High resolution identification and quantification of diffuse deep groundwater discharge in mountain rivers using continuous boat-mounted helium measurements

Connor P. Newman a,*,1, C. Eric Humphrey b, Matthias S. Brennwald c, W. Payton Gardner d,2, Kelli M. Palko a, Michael Gooseff e, D. Kip Solomon f,3

a U.S. Geological Survey, Colorado Water Science Center, USA
b U.S. Geological Survey, Utah Water Science Center, USA
c Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
d University of Montana, Department of Geosciences, USA
e University of Colorado at Boulder, Department of Civil, Environmental and Architectural Engineering, USA
f University of Utah, Department of Geology and Geophysics, USA

ARTICLE INFO

Keywords:
- Noble gases
- Deep groundwater
- Portable mass spectrometer
- Diffuse groundwater discharge
- Saline geothermal systems

ABSTRACT

Discharge of deeply sourced groundwater to streams is difficult to locate and quantify, particularly where both discrete and diffuse discharge points exist, but diffuse discharge is one of the primary controls on solute budgets in mountainous watersheds. The noble gas helium is a unique identifier of deep groundwater discharge because groundwater with long residence times is commonly enriched in helium. In this study, a portable mass spectrometer was used to measure longitudinal variation in dissolved helium concentrations in two mountainous rivers at high spatial resolution not feasible with traditional sampling techniques. Helium profiles were then simulated using a mass-balance model to quantify longitudinal variation in groundwater discharge to the receiving rivers. Results indicate helium concentrations were enriched by multiple orders of magnitude above atmospheric equilibrium in both rivers and that this persisted for up to 18 km below observed pulse inputs in the Colorado River. Helium mass-balance models match observed longitudinal patterns with the exception of sharp initial increases in helium observed in the rivers. Increased longitudinal groundwater discharge rates correspond to mapped geologic structures in both watersheds that likely transport deep geothermal water. Models show variable sensitivity to spatial assignment of input variables representing the groundwater source, illustrating the importance of collecting data from discrete groundwater discharges where possible. The methodology shows promise for field experiments designed to assess air–water exchange rates and to quantify total groundwater discharge from a combination of discrete and diffuse sources.

1. Introduction

Groundwater discharge to streams maintains streamflow in periods of drought and represents an important source of water for human and ecological use (Gleeson and Richter, 2017; Swanson et al., 2021). Groundwater discharge also moderates stream temperatures and creates ecological refugia vital for fish habitat (Briggs et al., 2018). Sustainability of groundwater discharge to streams under possible future variation in climate will support increased understanding of surface water sustainability in areas where streamflow is groundwater dominated such as the central Rocky Mountains of the United States (USA; Rumsey et al., 2015). Groundwater discharge to streams may be more susceptible to climatic variation where groundwater discharge is young and sourced from a shallow aquifer, and recent analysis indicated that groundwater storage and discharge in mountainous catchments is a reliable predictor of total surface water discharge (Wolf et al., 2023). A

* Corresponding author.
E-mail address: cpnewman@usgs.gov (C.P. Newman).
1 ORCID: 0000-0002-6978-3440.
2 ORCID: 0000-0003-0664-001X.
3 ORCID: 0000-0001-6370-7124.

https://doi.org/10.1016/j.jhydrol.2024.131717

Available online 25 July 2024
0022-1694/Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
recent evaluation of the age of groundwater discharge globally suggests that most groundwater discharge to streams is young (less than three months old; Jasechko et al., 2016) and sourced from relatively shallow depths (less than 500 m; Ferguson et al., 2023), indicating that patterns of groundwater discharge could shift globally in response to climatic variability.

Despite recent evidence suggesting that discharge of old and deep groundwater to streams accounts for a minority of global streamflow (Ferguson et al., 2023), there are regions throughout the globe where discharge of deeply sourced groundwater has been identified as important to local and regional hydrologic budgets (Beisner et al., 2018; Frisbee et al., 2017; Gardner et al., 2011; Luijendijk et al., 2020; Smerdon et al., 2012). Additionally, recent work suggests that a substantial portion of the global solute discharge to the ocean is derived from groundwater (Wood et al., 2023). Because of long residence times in the subsurface, old groundwater may be highly concentrated in solutes from rock weathering (Beyer et al., 2016; Zhu, 2005). Discharge of highly concentrated old groundwater in the Colorado River Basin of the Rocky Mountains is partially responsible for high loads of salinity (Miller et al., 2017; Mast and Terry, 2019) which could cause economic and ecosystem damages. Therefore, although the volumetric flux of old groundwater may be small, this discharge may provide a substantial flux of solutes from the deep to shallow hydrosphere (Mayfield et al., 2021; Wood et al., 2023).

Because groundwater discharge to streams can occur via discrete and diffuse sources (Winter et al., 1999), it may be difficult to quantify all groundwater discharge within a given study reach. Both physical and chemical techniques may be used to identify and quantify groundwater discharge locations (Rosenberry and LaBaugh, 2008). Of the variety of chemical techniques, dissolved gases have utility in estimating groundwater discharges because of disequilibrium between gases in influent groundwater and gases in surface water that are generally in equilibrium with the atmosphere (Adyasari et al., 2023). Dissolved gas tracers are useful for both quantifying the volume and understanding age distributions of groundwater discharging to streams (Sanford et al., 2015; Stolp et al., 2010). Tracers that are unique indicators of young groundwater discharge include radon (Rn; days to weeks old; Adyasari et al., 2023) and tritium (H; months to decades old; Duvert et al., 2016) whereas discharge of old groundwater (centuries to millennia old) may be identified by carbon-14 (14C; Bourke et al., 2014) or helium-4 (He; Beisner et al., 2018; Gardner et al., 2011; Gardner and Susong, 2019). Quantifying both discrete and diffuse groundwater discharge to streams is important for the Colorado River Basin where the majority of streamflow during low-flow periods may be derived from groundwater discharge (Rumsey et al., 2015) and where saline geothermal systems contribute substantial salinity (Miller et al., 2017).

