) had the highest likelihood of success, 73.6%. Other likelihoods of success were 15.9% for reach A, 28.3% for reach B, 24.1% for reach C1, 5.5% for reach C2, and 0.2% for reach D2.

Groundwater samples collected with streambed piezometers in three reaches had SF_6 concentrations within or close to the C_{gw} 68% confidence intervals from the Monte Carlo analysis (Fig. 4) in part because the 68% confidence intervals were quite wide. The streambed groundwater samples had SF_6 concentrations from 0.12 to 3.02 fmol/kg, representing piston flow ages between 3 and 46 years (from TracerLPM). It is reasonable to infer these ages based on the piston flow model (PFM) since these samples were collected through short 5-cm screens in the streambed (Kennedy et al. 2009; Gilmore et al. 2016b). Flow-weighting the PFM ages of these samples as explained in the section "Data collection" results in MTT values of 39 years in reach B, 18 years in reach C1, and 41 years in D1.

The groundwater samples from reach C1 are arguably the most comparable to the reach mass-balance C_{gw} because they were collected on the same day as the reach mass-balance data. The SF_6 concentrations of these samples collected on a transect across the channel near the center of reach C1 (the 13,235 m mark) range from 0.28 to 3.02 fmol/kg (Fig. 4); the flow-weighted mean concentration was 1.82 fmol/kg. In comparison the C_{gw} calculated in Eq. (1) was 3.50 fmol/kg.

The groundwater samples from a transect at 20,300 m in reach D1 were collected 7 days before the reach mass-balance work and had a narrower range of SF_6 concentrations (Fig. 4) with a flow-weighted mean of 0.26 fmol/kg. Thus, as in reach C1, the streambed groundwater samples in reach D1 had lower SF_6 and greater age than the C_{gw} and MTT from reach mass-balance (1.48 fmol/kg and 27 years).

Parameter	Reach					
	A	B	C1	C2	D1	D2
n	1%	-1%	1%	0%	-2%	-1%
Q_u	4%	-2%	-1%	-1%	-10%	2%
Q_d	1%	-2%	-3%	-3%	-6%	-3%
Q_{gw}	-3%	-11%	-73%	-19%	-42%	-54%
τ	2%	2%	4%	1%	2%	-1%
D_{SF6}	3%	2%	-3%	-1%	-2%	-1%
D_{C3H8}	3%	0%	-1%	0%	1%	1%
k_{C3H8}	-6%	-2%	0%	-1%	-5%	0%
C_u	1%	-10%	-8%	-13%	-2%	-2%
C_d	-43%	-40%	1%	-29%	-18%	-12%
C_{eq}	-6%	1%	4%	0%	1%	-2%

Fig. 5 Percent change in the width of the 68% confidence interval from the Monte Carlo analysis when the uncertainty in a single input variable was reduced to zero. A negative change (shown in red) means that C_{gw} became more certain

Sensitivity of C_{gw} to each variable in Eq. (1)

To identify the main sources of the uncertainty in the C_{gw} values from reach mass-balance, Monte Carlo analysis of each reach was rerun 11 times, each time with the uncertainty in one input variable set to zero, to show how much the confidence limits on C_{gw} would change. The results show how much the uncertainty in each variable contributes to the uncertainty in C_{gw}. For instance, for a given reach, the uncertainty of Q_u was set to zero in the Monte Carlo analysis, and the percent change in the C_{gw} confidence interval for the reach (compared to the confidence interval in Fig. 4) was determined. This was repeated for each variable in Eqs. (1)–(3), for each reach.

Q_{gw} was the largest single source of uncertainty in C_{gw} in three of the six reaches (C1, D1, and D2) and the second largest source in two reaches (B and C2) (Fig. 5). This is related to Q_{gw} being computed as a relatively small difference between two much larger numbers, Q_d and Q_u. Q_{gw} in reaches B, C1, C2, and D2 was only 7–8% of Q_d.

In reach A, the large influence of uncertainty in C_d (Fig. 5) is due to the unusually large difference between the replicate samples for C_d. Beyond that, the importance of uncertainty in C_d seems surprising: it was the largest contributor to C_{gw} uncertainty in reaches B and C2 as well as the second largest contributor in reaches D1 and D2. Also, uncertainty in C_{gw} was more affected by C_d than by C_u. This seems nonintuitive, given that uncertainties in C_d and C_u were low and identical to each other; SF_6 concentrations in stream water at the upstream and downstream ends of each reach were determined by the exact same sampling and laboratory processes and have the same percentage uncertainty.

Table 3 Uncertainty in key variables in recent applications of reach mass-balance SF_6 groundwater age-dating. The only major difference between the present study and Solomon et al. (2015) is the larger uncertainty in Q_{gw} in the Sand Hills study

Variable	Average uncertainty (present study)	Average uncertainty (Solomon et al. 2015)
Tracer used	SF_6	SF_6
C: tracer concentrations	5%	5%
C_{eq}: equilibrium concentration	2.4%	3%
k: gas exchange rate constant	21.8%	15%
Q: stream discharge	2.7%	3%
Q_{gw}: groundwater discharge	31.8%	3%

The nonintuitive outcome arises because contributions to the uncertainty in C_{gw} depend on the mathematical sensitivities of C_{gw} to each input variable (i.e., the value of the partial derivative of C_{gw} with respect to each variable), not just the uncertainty in each variable (e.g., Taylor 1982; Kline 1985; Genereux 1998). From Eq. (1), $\partial C_{gw}/\partial C_d$ has a larger absolute value than $\partial C_{gw}/\partial C_u$ (see ESM), making C_{gw} more sensitive to uncertainty in C_d than to uncertainty in C_u.