Given the utility of He in identifying diffuse groundwater discharge (Beisner et al., 2018; Gardner et al., 2011; Gardner and Susong, 2019) and the importance of diffuse saline geothermal groundwater discharge within the Colorado River Basin, a portable noble gas mass spectrometer (miniRUEDI, Brennwald et al., 2016) was tested for boat-mounted use in mountainous river settings to both spatially identify discharge zones and estimate total groundwater fluxes. This application is novel in the high-resolution measurements of noble gases possible using the portable mass spectrometer, and the fine-scale resolution of groundwater flux estimates that may be achieved using these data. Previous studies (Avery et al., 2018; Gardner et al., 2011; Gardner and Susong, 2019; Gleeson et al., 2018; Smerdon and Gardner, 2022) have demonstrated the use of noble gases in both definition and quantification of groundwater discharge, but those studies generally utilized discrete samples collected in the field and analyzed in laboratories. Discrete sample collection may limit the spatial resolution of dissolved gas data collection, whereas when the portable mass spectrometer is used, dissolved gas data may be collected at high spatial resolution and in near real time. The portable mass spectrometer also has been applied to high resolution characterization in previous studies of surface waters (Doda et al., 2024; Langenegger et al., 2019; Maier et al., 2022; Weber et al., 2018) but not to quantify diffuse groundwater discharge rates. This work tests the method’s application to high gradient surface waters to provide a highly spatially detailed understanding of groundwater discharge. In contrast to the methodology explored herein, it may be possible to identify saline geothermal inputs to rivers using specific conductance or temperature probes. Although geothermal groundwater discharge displays high contrasts with surface water in these parameters, the contrast is not as large as that between discharging groundwater and surface water with respect to noble gas concentrations. Additionally, mixing of groundwater discharge with river water convolutes and spreads signals in both specific conductance and temperature, whereas the degassing behavior of noble gases provides more unique identification of discrete discharge zones (Avery et al., 2018; Beisner et al., 2018). This study was therefore focused on using dissolved He data alone to evaluate the ability of this dataset to identify and quantify groundwater discharge in mountain rivers.

1.1. Study sites

Two study sites were selected for data collection during a low flow period in October 2022 (Fig. 1). Data collection at low flow when groundwater discharge likely comprises a large portion of streamflow (Rumsey et al., 2015) increasing the relative signal of the discharging groundwater. Both study sites are river reaches where there are known geothermal inputs to surface waters and where there was existing noble gas data for discrete springs (Nelson et al., 2009; Newman et al., 2024; U.S. Geological Survey, 2024). These two study sites are the Virgin River near Pah Tempe Springs, Utah, USA and the Colorado River near Glenwood Springs, Colorado, USA.

The Virgin River begins in southwestern Utah, USA, and flows southwest to a confluence with the Colorado River at Lake Mead. Streamflow in the Virgin River within the study area upstream from the geothermal system is approximately 150 L per second (L/s) during the low-flow period of this study (U.S. Geological Survey, 2024). Near La Verkin, Utah, the Virgin River flows through Timpoweap Canyon over Permian-aged carbonate rocks where it intersects Pah Tempe Hot Springs (Gerner and Thiros, 2014). Pah Tempe Hot Springs (also known as La Verkin Springs or Dixie Hot Springs) discharges from highly fractured carbonate rocks along the footwall damage zone of the Hurricane Fault (Lund et al., 2002; Biek, 2003; Dutson, 2005). The entire spring complex is located along approximately 600 m of the stream and discharges about 260 L/s of water at a temperature of 40 °C with an estimated 90 kilotons of total dissolved solids per year (Gerner and Thiros, 2014). Water from Pah Tempe hot springs is a mixture of deep meteoric groundwater, with a circulation depth >3–5 km into basement rocks and shallow groundwater recharged from the Virgin River (Nelson et al., 2009; Gerner and Thiros, 2014). Discharge from Pah Tempe occurs at over 10 discrete spring orifices as well as diffuse inflow along the riverbed as evident by the formation of gas bubbles. The springs have been well-characterized from previous studies and provide an ideal site to evaluate the miniRUEDI for identifying and quantifying geothermal discharge.

The area near Glenwood Springs is characterized by a several hundred meter deep canyon (Glenwood Canyon) that has been cut by the Colorado River over a period of approximately 1.4 million years (Polyak et al., 2013). Discharge of the Colorado River within the study area is approximately 39,700 L/s during the low-flow period of this study (U.S. Geological Survey, 2024). Saline geothermal springs occur at both the eastern and western ends of Glenwood Canyon, with the Dotsero Springs geothermal system on the east and the Glenwood Springs geothermal system on the west (Eisenbauer, 1983; 1986). These springs constitute several of the largest point sources of salinity to the Upper Colorado River (Miller et al., 2017). The salinity of the springs combined with geologic mapping has been used to develop a conceptual model wherein spring water is derived from both the Eagle Valley Evaporite (composed
of interlayered gypsum, halite, and other salts) and the Leadville Limestone (Eisenhauer, 1986; Kirkham et al., 2002; Maslyn et al., 2017). More than 20 discrete springs discharge along the riverbank in the Glenwood Springs system in less than 2 km, including one of the largest springs by discharge volume in the state of Colorado (Eisenhauer, 1986), with discharge of approximately 185 L/s (Geldon, 1989). In addition to these numerous mapped discrete springs, field investigation also indicates areas of diffuse discharge along the riverbed, which can be identified by formation of gas bubbles and handheld infrared thermal cameras. These characteristics of both discrete and diffuse geothermal groundwater discharge to the river makes the Colorado River near Glenwood Springs an ideal case study for the miniRUEDI.