Uncertainty in Q_{gw}, and planning reach mass-balance work

Overview

The reach mass-balance results and uncertainty analyses (Tables 1–2; Figs. 4–5), along with a comparison to the only other reach mass-balance age-dating study using SF_6 (Table 3), suggest that reach mass-balance estimation of groundwater age and MTT is a useful technique if Q_{gw} into the stream reach is not too small relative to the stream discharge rates at the ends of the reach (Q_d and Q_u). When Q_{gw} is small relative to Q_d and Q_u, the uncertainty in the Q_{gw} value is large owing to the standard calculation of uncertainty in Q_{gw} as the square root of the sum of the squared uncertainties in Q_d and Q_u (see ESM). The ratio of Q_{gw} to Q_d ranged from 0.07 to 0.57 (mean of 0.17) in the six Sand Hills reaches, resulting in a mean uncertainty of about 32% in the Q_{gw} values for these reaches; this proved problematic for constraining MTT. Meanwhile Solomon et al. (2015) reported an uncertainty of 3% in Q_{gw} which was associated with $Q_{gw}/Q_d = 0.5$ and a well constrained value of MTT (Table 3).

Those observations, along with a standard calculation of uncertainty in Q_{gw}, suggest that Q_{gw}/Q_d greater than or equal to about 0.4 would generally give reasonable uncertainty

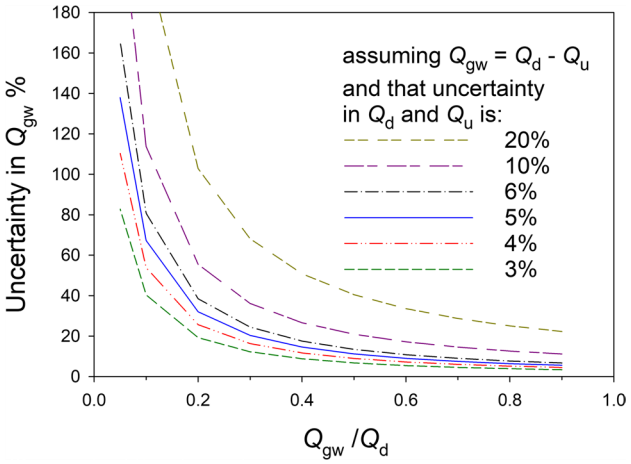

Fig. 6 Uncertainty in Q_{gw} as a function of Q_{gw}/Q_d, for assumed uncertainties of 3–20% in Q_d and Q_u

in Q_{gw} (Fig. 6). For example, $Q_{gw}/Q_d = 0.4$ results in <15% uncertainty in Q_{gw} if uncertainty in stream discharge is 5% or less. While the significance of and tolerance for uncertainty in Q_{gw} may vary among studies, a Q_{gw}/Q_d value of about 0.4 can be considered a reasonable design target in planning reach mass-balance studies to determine C_{gw} for age-dating tracers or other solutes. Uncertainty obviously exists in other variables and may be significant. But uncertainty in Q_{gw} stands out as potentially critical: it can be either harmfully large or acceptably small (Fig. 6), depending on the choice of field site and measurement locations.

Q_{gw}/Q_d above a certain threshold is not the only consideration for good outcomes from reach mass-balance. For example, the suggestion that the α parameter be small, preferably <<10 (Solomon et al. 2015), is also relevant. Cook (2013) makes a number of suggestions related to successful determination of Q_{gw} from C_{gw} (the "inverse" of the problem discussed here), some of which are also relevant to determination of C_{gw} from Q_{gw} and other data.

With regard to the recommendation here concerning selection of reaches with large groundwater input, two new planning concepts and tools are proposed:

1. The ratio of groundwater recharge rate (R) to stream density (D) may be a general screening metric for assessing the relative likelihood of successful reach mass-balance work: successful outcomes are more likely in watersheds with higher R/D.
2. Within a particular watershed, analysis of watershed areas and/or stream lengths using a digital elevation model (DEM) in a geographic information system (GIS) can be used to estimate whether long-term average Q_{gw}/Q_d in a specific reach is likely to exceed a desired minimum (e.g., the value of 0.4 mentioned previously or another chosen threshold).

Table 4 Recharge rate (R) and stream density (D) estimates for the SBML River watershed (Nebraska Sand Hills) and the West Bear Creek watershed (North Carolina coastal plain)

Parameter	SBML River, NE (present study)	West Bear Creek, NC (Solomon et al. 2015)
Recharge rate, R (mm/year)	50	210
Total perennial stream length (km)	105.5	55.5
Watershed area (km^2)	225	61
Stream density, D (km/km^2)	0.47	0.91
R/D (m^2/year)	107	231

The following sections describe the application of these concepts at the Sand Hills study site and outline their use as planning tools in selecting locations and measurement sites for field application of reach mass-balance.

R/D for screening among potential study watersheds

The greater the recharge per unit area of watershed (R), and the smaller the cumulative length of stream channels available for groundwater discharge per unit area of watershed (D), the greater Q_{gw} is likely to be for a given stream reach, and the less likely it would be that large uncertainty in Q_{gw} interferes with estimation of C_{gw} and groundwater MTT.

To test the viability of R/D as a screening metric, R and D from the Sand Hills study site were compared with the only previously published reach mass-balance study using SF_6 for groundwater age-dating (Solomon et al. 2015). Recharge estimates came from available literature (Szilagyi et al. 2011; Szilagyi and Jozsa 2012; Solomon et al. 2015; Gilmore et al. 2018). The National Hydrography Dataset (NHD) (US Geological Survey 2021) and 10 m DEMs were used to calculate D. Details are in section S5 of the ESM.

R/D at West Bear Creek was estimated to be 231 m^2/year (Table 4) and is associated with a well-constrained application of reach mass-balance; Solomon et al. (2015) reported a C_{gw} of 0.55 fmol/kg and 68% confidence limits of about ± 0.21 fmol/kg, with corresponding uncertainty of about ±4 years in MTT. R/D in the Nebraska Sand Hills was 107 m^2/year and was associated with a much less well-constrained application of reach mass-balance (this paper). For even the best case among the six Sand Hills reaches (reach D1, Fig. 4), the 68% confidence limits were much broader (about ± 1.1 fmol/kg); in some cases, confidence limits for the other Sand Hills reaches reached implausible values, e.g., above C_{eq} (implying future age) or below zero (not possible). These results are consistent with the suggestion that successful reach mass-balance outcomes are more likely in watersheds of greater R/D.