2. Methods

2.1. Dissolved gas sampling and analysis

Continuous river gas concentration measurements were made using a portable mass spectrometric system, known as the “miniRUEDI” (Brennwald et al., 2016), that utilizes the gas-equilibrium membrane-inlet mass-spectrometry (GE-MIMS) technique for quantification of dissolved gases in water. The miniRUEDI measures the partial pressures
of dissolved gas species (He, Ar, Kr, N\textsubscript{2}, O\textsubscript{2}, CO\textsubscript{2}, CH\textsubscript{4}, etc.) for determination of gas concentrations in aqueous matrices, can be calibrated on-site via ambient air, and has been utilized in springs, groundwater, lakes, and rivers (Doda et al., 2024; Langenegger et al., 2019; Lightfoot et al., 2022; Maier et al., 2022; Schilling et al., 2023; Weber et al., 2018). For use in aqueous settings, the miniRUEDI requires water to be pumped through the GE-MIMS membrane module. The gas species dissolved in the water equilibrate with a small gas headspace in the module. A small fraction of this headspace gas is transferred through a pressure reducing system into the mass spectrometer where the gases are analyzed. Analytical uncertainty is estimated to be 1–3% (Bremwald et al., 2016).

In the Virgin River and Colorado River study areas, the miniRUEDI was secured to inflatable rafts to make continuous gas measurements along the rivers. River water was pumped through a high-capacity particle filter and into the GE-MIMS module using a positive pressure submersible pump and out a discharge tube located at the back of the raft. The miniRUEDI sample data were calibrated using ambient air as a standard gas measured after every 20 sample measurements (approximately once per hour). Although the miniRUEDI makes continuous measurements, the time required to detect individual peaks means that a particular gas species is measured every 3–5 min. A global positioning system (GPS) device was used to precisely locate river gas measurements.

In addition to on-site dissolved gas analysis, discrete samples were collected from springs within both watersheds. These discrete samples provide the elemental and isotopic noble-gas composition of the groundwater endmember used in mass-balance modeling. In the Virgin River/Pah Tempe study area, discrete spring samples are reported in Nelson et al. (2009). In the Colorado River/Glenwood Springs area, discrete samples were collected from six springs as part of this study, four of which discharge directly to the river whereas the other two are located between 200 m and 1 km away from the river. Discrete samples were collected for laboratory noble gas analysis in copper tubes as described in Aeschbach-Hertig and Solomon (2013) and were analyzed by magnetic-sector mass spectrometry and ultralow vacuum extraction line according to Hunt (2015). Description of discrete springs and analytical data used as inputs to the helium mass-balance modeling are reported in Table 1. All analytical results are published in Newman et al. (2024).

2.2. Helium mass-balance modeling

The dissolved He concentrations in river water derived from high-resolution analysis with the portable mass spectrometer were used to estimate groundwater inflow (Fig. 2) according to generalized mass balance (Eq. (1) and water mass balance (Eq. (2); Beisner et al., 2018; Smerdon and Gardner, 2021; Smendor and Gardner, 2022):

\[
\frac{\partial C}{\partial x} = \frac{A}{Q} \left( \frac{\partial C}{\partial x} \right) + \frac{q_{gw} w}{Q} (C_{gw} - C) + \frac{Q_{tr}}{\Delta x Q} (C_{tr} - C) - \frac{k w}{Q} (C - C_{AEW}) - \frac{A}{Q} \frac{\partial C}{\partial x},
\]

(1)

where \(C\) is the He concentration in the stream (mol/m\textsuperscript{3}), \(x\) is the flow distance (m), \(D\) is the longitudinal hydrodynamic dispersivity (m\textsuperscript{2}/s), \(A\) is the stream cross-sectional area (m\textsuperscript{2}), \(Q\) is the stream discharge (m\textsuperscript{3}/s), \(q_{gw}\) is the specific groundwater influx (m/s), \(w\) is the stream width (m), \(C_{gw}\) is the He concentration in the inflowing groundwater (mol/m\textsuperscript{3}), \(C_{go}\) is the river loss to groundwater (m/s), \(k\) is the gas exchange velocity (m/s), \(C_{gw}\) is the equilibrium atmospheric concentration (mol/m\textsuperscript{3}) given by Henry’s Law \([C_{gw} = H(T,S)P_{am}, P_{am}\) is the atmospheric partial pressure of the gas (atm), \(H(T,S)\) is the Henry coefficient (mol/m\textsuperscript{3}/atm)], \(\lambda\) is the decay coefficient (s\textsuperscript{-1}), \(Q_{tr}\) is the influent tributary discharge (m\textsuperscript{3}/s), \(C_{tr}\) is the He concentration in the tributary water (mol/m\textsuperscript{3}), and \(\Delta x\) is the model grid cell spacing. Surface water mass balance (Eq. (2)) describes the influence of precipitation (P) and evaporation (E) on water balance. Equations (1) and (2) describe the mass balance of tracer and water for a stream with groundwater exchange, tributary inflow and extraction, gas exchange and first order decay. Derivations of the mass-balance equations are given in the supplemental material of Smerdon and Gardner (2022). Several simplifications can be made to these equations, namely that no decay occurs because He is a stable gas (making term \(\frac{\partial C}{\partial x}\) in Eq. (1) equal to zero), that no surface water is lost to groundwater in the study reach (making term \(q_{gw} w\) in Eq. (2) equal to zero), and that no precipitation occurs during the sampling (making term \(P - w\) in Eq. (2) equal to zero).