In the preceding comparison, only perennial streams were considered in the calculation of R/D. The exclusion of ephemeral streams seemed reasonable given that the West Bear Creek data (Solomon et al. 2015) were collected during low summer baseflow in July when ephemeral streams were not active and Q_{gw}/Q_d was 0.5 (there are no ephemeral streams identified by the NHD in the SBML River watershed in the Nebraska Sand Hills, so R/D there was necessarily based only on perennial streams).

An attempt to determine groundwater MTT at West Bear Creek under early spring baseflow in March (stream flow 10 × higher than in July) was not successful as C_{gw} was too similar to C_{eq} (Solomon et al. 2015). Intermittent and ephemeral streams were likely active at West Bear Creek in March; Q_{gw}/Q_d in March was only 0.17 and R/D considering both perennial and ephemeral streams was 104 m^2/year, much lower than the July value (231) and close to the value for the Sand Hills (107) (Table 4). This again emphasizes the potential significance and usefulness of R/D, because the decrease in R/D at West Bear Creek, from July to March, was associated with a change from favorable to unfavorable conditions and outcomes for reach mass-balance. In some places the relevant value of D may be higher under wet conditions if ephemeral channels become active. The West Bear Creek results suggest that reach mass-balance in a watershed with a significant length of ephemeral channels may be more successful under low baseflow conditions.

There is significant variation in R/D across the 48 contiguous US states (Fig. 7). Locations of higher R/D include the northeastern US, the Laurentian Great Lakes region, northern Florida, the Pacific Northwest and some areas in the Mountain West. Given the R/D values above for the Sand Hills and West Bear Creek (107 and 231 m^2/year, respectively), the yellow zone in Fig. 7 (R/D of 111–230 m^2/year) may represent the transition between areas of high and low potential for successful reach mass-balance. When only perennial streams are considered (Fig. 7b), parts of the American West have high R/D values because of the relatively low stream density for perennial channels. However, high R/D in these desert-like areas due to very low D (e.g., in west Texas, and central New Mexico) is unlikely to be associated with successful reach mass-balance. Given that the Nebraska Sand Hills have higher recharge than much of the American West, and MTT was not well constrained by reach mass-balance in the Sand Hills, reach mass-balance may not be feasible in most places in the US west.

While the NHD is a useful dataset, variations in the input data affect the calculation of stream density. For example, in the NHD, 1:24,000 topographic maps were generally used to identify streams. However, if even a single 1:24,000-scale map was not available in a 1:100,000-scale quadrangle, then