Table 1
Discrete noble-gas samples from the Colorado River and Virgin River study areas. R/R\textsubscript{He} denotes the helium-3 to helium-4 ratio in the sample (R) to the helium-3 to helium-4 ratio in air (R\textsubscript{He}, a value of 1.38 x 10\textsuperscript{-6} was used in this study). STP denotes standard temperature and pressure, which for the portable noble gas mass spectrometer used in this study is defined as 22.414 cm\textsuperscript{3} STP=1 Mol.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Site Name</th>
<th>Type</th>
<th>Longitudinal Distance along river (m)</th>
<th>Sample Date</th>
<th>(^{3}\text{He} \text{(cm}^{3}\text{STP/g H}_{2}\text{O}))</th>
<th>R/R\textsubscript{He}</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado River</td>
<td>S. Spring 4</td>
<td>Spring</td>
<td>6,595</td>
<td>12 November 2021</td>
<td>6.63 x 10\textsuperscript{-5}</td>
<td>0.099</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Colorado River</td>
<td>S. Spring 5</td>
<td>Spring</td>
<td>6,916</td>
<td>12 November 2021</td>
<td>5.98 x 10\textsuperscript{-5}</td>
<td>0.098</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Colorado River</td>
<td>Travertine Island</td>
<td>Spring</td>
<td>7,041</td>
<td>19 October 2022</td>
<td>6.48 x 10\textsuperscript{-5}</td>
<td>0.101</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Colorado River</td>
<td>S. Spring 3</td>
<td>Spring</td>
<td>7,063</td>
<td>10 August 2021</td>
<td>5.75 x 10\textsuperscript{-5}</td>
<td>0.095</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Colorado River</td>
<td>CDOT Spring</td>
<td>Spring</td>
<td>8,286</td>
<td>27 October 2021</td>
<td>2.81 x 10\textsuperscript{-5}</td>
<td>0.098</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Colorado River</td>
<td>South Canyon HS</td>
<td>Spring</td>
<td>14,848</td>
<td>11 August 2021</td>
<td>3.09 x 10\textsuperscript{-5}</td>
<td>0.156</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Virgin River</td>
<td>LV101</td>
<td>Well</td>
<td>0</td>
<td>9 November 2018</td>
<td>2.43 x 10\textsuperscript{-5}</td>
<td>0.101</td>
<td>Newman et al. (2024)</td>
</tr>
<tr>
<td>Virgin River</td>
<td>Pah Tempe 1.2</td>
<td>Spring</td>
<td>140</td>
<td>14 March 2003</td>
<td>2.13 x 10\textsuperscript{-5}</td>
<td>0.099</td>
<td>Nelson et al. (2009)</td>
</tr>
<tr>
<td>Virgin River</td>
<td>Virgin 2</td>
<td>Spring</td>
<td>128</td>
<td>14 March 2003</td>
<td>2.94 x 10\textsuperscript{-5}</td>
<td>0.102</td>
<td>Nelson et al. (2009)</td>
</tr>
</tbody>
</table>
Inverse modeling to estimate $q_{gi}$ was completed using the StreamTran package (Smeardon and Gardner, 2022) of Python (Van Rossum and Drake, 2009). All input data and the model code are provided in Newman et al. (2024). The values for river width ($w$) were extracted from the surface water and ocean topography mission (SWOT) online database (Altenau et al., 2021) for the Colorado River through Glenwood Springs. The Virgin River is not observable from SWOT, the width was instead estimated from satellite imagery. Values for depth ($d$) were estimated during field investigations, and the cross-sectional area was calculated as $w \times d$. Values of $Q$ and $Q_{tr}$ were measured at specific locations along both river transects. Cross-sectional area was then combined with $Q$ to produce a mean velocity vector. The values of $q_{go}$ were set to zero. The concentration of gas in equilibrium with the atmosphere ($C_{aew}$) was calculated using the average reach elevation and measured water temperature using solubility relations compiled in Ballentine et al. (2002). The value of $C_{tr}$ was assumed to equal $C_{aew}$ based on no observed geothermal discharge within the tributaries. Values of $C_{gw}$ were allowed to vary spatially based on measurement of noble gases in discrete springs (Table 1). With the assignment of variables as described above, there are two primary sources of uncertainty in the mass-balance model, the $^4$He content of discharging groundwater ($C_{gw}$) and gas exchange velocity ($k$).

Model sensitivity analysis was completed to determine the effect of groundwater gas concentrations ($C_{gw}$) on results of both simulated $^4$He in the rivers and of the groundwater discharge flux ($q_{gi}$). Discrete springs in both the Virgin and Colorado Rivers allow for direct measurement of $C_{gw}$, but spatial heterogeneity in $C_{gw}$ may be a limitation of the modeling approach (Atkinson et al., 2014). Spatial variation in groundwater geochemistry is evident in springs of the Glenwood Springs geothermal system, which have $^4$He concentrations that range over three orders of magnitude (Table 1). To evaluate the sensitivity of model results to heterogeneous $C_{gw}$, one model iteration used the spatially variable $C_{gw}$ values (spatially variable model hereafter), and one used the median of all observed $C_{gw}$ to represent input of groundwater with a homogeneous composition (homogeneous model hereafter) along the study reaches.

To account for uncertainty in gas exchange, an initial $k$ was estimated based on methods described in Raymond et al. (2012). Although the method in Raymond et al. (2012) allows estimation of $k$ based on the characteristics of the site (namely slope, velocity, depth, and discharge), compilation of $k$ constants from various environments (Gleeson et al., 2018) highlights the uncertainty in this variable. Uncertainty in $k$ increases uncertainty in the $^4$He mass-balance model. To account for uncertainty in $k$, a Monte Carlo analysis was conducted centered on the value calculated using the method of Raymond et al. (2012), $k = 4.94$ m/day for the Virgin River and $k = 21.6$ m/day for the Colorado River, and using a standard deviation $= 40$ m/s. These $k$ values, calculated using the method of Raymond et al. (2012), are within the range of values summarized in Gleeson et al. (2018) for a range of environments. One hundred samples of the $k$ distribution were used to create He mass-balance model realizations. This method therefore accounts for uncertainty in $k$ and propagates that uncertainty to understand uncertainty in the resulting calculated groundwater discharge rates. Beisner et al. (2018) completed a Bayesian uncertainty analysis of the effect of unknown $k$ values on groundwater discharge in a similar river environment and found that the resulting distribution of calculated groundwater inflow values was not highly sensitive to the $k$ value.