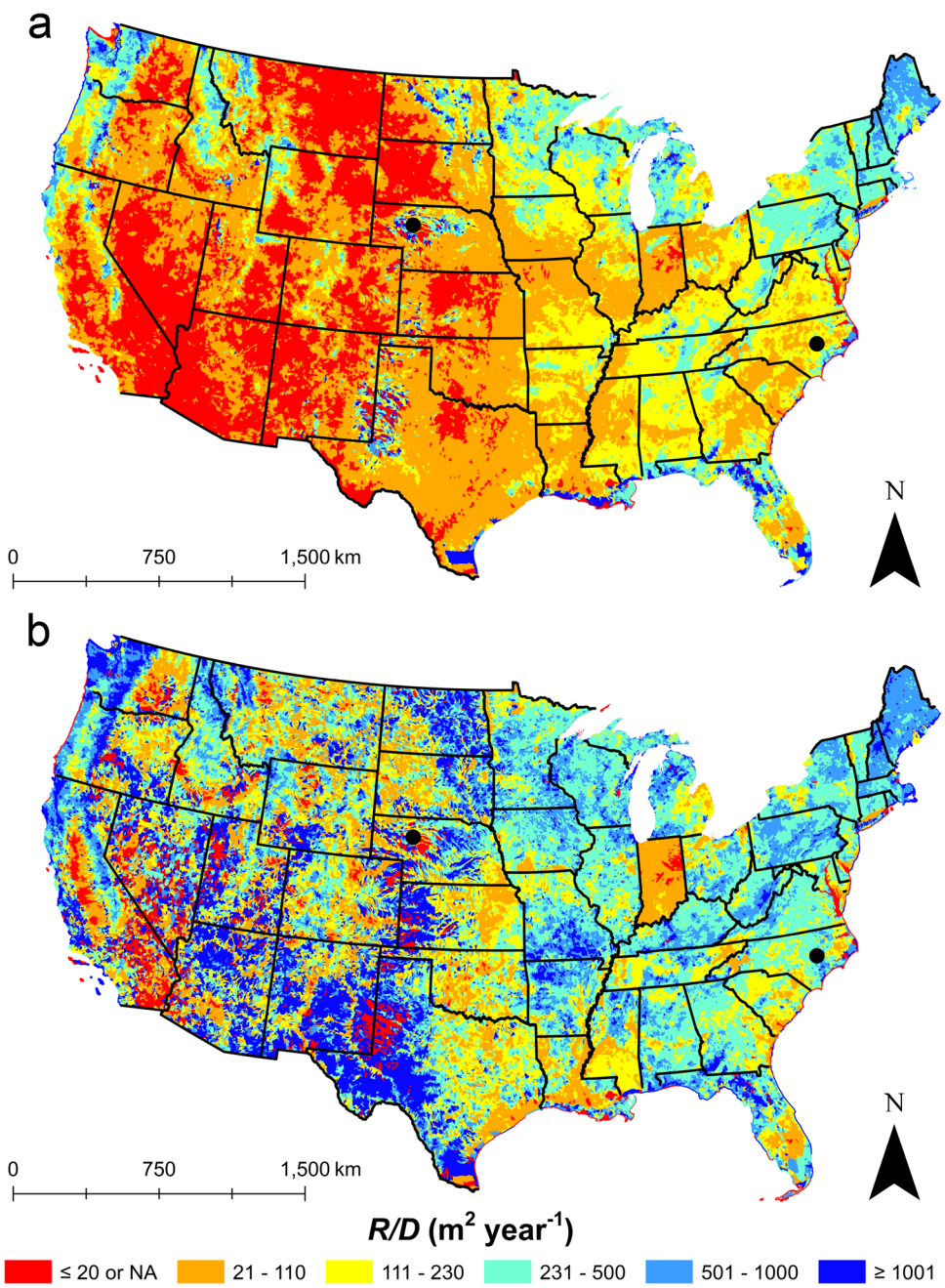

Fig. 7 *R*/*D* in the 48 contiguous US states in m^2/year. **a** All streams identified by the National Hydrography Dataset (NHD) were included in the analysis; **b** Only perennial streams were included in the analysis (intermittent and ephemeral streams were removed). A raster of recharge in the 48 contiguous US states (Reitz et al. 2017) was averaged for each 12-digit hydrologic unit (HUC12) in the NHD. The stream density for each HUC12 unit was calculated using the NHD and the methods of US Environmental Protection Agency (2016). For each HUC12 unit, the average recharge was divided by the stream density. The black dots represent the field sites in the Nebraska Sand Hills (this paper) and in North Carolina at West Bear Creek (Solomon et al. 2015)

less detailed 1:62,500-scale topographic maps were used instead for the entire 1:100,000-scale quadrangle (Anne Neale at the US Environmental Protection Agency, personal communication, 19 August 2020). Also, some state-to-state variation in the definition of stream channels and thus *R*/*D* is apparent (e.g., the low *R*/*D* in Indiana).

Figure 7 provides an example showing that the data required for use of the *R*/*D* metric are broadly available, that *R*/*D* can be computed and mapped using standard GIS files and methods, and that there is a significant range of *R*/*D* at the continental scale likely controlled by climate and other factors; it also gives a broad indication of areas where *R*/*D* is higher in the US and therefore reach mass-balance is more likely to be successful. However, carefully determined values at a smaller scale (e.g., Table 4) may be more useful for planning reach mass-balance work in a specific area.

Reach-specific screening based on watershed areas or stream lengths

In a watershed with acceptably high *R*/*D*, a simple geospatial analysis may be useful in selection of a stream reach for reach mass-balance work. The analysis proposed is based

on watershed areas or stream lengths that can generally be estimated via GIS.

If a watershed experiences neither net gain nor loss of water by groundwater flow beneath topographic divides, the long-term average rate of volumetric groundwater discharge to a stream will be RA, the area of its watershed (A) times the average groundwater recharge rate (R) in the watershed. Assuming uniform recharge within the watershed, the long-term average rate of groundwater discharge to a stream reach, Q_{gw}, is:

$$Q_{gw} = Q_d - Q_u = R\left(A_d - A_u\right) \tag{6}$$

where "d" and "u" refer to the downstream and upstream ends of the reach, respectively, Q is the long-term average rate of baseflow (the stream flow due to groundwater discharge), and A is the watershed area upstream of "d" or "u". $A_d - A_u$ can be considered the contributing area of the reach. Assuming that the recharge in the contributing area is eventually discharged to the reach, the long-term average Q_{gw}/Q_d for the reach is:

$$\frac{Q_{gw}}{Q_d} = \left(1 - \frac{A_u}{A_d}\right) \tag{7}$$

Given that A_u and A_d are defined by the locations of the upstream and downstream ends of the stream reach, Eq. (7) implicitly gives the value of Q_{gw}/Q_d for a reach of any chosen length at any place in the watershed. Equation (7) suggests a GIS-based approach for defining a stream reach for reach mass-balance work:

1. Choose the desired minimum "threshold" value of Q_{gw}/Q_d for the study reach (see the "Overview" section).
2. Pick two points on the stream to serve as the upstream and downstream ends of a reach.
3. Delineate the watersheds for both points using a DEM in GIS.
4. Calculate A_u/A_d from the delineated watersheds. Use Eq. (7) to predict Q_{gw}/Q_d.
5. Compare the predicted Q_{gw}/Q_d from Eq. (7) with the threshold value. If the predicted value is lower than the threshold value, lengthen the reach by moving at least one end of it, and/or move the reach farther upstream.
6. Repeat steps 3–5 until the predicted value of Q_{gw}/Q_d from Eq. (7) meets or exceeds the desired minimum value.

Groundwater recharge rate does not appear in Eq. (7) because the equation assumes uniform recharge. If there is a difference in groundwater recharge rate between the watershed areas A_u and A_d (e.g., due to an elevational gradient in R), the different rates R_u and R_d can be taken into account if the ratio R_u/R_d can be estimated from hydrologic information (Eq. 8). If the stream reach of interest includes a confluence with a tributary, the influence of the tributary can be subtracted to keep the focus of the analysis specifically on the groundwater discharge to the channel of the main stream (Eq. 9).

A similar analysis can be carried out using stream length rather than watershed area. R/D can be considered the long-term average volume of groundwater per stream length available for discharge to the stream, per unit time. Thus, long-term average baseflow is the product of R/D and L, where L is the total stream length (including tributaries) upstream of a given point in the stream channel. If R/D is uniform within the watershed, then Q_{gw}/Q_d can be estimated using Eq. (10), which is analogous to Eq. (7). Versions of Eq. (10) appropriate for $R_u \neq R_d$ and reaches with tributary confluences can be written (Eqs. 11 and 12).

L_u and L_d can be calculated within a GIS program by summing the lengths of all streams and tributaries upstream of the two ends of the stream reach. To do this, the local stream network can be obtained from a GIS shapefile, aerial imagery, or a watershed delineation. L_u/L_d may be useful in situations where a high-resolution DEM is not available for watershed delineation or where groundwater flow is not closely related to topographic boundaries.