Fig. 3. Results of Virgin River dissolved gas survey including (a) map of sampling locations and $^4$He concentrations and (b) longitudinal profile of $^4$He in river water compared to discrete spring values and atmospheric equilibrium.
3. Results

3.1. Virgin River/Pah Tempe geothermal system

Results from the longitudinal dissolved-gas survey of the Virgin River as it flows past the Pah Tempe geothermal system (Fig. 3) indicate that the river is in approximate equilibrium with $^4\text{He}$ in the atmosphere upstream from the eastern-most mapped springs, followed by an increase in the order of magnitude along only several hundred meters. This large pulse input of elevated $^4\text{He}$ then degasses along the next several hundred meters before another input causes $^4\text{He}$ concentrations to increase again between 500 and 750 m downstream. At approximately 960 m downstream, there is a step change in the data where data collection on the first day ended and began in the same location on the next day. At this location, there is an apparent decrease in the $^4\text{He}$ concentration by a factor of approximately 20. Changes in $^4\text{He}$ concentrations in river water from $7.38 \times 10^{-7}$ cm$^3$STP/g at the conclusion of the first day to $3.23 \times 10^{-8}$ cm$^3$STP/g at the start of the second day (at sampling locations that are 10 m apart; approximately 960 m downstream along the study reach) could be linked to incomplete lateral mixing of He in the river, to a lesser degree diel temperature cycles and the temperature dependent solubility of He, or to sample equilibration within the portable mass spectrometer. These several data points were not included in the model because they are likely not representative of the river. Even disregarding the $^4\text{He}$ concentrations of approximately $3.00 \times 10^{-8}$ cm$^3$STP/g as an artifact of incomplete mixing, there is a step change in this location. The approximately 5-fold decrease is more consistent with solubility changes or analytical latency. Solubility of He is temperature dependent, with increased solubility at lower temperature (Weiss, 1971). It would be expected that water temperature in the river during the morning would be lower than at the end of the previous day, but this would lead to a step change in the opposite direction of that observed. Temperature of the river water was not recorded during the survey, but understanding how river temperatures changed between data collected on sequential days could assist with interpretation of step changes. Additional studies could consider collecting continuous water temperature datasets along with noble gas surveys. Lastly, the equilibration period for the gas membrane unit of the mass spectrometer is approximately 15 min although analysis of gas samples can be completed every 3 to 5 min (Brennwald et al., 2016). Data collection periods ranged from 3 to more than 15 min; and therefore, it is possible that some of the measurements completed in more rapid succession are influenced by minor carryover from previous measurements. Additional studies could consider spacing all measurements by 15 min to allow full equilibration within the membrane.

Comparison of observed and simulated $^4\text{He}$ profiles for the Virgin River (Fig. 4) illustrates that the numerical model cannot reproduce the rapid increase in $^4\text{He}$ concentrations along the short distance observed in the data. Concentrations of $^3\text{He}$ in the river water increase two orders of magnitude in approximately 100 m, whereas model simulations require approximately 250 m to increase. Simulated groundwater inflow at the upper end of the model domain is also elevated, even though there are no thermal springs in that reach. High initial simulated groundwater inflow rates are an artifact of a few data points above the thermal inflow zone. Additional studies could consider beginning data collection at a greater distance from known groundwater discharge points (i.e., several hundred meters as was completed for the Colorado River area discussed next). This distance can provide the model with an equilibration period and could improve model fit. Downstream from the simulated peak in $^4\text{He}$ concentrations, the observations continue to be on the edge of the confidence intervals. Slight overprediction may be caused by including data from two different days in the model runs where river concentrations were affected by possible incomplete mixing, diel temperature changes, or latency in the gas equilibration module. The modeling framework does not currently have the ability to simulate spatially variable water temperatures, though this addition could allow for data from the different days to be more realistically simulated. Comparison of the spatially variable and homogeneous groundwater $^4\text{He}$ models for the Virgin River indicated no substantial difference in simulation output based on the model framework, therefore only the results from the spatially variable model are illustrated in Fig. 4. All model results are available in Newman et al. (2024). The similarity in simulations regardless of the assumption of the $^4\text{He}$ content of discharging groundwater is likely because there is less spatial variability in $^4\text{He}$ observed in springs in the Virgin River area than in the Colorado River area (Table 1). Although the model does not reproduce the exact observed

---

Fig. 4. Results of helium mass-balance model for the Virgin River with spatially heterogeneous $C_{gw}$ including (a) simulated and observed longitudinal profile of $^4\text{He}$, and (b) simulated longitudinal profile of groundwater discharge ($q_{gw}$). Only results of the spatially variable groundwater $^4\text{He}$ model are shown.
3.2. Colorado River/Glenwood springs geothermal system

Results from the longitudinal dissolved-gas survey of the Colorado River as it flows past the Glenwood Springs geothermal system indicate that the river is in approximate equilibrium with \(^{4}\text{He}\) in the atmosphere upstream from the eastern-most mapped springs (Fig. 5). As the river water passes several large springs along the southern bank, the \(^{4}\text{He}\) in river water increases by an order of magnitude along approximately 1 km. After this large pulse of \(^{4}\text{He}\) enters the river, the concentration decreases steadily along the next 18 km, though at the end of the study reach the concentration is still above atmospheric equilibrium (Fig. 5). This substantial pulse of \(^{4}\text{He}\) and long-lived anomaly is likely the first such dataset described in the literature. The extended pulse of \(^{4}\text{He}\) in the Colorado River (Fig. 5) compared to the Virgin River (Fig. 3) could be caused by sustained groundwater discharge to the Colorado River or slower degassing in the Colorado River owing to the greater water column depth than the Virgin River.

Helium mass-balance model results for the Colorado River (Fig. 6) indicate that both models of spatially variable and homogeneous \(^{4}\text{He}\) in groundwater generally match the overall pattern observed in the longitudinal profile of \(^{4}\text{He}\) in the river, with the exception of the distance over which the initial pulse of \(^{4}\text{He}\) is observed and the maximum values. Measurements indicate that \(^{4}\text{He}\) concentrations increase along approximately 1 km whereas the model simulations require approximately 2.5 km to increase from atmospheric equilibrium to peak concentrations. This lag effect could be caused by the model discretization or by the treatment of boundary conditions such as \(k\), \(A\), and \(d\) in a homogeneous manner within the model. Sensitivity analysis using variable discretization (\(\Delta x\) in Eq. (1)) indicated that the model was relatively insensitive to discretization. Model runs using both 1,000 cells and 100,000 cells illustrated that in both cases a lag existed in the simulated \(^{4}\text{He}\), though the model with more cells required substantially longer runtimes. The \(^{4}\text{He}\) mass-balance equation (Eq. (1)) requires estimates of river depth and width, which in turn dictate velocity through integration with discharge. Width was assigned based on data from SWOT (Altenau et al., 2021), but a single depth value was estimated for the entire study reach. It is possible that more granular definition of river depth, and therefore velocity, from observations along the study section could improve model fit, particularly around the initial pulse of \(^{4}\text{He}\).

Simulated \(q_{\text{gi}}\) values differ substantially between the spatially variable and homogeneous models. Both sets of simulations indicate substantial groundwater discharge at approximately 5,000 m, approximately coincident with and slightly upstream from the pulse of observed \(^{4}\text{He}\) in the river. Downstream from 5,000 m, the simulated groundwater discharge in the models diverges, with the spatially variable model indicating multiple zones of increased groundwater discharge at approximately 7,000 to 9,000 m and 13,000 to 15,000 m whereas groundwater discharge in the homogeneous model decreases throughout the remainder of the stream reach. The most pronounced difference between the models occurs at the downstream end of the reach where simulated groundwater discharge in the spatially variable \(^{4}\text{He}\) model increases by two orders of magnitude. This increase in simulated groundwater discharge is caused by the inclusion of the sample from South Canyon Hot Springs in the model inputs. This hot spring is located approximately 1,000 m from the river (site S6 on Fig. 1) and has substantially lower \(^{4}\text{He}\) concentrations than other hot springs occurring immediately adjacent to or within the Colorado River (Table 1). Although this spring represents local geothermal discharge, based on the difference in geochemical composition, it is unlikely that it represents discharge from the same geothermal upwelling zone that discharges directly to the Colorado River in the Glenwood Springs area. The differences in model output depending on which springs are considered in model inputs illustrates the importance of collecting spatially distributed information on groundwater inflows, specifically the \(^{4}\text{He}\) content and spring discharge estimates (if available). Simulated groundwater discharge also increases slightly in the vicinity of the West Glenwood Fault, similar to the simulation of the Virgin River near the vicinity of the Hurricane Fault Zone (Fig. 4b), are consistent with previous studies indicating structurally controlled groundwater discharge (Lund et al., 2002; Biek, 2003; Dutson, 2005).

Fig. 5. Results of Colorado River dissolved-gas survey including (a) map of sampling locations and \(^{4}\text{He}\) concentrations and (b) longitudinal profile of \(^{4}\text{He}\) in river water compared to discrete spring values and atmospheric equilibrium.
4. Identifying and quantifying deep groundwater discharge.

Recent research has indicated that groundwater discharge to streams may be dominated by groundwater sourced from relatively shallow depths (<500 m; Ferguson et al., 2023). These estimates are corroborated by observation of shallow active groundwater circulation in mountain watersheds dominated by sedimentary lithologies (10 s of m; Manning et al., 2020). Despite the possibility that global groundwater discharge to streams may be dominated by shallow groundwater, the discharge of old, deep groundwater within geothermal systems and highly incised mountainous terrain has been well established (Diamond et al., 2018; Gleeson and Manning, 2008; McIntosh and Ferguson, 2021). In some instances, groundwater discharge from deep systems contributes substantial solute loads to receiving water bodies (Mast and Terry, 2019; Wood et al., 2023), potentially causing ecological harm and degrading water quality. The somewhat contradictory nature of the established knowledge of deep groundwater circulation and discharge with small estimates of global deep discharge volumes may be due to the difficulty of observing deep groundwater flow paths (Condon et al., 2020), and the difficulty and complexity of quantifying the discharge of old groundwater to surface waters (Bourke et al., 2014). Differential gaging (Chafin and Butler, 2002), specific conductance mass balance (Rumsey et al., 2015), or stable isotopes (Tuttle and Grauch, 2009) each provide methods by which to estimate deep groundwater inputs (if appropriate endmembers may be estimated), but these methods generally have high uncertainty or require complex laboratory analyses. Smerdon and Gardner (2022) recently illustrated the utility of environmental tracer (noble gas, $^4$He, etc.) sampling of river waters for hydrologic investigation of large areas. Application of the portable mass spectrometer extends this methodology.

The ability of river-borne portable mass spectrometer measurements to identify deep groundwater discharge with elevated He, in this study manifesting as geothermal inflows, with high spatial precision is observed in both the Virgin (Fig. 3) and Colorado (Fig. 5) Rivers. The approach is scalable over multiple orders of magnitude in river discharge, with the discharge of the Virgin River upstream from the geothermal inputs being approximately 150 L/s, whereas the Colorado River discharge was approximately 39,700 L/s (U.S. Geological Survey, 2024). This approach allowed for more than 50 noble gas measurements to be made in both river systems (with study reaches ranging from 3 km to 25 km), a dataset that would be logistically difficult for most studies using standard copper tube or diffusion samplers collected for laboratory analysis unless precise locations of step changes in He content of river water could be known a priori. These datasets illustrate the high degree of spatial variability in noble gas compositions of river water receiving geothermal inputs, a pattern that would most likely not be observable with less spatially robust data. In particular, the ability of the portable mass spectrometer to define spatial variability means that the method shows promise for reconnaissance studies of large watersheds (e.g., Smerdon and Gardner, 2022). It is also possible that specific conductance or temperature could be used to identify groundwater discharge zones (Cook, 2013), but because these tracers are more conservatively transported in the surface water than the degassing He, it is unlikely that high-resolution identification of specific groundwater discharge zones, such as those observed in the Virgin River, could be uniquely identified.