For the six Sand Hills study reaches, the predicted Q_{gw}/Q_d values based on Eqs. (7) and (10) are generally lower than the field measured Q_{gw}/Q_d values based on stream discharge measurements (Fig. 8), but the relative variation among the reaches is similar (i.e., values are relatively higher for reaches A and D1 compared to the other four reaches). Based on this, reaches A and D1 seem most suitable for reach mass-balance. This held true for D1, the only reach where a realistic C_{gw} value was calculated; a realistic C_{gw}

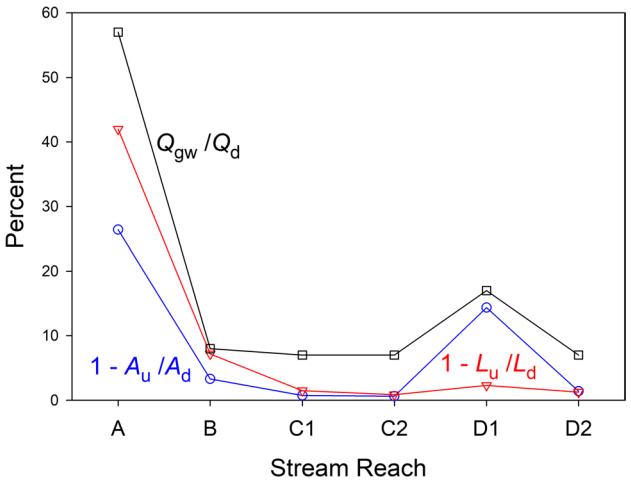

Fig. 8 Field measured Q_{gw}/Q_d compared to the Q_{gw}/Q_d values predicted by $1 - A_u/A_d$ (Eq. 7) and $1 - L_u/L_d$ (Eq. 10) at the Nebraska Sand Hills study site

Table 5 Models for prediction of Q_{gw}/Q_d from watershed areas and stream lengths (see text). The subscript "t" refers to the watershed of a tributary

Equation number	Equation for predicting Q_{gw}/Q_d	Conditions
Eq. (7)	$\frac{Q_{gw}}{Q_d} = 1 - \frac{A_u}{A_d}$	Uniform recharge
Eq. (8)	$\frac{Q_{gw}}{Q_d} = 1 - \frac{R_u A_u}{R_d A_d}$	$R_u \neq R_d$
Eq. (9)	$\frac{Q_{gw}}{Q_d} = 1 - \frac{R_u A_u}{R_d A_d} - \frac{R_t A_t}{R_d A_d}$	$R_u \neq R_d$ and tributary input to study reach
Eq. (10)	$\frac{Q_{gw}}{Q_d} = 1 - \frac{L_u}{L_d}$	Uniform recharge
Eq. (11)	$\frac{Q_{gw}}{Q_d} = 1 - \frac{R_u L_u}{R_d L_d}$	$R_u \neq R_d$
Eq. (12)	$\frac{Q_{gw}}{Q_d} = 1 - \frac{R_u L_u}{R_d L_d} - \frac{R_t L_t}{R_d L_d}$	$R_u \neq R_d$ and tributary input to study reach

was not calculated in reach A, mainly due to unusual uncertainty in C_d in reach A, not uncertainty in Q_{gw} (Fig. 5). Consistent with Fig. 8, reach A was the only reach in which Q_{gw} did not contribute much toward the total uncertainty in C_{gw}.

The simple analysis in Table 5 is based on long-term averages; it is focused on baseflow, not temporal variation associated with changing hydrologic conditions. However, the results (Fig. 8) suggest Eqs. (7–12) may provide at least a relative index that could be useful in choosing the location and length of a stream reach for successful reach mass-balance groundwater age dating or reach mass-balance work for a solute other than an age-dating tracer.

Summary and conclusions

Atmospherically derived SF_6 was used as an age-dating tracer in the reach mass-balance analysis of six stream reaches in the Nebraska Sand Hills. The goal was to estimate C_{gw}, the flow-weighted mean SF_6 concentration in the groundwater discharge to each reach, and to estimate groundwater MTT from C_{gw}. The calculated C_{gw} in reach D1 (1.48 fmol/kg) was consistent with a groundwater MTT of 27 years, roughly triple the minimum groundwater age in this reach based on stream water SF_6 concentrations uncorrected for gas exchange (Fig. 3). Estimated C_{gw} values from reach mass-balance in the other reaches were not plausible. From the Monte Carlo analyses, it is apparent that the uncertainty in computed C_{gw} values was large. Low rates of groundwater inflow to the stream reaches (Q_{gw}) led to large uncertainty in Q_{gw}, which in turn was a major contributor to uncertainty in C_{gw}. Reaches where uncertainty in Q_{gw} is high are, in general, not good candidates for reach mass-balance work.

In reach mass-balance, Q_{gw} is calculated as the difference between volumetric stream flow at the upstream and downstream ends of a stream reach (Q_d–Q_u). When Q_{gw} is small compared to stream flow, Q_{gw} will be highly uncertain as a small difference between two much larger numbers. When planning reach mass-balance, emphasis should be placed on maximizing Q_{gw}/Q_d in order to better constrain C_{gw} and MTT. Field measurements of stream flow and/or groundwater discharge are the best way to determine Q_{gw}/Q_d and the suitableness of a stream reach for reach mass-balance. In the absence of such data, two screening metrics may help identify suitable reaches prior to the start of field work and increase the likelihood of successful outcomes in reach mass-balance. The first metric is for broad screening related to climate (recharge rate) and stream networks (the "density" of drainage, cumulative channel length per watershed area). The second metric is designed to be applied to a specific stream reach, preferably but not necessarily in a watershed or region that has passed the first metric. Thus, the two metrics are useful in different ways, and offer researchers interested in reach mass-balance two options for screening potential field sites.