The nature of the identified geothermal discharge zones is also of note. In both the Virgin River and Colorado River, $^4$He anomalies manifested both as broad zones that were apparently unrelated to geologic structures and also as focused anomalies that occurred in close proximity to mapped faults. In some respects, the geothermal discharge along faults could be considered discrete discharge zones because discharge is focused along discrete physical features. In the hydrologic sense however, these discharge zones are still diffuse because geothermal water discharges across large areas of the riverbed, as verified by field observations that show bubble formation and thermal anomalies over zones that are multiple meters in length. The portable mass spectrometer is able to detect $^4$He anomalies from both discrete spring discharges along the edges of the river and from diffuse zones...
along the riverbed.

Mass-balance model simulations were used to quantify groundwater discharge to both the Virgin and Colorado Rivers. The models represent the spatially variable patterns of $^4$He in both river systems relatively well (as illustrated by the overlap of 95% confidence intervals with data points). However, the models were unable to reproduce rapid increases in $^4$He increases at the upstream end of both study reaches where large discrete springs are known to contribute geothermal water. Simulated groundwater discharges in both study reaches correspond broadly to known or hypothesized structural conduits for groundwater flow. With this knowledge of the general locations of geothermal discharge points, the portable mass spectrometer could be reused on a finer spatial scale to map localized discharge zones along these faults. These results corroborate previous applications (Beisner et al., 2018; Gardner and Susong, 2019; Smerdon and Gardner, 2022) of the mass-balance modeling methodology being appropriate for estimation of diffuse deep groundwater discharge to streams where there is uncertainty in model parameters. It is noteworthy that the model results are sensitive to the product of groundwater discharge ($q_{gw}$) and groundwater concentration ($C_{gw}$), see Eq. (11) and that recalcing the results in combined discharge-concentration space could be used to further explore model uncertainty, where spatially distributed groundwater concentration data are available.

In addition to precise spatial identification of discharge zones, the high-resolution sampling also likely reduces the error in groundwater inflow estimates derived from the $^4$He mass-balance model. Cook (2013) derives relations between sample spacing and groundwater discharge uncertainty, with results indicating that inflow estimates have greater error at greater sample spacings. Most longitudinal studies of dissolved gases in rivers use sample spacings of multiple kilometers (Beisner et al., 2018; Cook, 2013; Gardner et al., 2011; Smerdon and Gardner, 2022). Utilizing the portable mass spectrometer allows sample spacings on the order of 10 s to 100 s of meters, even for long study reaches, which likely reduces the uncertainty of groundwater discharge estimates.

4.2. Applicability and possibilities for additional studies

The ability of both models to fit abrupt increases in observed $^4$He concentrations could be improved by additional data collection in the field. The mass-balance model uses the geometry of the river in model inputs, namely depth, width, and discharge. These values are used to define the velocity field for the river which in turn is used in the calculation of $k$, the gas exchange velocity, from $^4$He inputs to quantify the product of measured dissolved $^4$He concentrations and groundwater discharge uncertainty, with results indicating that inflow estimates have greater error at greater sample spacings. Most longitudinal studies of dissolved gases in rivers use sample spacings of multiple kilometers (Beisner et al., 2018; Cook, 2013; Gardner et al., 2011; Smerdon and Gardner, 2022). Utilizing the portable mass spectrometer allows sample spacings on the order of 10 s to 100 s of meters, even for long study reaches, which likely reduces the uncertainty of groundwater discharge estimates.

Although this study directly identifies the presence of deep groundwater discharge to both the Virgin and Colorado Rivers and estimates the discharge quantity, the age of the discharging groundwater was not addressed. Recent studies utilizing distributed samples of environmental tracers in streams (Beisner et al., 2018; Smerdon and Gardner, 2022), with larger spatial gaps between samples than the current study, illustrate that the age of discharge groundwater may vary substantially along a river reach. Spatially variable groundwater ages may influence the sustainability of streamflow where streams are groundwater supported (Gleeson and Richter, 2017). The portable noble gas mass spectrometer could be combined with sampling of environmental tracers such as $^3$H, CFCs, and SF$_6$ to quantify both groundwater discharge volumes and the age of the discharging groundwater (i.e., Smerdon and Gardner, 2022). Inclusion of tracers of old groundwater ($^4$He, $^{88}$Kr, $^{14}$C) could be combined with tracers of young groundwater ($^3$H, SF$_6$, $^{39}$Ar) to more fully understand distributions of groundwater ages in watersheds where the combined tools of discrete samples and continuous profiling using the portable mass spectrometer were applied. In the Virgin River and Colorado River study areas, the discharging deep groundwater is of geothermal origin, and unique estimates of residence times of geothermal water are difficult to quantify given the complexity of geochemical reactions within the reservoir (Qiu et al., 2018; Stefánsson et al., 2019). Although constraining watershed-scale residence times in these settings may be complex given the geothermal systems, the approach of combining high-resolution noble gas measurements with discrete samples of other environmental tracers could be utilized in watersheds with old, non-geothermal groundwater upwelling such as the Paradox Valley of Colorado (Mast and Terry, 2019).

As described by previous studies utilizing noble gases as tracers of groundwater/surface-water interactions, one of the primary unknowns in the mass-balance model is the gas exchange velocity, $k$ (Beisner et al., 2018; Gleeson et al., 2018). Some studies directly quantify $k$ for specific study reaches using gas injection approaches (Avery et al., 2018) or estimate via modeling of multiple tracers simultaneously (Gleeson et al., 2018)
to estimate gas exchange velocity is to log dissolved oxygen for several days at one or two sites along the stream reach and fit a similar gas exchange and stream metabolism model after Hall et al. (2016). Another potential experiment to derive site-specific k values could be to pair the portable mass spectrometer with a continuous dye-tracer injection test. Given appropriate considerations for dye mass balance (Runkel, 2015), spatially distributed measurements of dye concentrations could be used to quantify diffuse inflows with relatively low uncertainty because the influent groundwater is known to have a dye concentration of zero. Noble gas concentration observations made with the portable mass spectrometer could then be used with the modeling approach described herein to match the groundwater discharge measured using dye dilution, making the k values the objective of the iterative approach. Another approach to estimate k values could be made by injecting a noble gas species known to be absent from diffuse discharge, such as Ar or Kr in the river reaches within this study. Simultaneous measurement of the loss of injected gas through degassing with gases derived from deep groundwater discharge (He) could allow the gas exchange behavior of both gases to be estimated using the equations in Raymond et al. (2012). It may be beneficial to include gases with substantially different masses (such as He and Kr) because this may allow for differences in gas exchange behavior to be quantified with higher precision. Finally, if groundwater inflow were known a priori to be zero within a specific stream reach using other methodologies, then the term q_m in Eq. (1) could be eliminated, allowing direct calculation of k with spatially distributed longitudinal datasets. It is noteworthy that each of these potential approaches could quantify spatial changes in k, which is presently unaddressed in the literature.

It is also useful to consider the applicability of the methodology to rivers of variable size and with different ⁴He sources. Streamflow of the rivers considered in this study ranged over several orders of magnitude (150 L/s in the Virgin River and 39,700 L/s in the Colorado River). These rivers are relatively small compared with other rivers in North America and around the world that may be experiencing inflow of old groundwater. For instance, low-flow discharge of the Columbia River in the northwestern United States is approximately 3 x 10⁶ L/s (U.S. Geological Survey, 2024), multiple orders of magnitude larger than the Colorado River. Groundwater in the Columbia River basalt aquifers can have terrigenic ⁴He concentrations up to 2.5 x 10⁴ cm³STP/g (Johnson et al., 2024). If the dynamics of groundwater discharge and air–water exchange were similar between the Columbia River and the Colorado River, then the hypothetical river ⁴He concentration could be considered by simple dilution. Observed ⁴He in the Colorado River reaches 11.25 times the concentration of air equilibrated water (Fig. 5). Dilution of this signal by the approximate greater discharge in the Columbia River versus the Colorado River indicates that ⁴He concentrations in the Columbia could be expected to be up to 15 % greater than atmospheric equilibrium. Given that the analytical uncertainty of the portable mass spectrometer is 1 to 3 % (Brennwald et al., 2016), it is possible that the signal of discharging groundwater could be uniquely identified even in a river as large as the Columbia. These calculations incorporate several simplifying factors, such as the exact dynamics of groundwater discharge, the ⁴He content of the discharging groundwater, and gas exchange with the atmosphere, which are likely to differ between different river systems. Nevertheless, the method could be viable across large ranges in river dynamics and groundwater inflow regimes.

5. Conclusions

Groundwater discharge to streams is a critical control on streamflow generation, water temperature variability, and solute budgets. Recent investigations suggest that most groundwater discharge to streams is young (Ferguson et al., 2023; Jasechko et al., 2016), yet the discharge of old groundwater with high solute concentrations is important from a global mass balance perspective (Mayfield et al., 2021; Wood et al., 2023) and for watershed-scale processes (Frisbee et al., 2017; Mast and Terry, 2019). Diffuse discharge of groundwater, particularly old groundwater, is commonly difficult and costly to quantify (Bourke et al., 2014; Smerdon and Gardner, 2022). In this study, a portable mass spectrometer was used to measure ⁴He concentrations in the Virgin River and Colorado River, both of which have geothermal inputs that increase salinity (degrading water quality) and cause substantial economic damage (Miller et al., 2017). Longitudinal profiles of ⁴He concentrations with high spatial resolution were used in an inverse mass-balance modeling approach to estimate the total groundwater discharge to both reaches. Longitudinal profiles of ⁴He in both rivers indicate substantial ⁴He influx leading to river concentrations enriched multiple orders of magnitude greater than the atmosphere for hundreds of meters in the Virgin River to more than 10 km in the Colorado River. Large fluctuations in ⁴He concentrations along short distances in the Virgin River near Pah Tempe springs illustrate concentrated groundwater discharge zones and relatively rapid degassing of ⁴He to the atmosphere. Contrastingly, the Colorado River near Glenwood Springs is characterized by one substantial pulse of ⁴He that is maintained downstream. This pattern is more characteristic of broad groundwater discharge zones and slower degassing likely linked to the deeper water column in the Colorado River.

Comparison of observed and simulated profiles of ⁴He illustrate that although the mass-balance model is robust for estimating groundwater discharge and for evaluating gas exchange dynamics, additional field data collection could improve model fits. Collection of river geometry measurements coincident with noble gas data could constrain the channel dimensions throughout the study reach, thereby potentially improving model fit through more detailed definition of boundary conditions. Additionally, collecting spatially distributed Rn, ³⁷H, and CFC data could provide another dataset for inclusion in the inversion process and allow spatially variable groundwater ages to be estimated. Finally, the portable mass spectrometer shows promise for use in direct quantification of gas exchange rates using field experiments where other quantities in the mass-balance model can be directly quantified.

CRediT authorship contribution statement

Connor P. Newman: Writing – original draft, Visualization, Investigation, Funding acquisition, Formal analysis, Conceptualization. C. Eric Humphrey: Writing – review & editing, Visualization, Methodology, Investigation, Data curation. Matthias S. Brennwald: Writing – review & editing, Validation, Software, Resources, Methodology, Investigation, Formal analysis. W. Payton Gardner: Writing – review & editing, Validation, Software, Methodology, Investigation, Formal analysis. Kelli M. Palko: Writing – review & editing, Resources, Investigation. Michael Goosell: Writing – review & editing, Resources, Methodology, Investigation. D. Kip Solomon: Writing – review & editing, Validation, Resources, Methodology, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and models will be available at this DOI upon article publication: 10.5066/P9MP8PB9 through a U.S. Geological Survey data release.

Acknowledgements

This study was funded by the U.S. Geological Survey Water Mission...