Successful outcomes in reach mass-balance may be more likely in watersheds with a higher ratio of groundwater recharge rate (R) to stream density (D). R/D values for the Nebraska watershed studied here and a North Carolina watershed previously studied with the same approach (Solomon et al. 2015) were consistent with this, when only perennial streams were considered: a better outcome was achieved in the North Carolina watershed, which had higher R/D than the Nebraska watershed (231 vs. 107 m^2/year). Estimates of recharge can often be obtained from a variety of methods and existing literature, while stream density can be calculated in a GIS using publicly available data.

Going further to focus on reaches of specific length in specific locations in a potential study watershed, a GIS based analysis of watershed areas at the upstream and downstream ends of a reach (A_u and A_d respectively) can be used in the planning stages of reach mass-balance work to assess whether Q_{gw}/Q_d for the reach may reach a desired minimum value. A similar analysis could be based on cumulative stream lengths upstream of the two ends of the reach (L_u and L_d). Many proposed reaches could be screened for suitable Q_{gw}/Q_d prior to field measurements, considering the broad availability of DEMs and stream network data.

Thus, selecting watersheds with high R/D and, within those watersheds, reaches with a large predicted Q_{gw}/Q_d (Eqs. 7–12) is recommended in order minimize the

uncertainty in Q_{gw} and increase the probability of a successful application of reach mass-balance. Further development and application of these screening metrics may increase confidence in reach selection in future studies, allowing reach mass-balance to become a more versatile tool in hydrogeology.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10040-021-02435-8.

Funding The authors gratefully acknowledge financial support of this work from the U.S. National Science Foundation through awards 1744714 (NC State University), 1744719 (University of Nebraska), and 1744721 (University of Utah).

Declarations

Conflicts of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Alvarado JAC, Purtschert R, Hinsby K, Troldborg L, Hofer M, Kipfer R, Aeschbach-Hertig W, Arno-Synal H (2005) ^{36}Cl in modern groundwater dated by a multi-tracer approach (^{3}H/^{3}He, SF_6, CFC-12 and ^{85}Kr): a case study in Quaternary sand aquifers in the Odense Pilot River Basin, Denmark. Appl Geochem 20(3):599–609. https://doi.org/10.1016/j.apgeochem.2004.09.018

Banks VJ, Palumbo-Roe B (2010) Synoptic monitoring as an approach to discriminating between point and diffuse source contributions to zinc loads in mining impacted catchments. J Environ Monit 12:1684–1698. https://doi.org/10.1039/c0em00045k

Billesbach DP, Arkebauer TJ (2012) First long-term, direct measurements of evapotranspiration and surface water balance in the Nebraska SandHills. Agric For Meteorol 156:104–110. https://doi.org/10.1016/j.agrformet.2012.01.001

Böhlke JK, Denver JM (1995) Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic coastal plain, Maryland. Water Resour Res 31(9):2319–2339. https://doi.org/10.1029/95WR01584

Bullister JL, Wisegarver DP, Menzia FA (2002) The solubility of sulfur hexafluoride in water and seawater. Deep Sea Res Part I 49:175–187. https://doi.org/10.1016/S0967-0637(01)00051-6

Busenberg E, Plummer LN (1992) Use of chlorofluorocarbons (CCl_3F and CCl_2F_2) as hydrologic tracers and age-dating tools: the alluvium and terrace system of central Oklahoma. Water Resour Res 28(9):2257–2283. https://doi.org/10.1029/92WR01263

Busenberg E, Plummer LN (2000) Dating young groundwater with sulfur hexafluoride: natural and anthropogenic sources of sulfur hexafluoride. Water Resour Res 36(10):3011–3030. https://doi.org/10.1029/2000WR900151

Chacha N, Njau KN, Lugomela GV, Muzuka ANN (2018) Groundwater age dating and recharge mechanism of Arusha aquifer, northern Tanzania: application of radioisotope and stable isotope techniques. Hydrogeol J 26:2693–2706. https://doi.org/10.1007/s10040-018-1832-0

Chen M, Parton WJ, Hartman MD, Del Grosso SJD, Smith WK, Knapp AK, Lutz S, Derner JD, Tucker CJ, Ojima DS, Volesky JD, Stephenson MB, Schacht WH, Gao W (2019) Assessing precipitation, evapotranspiration, and NDVI as controls of U.S. Great Plains plant production. Ecosphere 10(10):1–17. https://doi.org/10.1002/ecs2.2889

Choi J, Hulseapple SM, Conklin MH, Harvey JW (1998) Modeling CO_2 degassing and pH in a stream–aquifer system. J Hydrol 209(1):297–310. https://doi-org.prox.lib.ncsu.edu/10.1016/S0022-1694(98)00093-6

Cook PG (2013) Estimating groundwater discharge to rivers from river chemistry surveys. Hydrol Process 27:3694–3707. https://doi.org/10.1002/hyp.9493

Cook PG, Solomon DK (1997) Recent advances in dating young groundwater: chlorofluorocarbons, ^{3}H/^{3}He and ^{85}Kr. J Hydrol 191(1):245–265. https://doi.org/10.1016/S0022-1694(96)03051-X

Cook PG, Favreau G, Dighton JC, Tickell S (2003) Determining natural groundwater influx to a tropical river using radon, chlorofluorocarbons and ionic environmental tracers. J Hydrol 277:74–88. https://doi.org/10.1016/S0022-1694(03)00087-8

Delin GN, Healy RW, Lorenz DL, Nimmo JR (2007) Comparison of local- to regional-scale estimates of ground-water recharge in Minnesota, USA. J Hydrol 334(1):231–249. https://doi-org.prox.lib.ncsu.edu/10.1016/j.jhydrol.2006.10.010

Ekwurzel B, Schlosser P, Smethie WM Jr, Plummer LN, Busenberg E, Michel RL (1994) Dating of shallow groundwater: comparison of the transient tracers ^{3}H/^{3}He, chlorofluorocarbons, and ^{85}Kr. Water Resour Res 30(6):1693–1708. https://doi.org/10.1029/94WR00156

Genereux D (1998) Quantifying uncertainty in tracer-based hydrograph separations. Water Resour Res 34(4):915–919. https://doi.org/10.1029/98WR00010

Genereux DP, Hemond HF (1990) Naturally occurring radon 222 as a tracer for streamflow generation: steady state methodology and field example. Water Resour Res 26(12):3065–3075. https://doi.org/10.1029/WR026i012p03065

Genereux DP, Hemond HF (1992) Determination of gas exchange rate constants for a small stream on Walker Branch Watershed, Tennessee. Water Resour Res 28(9):2365–2374. https://doi.org/10.1029/92WR01083

Genereux DP, Hemond HF, Mulholland PJ (1993) Use of radon-222 and calcium as tracers in a three-end-member mixing model for streamflow generation on the West Fork of Walker Branch Watershed. J Hydrol 142:167–211. https://doi.org/10.1016/0022-1694(93)90010-7

Gilmore TE, Genereux DP, Solomon DK, Solder JE, Kimball BA, Mitasova H, Birgand F (2016a) Quantifying the fate of agricultural nitrogen in an unconfined aquifer: stream-based observations at three measurement scales. Water Resour Res 52(3):1961–1983. https://doi.org/10.1002/2015WR017599

Gilmore TE, Genereux DP, Solomon DK, Solder JE (2016b) Groundwater transit time distribution and mean from streambed sampling in an agricultural coastal plain watershed, North Carolina, USA. Water Resour Res 52(3):2025–2044. https://doi.org/10.1002/2015WR017600

Gilmore TE, Zlotnik V, Johnson M (2018) Recognition of regional water table patterns for estimating recharge rates in shallow aquifers. Groundwater 57(3):443–454. https://doi.org/10.1111/gwat.12808

Gosselin DC, Drda S, Harvey FE, Goeke J (1999) Hydrologic setting of two interdunal valleys in the central Sand Hills of Nebraska. Ground Water 37(6):924–933. https://doi.org/10.1111/j.1745-6584.1999.tb01192.x

Gosselin DC, Sridhar V, Harvey FE, Goeke JW (2006) Hydrological effects and groundwater fluctuations in interdunal environments in the Nebraska Sandhills. Great Plains Res 16(1):17–28

Holtzman R, Shavit U, Segal-Rozenhaimer S, Gavrieli I, Marei A, Farber E, Vengosh A (2005) Quantifying ground water inputs along the Lower Jordan River. J Environ Qual 34:897–906. https://doi.org/10.2134/jeq2004.0244

Hope D, Palmer SM, Billett MF, Dawson JJC (2001) Carbon dioxide and methane evasion from a temperate peatland stream. Limnol Oceanogr 46(4):847–857. https://doi.org/10.4319/lo.2001.46.4.0847

Jähne B, Heinz G, Dietrich W (1987) Measurement of the diffusion coefficients of sparingly soluble gases in water. J Geophys Res Oceans 92(C10):10767–10776. https://doi.org/10.1029/JC092iC10p10767

Jones JB, Mulholland PJ (1998) Carbon dioxide variation in a hardwood forest stream: an integrative measure of whole catchment soil respiration. Ecosystems 1(2):183–196. https://doi.org/10.1007/s100219900014

Jurgens BC, Böhlke JK, Eberts SM (2012) TracerLPM (version 1): an Excel® workbook for interpreting groundwater age distributions from environmental tracer data. US Geological Survey Techniques and Methods Report 4-F2, USGS, Reston, VA, p 60

Katz BG (2004) Sources of nitrate contamination and age of water in large karstic springs of Florida. Environ Geol 46(6):689–706. https://doi.org/10.1007/s00254-004-1061-9

Kennedy CD, Genereux DP, Corbett DR, Mitasova H (2009) Relationships among groundwater age, denitrification, and the coupled groundwater and nitrogen fluxes through a streambed. Water Resour Res 45:W09402. https://doi.org/10.1029/2008WR007400

Kilpatrick FA, Rathbun RE, Yotsukura N, Parker GW, DeLong LL (1989) Determination of stream reaeration coefficients by use of tracers. Techniques of Water-Resources Investigations of the United States Geological Survey, USGS, Denver, CO, pp 1–52

Kim H, Hemond HF, Krumholz LR, Cohen BA (1995) In-situ biodegradation of toluene in a contaminated stream, part 1: field studies. Environ Sci Technol 29(1):108–116. https://doi.org/10.1021/es00001a014

Kim H, Hemond HF (1998) Natural discharge of volatile organic compounds from contaminated aquifer to surface waters. J Environ Eng 124(8):744–751. https://doi.org/10.1061/(ASCE)0733-9372(1998)124:8(744)

Kimball BA, Runkel RL, Gemer LJ (2001) Quantification of mine-drainage inflows to Little Cottonwood Creek, Utah, using a tracer-injection and synoptic-sampling study. Environ Geol 40:1390–1404. https://doi.org/10.1007/s002540100320

King DB, Saltzman ES (1995) Measurement of the diffusion coefficient of sulfur hexafluoride in water. J Geophys Res Oceans 100(C4):7083–7088. https://doi.org/10.1029/94JC03313

Kline SJ (1985) The purposes of uncertainty analysis. J Fluids Eng 107(2):153–160. https://doi.org/10.1115/1.3242449

Lauffenburger ZH, Gurdak JJ, Hobza C, Woodward D, Wolf C (2018) Irrigated agriculture and future climate change effects on groundwater recharge, northern High Plains aquifer, USA. Agric Water Manag 204:69–80. https://doi.org/10.1016/j.agwat.2018.03.022

List RJ (1949) Smithsonian meteorological tables, 6th edn. Smithsonian Institution Press, Washington, DC, 527 pp

Loope DB, Swinehart JB (2000) Thinking like a dune field: geologic history in the Nebraska Sand Hills. Great Plains Res 10(1):5–35

Maiss M, Brenninkmeijer CAM (1998) Atmospheric SF_6: trends, sources, and prospects. Environ Sci Technol 32(20):3077–3086

Miller JA, Appel CL (1997) Segment 3: Kansas, Missouri, and Nebraska. In: Ground water atlas of the United States. US Geological Survey, Reston, VA, pp D1–D24

Natchimuthu S, Wallin MB, Klemedtsson L, Bastviken D (2017) Spatio-temporal patterns of stream methane and carbon dioxide emissions in a hemiboreal catchment in Southwest Sweden. Sci Rep 7:1–12. https://doi.org/10.1038/srep39729

NOAA (2020) Sulfur hexafluoride (SF_6): combined data set. Earth System Research Laboratory. https://www.esrl.noaa.gov/gmd/hats/combined/SF6.html. Accessed 7 December 2020

Oviedo-Vargas D, Genereux DP, Dierick D, Oberbauer SF (2015) The effect of regional groundwater on carbon dioxide and methane emissions from a lowland rainforest stream in Costa Rica. J Geophys Res Biogeosci 120(12):2579–2595. https://doi.org/10.1002/2015JG003009

Parker GW, DeSimone LA (1992) Estimating reaeration coefficients for low-slope streams in Massachusetts and New York 1985–88. US Geological Survey, Denver, CO

Poreda RJ, Cerling TE, Salomon DK (1988) Tritium and helium isotopes as hydrologic tracers in a shallow unconfined aquifer. J Hydrol 103(1):1–9. https://doi.org/10.1016/0022-1694(88)90002-9

Qi SL (2009) Digital map of aquifer boundary for the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming. US Geological Survey. https://water.usgs.gov/GIS/metadata/usgswrd/XML/ds543.xml#stdorder. Accessed 7 December 2020

Reitz M, Sanford WE, Senay GB, Cazenas J (2017) Annual estimates of recharge, quick-flow runoff, and evapotranspiration for the contiguous U.S. using empirical regression equations. J Am Water Resour Assoc 53(4):961–983. http://dx.doi.org.prox.lib.ncsu.edu/10.1111/1752-1688.12546

Sanford WE, Casile G, Haase KB (2015) Dating base flow in streams using dissolved gases and diurnal temperature changes. Water Resour Res 51(12):9790–9803. https://doi.org/10.1002/2014WR016796

Schlosser P, Stute M, Dörr H, Sonntag C, Münnich KO (1988) Tritium/^{3}He dating of shallow groundwater. Earth Planet Sci Lett 89(3):353–362. https://doi.org/10.1016/0012-821X(88)90122-7

Solomon DK, Gilmore TE, Solder JE, Kimball B, Genereux DP (2015) Evaluating an unconfined aquifer by analysis of age-dating tracers in stream water. Water Resour Res 51(11):8883–8899. https://doi.org/10.1002/2015WR017602

Solomon DK, Humphrey E, Gilmore TE, Genereux DP, Zlotnik V (2020) An automated seepage meter for streams and lakes. Water Resour Res 56:1–11. https://doi-org.prox.lib.ncsu.edu/10.1029/2019WR026983

Stolp BJ, Solomon DK, Suckow A, Vitvar T, Rank D, Aggarwal PK, Han LF (2010) Age dating base flow at springs and gaining streams using helium-3 and tritium: Fischa-Dagnitz system, southern Vienna Basin, Austria. Water Resour Res 46(7):1–13. https://doi.org/10.1029/2009WR008006

Szilagyi J, Zlotnik VA, Gates JB, Jozsa J (2011) Mapping mean annual groundwater recharge in the Nebraska Sand Hills, USA. Hydrogeol J 19(8):1503–1513. https://doi.org/10.1007/s10040-011-0769-3

Szilagyi J, Jozsa J (2012) MODIS-aided statewide net groundwater-recharge estimation in Nebraska. Groundwater 51(5):735–744. https://doi.org/10.1111/j.1745-6584.2012.01019.x

Taylor JR (1982) An introduction to error analysis: the study of uncertainties in physical measurements. University Science Books, Mill Valley, CA

US Environmental Protection Agency (2016) EnviroAtlas: stream density. https://epa.gov/enviroatlas. Accessed 5 July 2020

US Geological Survey (2018) SF_6 sampling. The Reston Groundwater Dating Laboratory. https://water.usgs.gov/lab/sf6/sampling/. Accessed 9 November 2021

US Geological Survey (2019) Specific conductance. In: Survey Techniques and Methods, book 9, chap. A6.3. USGS, Reston, VA, pp 1–15

US Geological Survey (2021) National hydrography dataset. https://www.usgs.gov/core-science-systems/ngp/national-hydrography/national-hydrography-dataset?qt-science_support_page_related_con=0#qt-science_support_page_related_con. Accessed 9 November 2021

Wallin MB, Öquist MG, Buffam I, Billett MF, Nisell J, Bishop KH (2011) Spatiotemporal variability of the gas transfer coefficient (K_{CO2}) in boreal streams: implications for large scale estimates of CO_2 evasion. Glob Biogeochem Cycles 25(3):1–14. https://doi.org/10.1029/2010GB003975

Weiss RF, Price BA (1980) Nitrous oxide solubility in water and seawater. Mar Chem 8(4):347–359. https://doi.org/10.1016/0304-4203(80)90024-9

Wise DL, Houghton G (1966) The diffusion coefficients of ten slightly soluble gases in water at 10–60°C. Chem Eng Sci 21(11):999–1010. https://doi.org/10.1016/0009-2509(66)85096-0

Zoellmann K, Kinzelbach W, Fulda C (2001) Environmental tracer transport (^3H and SF_6) in the saturated and unsaturated zones and its use in nitrate pollution management. J Hydrol 240(3):187–205. https://doi.org/10.1016/S0022-1694(00)00326-7