Title

Direct Observation of the Depth of Active Groundwater Circulation in an Alpine Watershed

Journal

Water Resources Research, 57(2)

ISSN

0043-1397

Authors

Manning, Andrew H

Ball, Lyndsay B

Wanty, Richard B

et al.

Publication Date

2021-02-01

DOI

10.1029/2020wr028548

Peer reviewed

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR028548

Direct Observation of the Depth of Active Groundwater Circulation in an Alpine Watershed

Andrew H. Manning[1], Lyndsay B. Ball[1], Richard B. Wanty[1], and Kenneth H. Williams[2,3]

[1]U.S. Geological Survey, Denver, CO, USA, [2]Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [3]Rocky Mountain Biological Lab, Gothic, CO, USA

Key Points:

- The depth of active groundwater circulation is a fundamental characteristic of mountain watersheds, yet is rarely well constrained
- Hydrogeologic, groundwater chemistry, and $^{3}H/^{3}He$ age data were collected at discrete depths from two boreholes 46 and 81 m deep
- Borehole data combined with shallow groundwater data and numerical modeling indicate an active circulation depth of ~20 m

Supporting Information:

- Supporting Information S1

Correspondence to:
A. H. Manning,
amanning@usgs.gov

Citation:
Manning, A. H., Ball, L. B., Wanty, R. B., & Williams, K. H. (2020). Direct observation of the depth of active groundwater circulation in an alpine watershed. *Water Resources Research*, *56*, e2020WR028548. https://doi.org/10.1029/2020WR028548

Received 4 AUG 2020
Accepted 10 DEC 2020

Abstract The depth of active groundwater circulation is a fundamental control on stream flows and chemistry in mountain watersheds, yet it remains challenging to characterize and is rarely well constrained. We collected hydraulic conductivity, hydraulic head, temperature, chemical, noble gas, and $^{3}H/^{3}He$ groundwater age data from discrete levels in two boreholes 46 and 81 m deep in an alpine watershed, in combination with chemical and age data from shallow groundwater discharge, to discern groundwater flow rates at different depths and directly observe active and inactive groundwater. Vertical head gradients are steep (average of 0.4) and thermal profiles are consistent with typical linear conductive continental geotherms. Groundwater deeper than ~20 m is distinct from shallow groundwater and creek water in its chemistry, noble gas signature, and age (dominantly >65 years compared to <9 years). Together these results suggest low vertical groundwater flow velocities and a relatively shallow active circulation depth of ~20 m. This hypothesis is tested with a simple 2-D numerical fluid flow and heat transport model representing a hillslope transect through the two boreholes. The modeling indicates that the subhorizontally bedded sedimentary rocks underlying the basin are highly anisotropic with low vertical hydraulic conductivity, and at most ~10% of bedrock recharge (equivalent to <2% of stream baseflow) flows below a depth of 20 m. The study demonstrates the considerable value of discrete-depth hydrogeologic, chemical, and age data for determining active circulation depth, and illustrates an approach for maximizing the utility of individual boreholes drilled for mountain bedrock aquifer characterization.

1. Introduction

Delineating the depth of active groundwater circulation, or the "bottom of the watershed," is essential for developing realistic conceptual hydrologic models of mountain watershed systems (Condon et al., 2020). The active circulation depth directly controls aquifer storage, groundwater residence times, and the disposition of groundwater to discharge locally in the headwater catchment versus exiting in the subsurface, becoming interbasin flow to a lower elevation parent watershed (Gleeson & Manning, 2008; Markovich et al., 2019; Schaller & Fan, 2009). The definition of active versus inactive groundwater can vary depending on study objectives. Here, we consider the depth of active circulation to be the depth below which the total annual groundwater flux is a sufficiently small fraction of the annual watershed hydrologic budget that its exclusion from watershed numerical models has a negligible impact on model outputs.

A growing number of watershed studies in mountainous regions demonstrate that: (a) groundwater years to decades old (or older) can be an important control on runoff generation and stream solute concentrations (e.g., Frisbee et al., 2011; Hale & McDonnell, 2016; Jasechko et al., 2016; Rumsey et al., 2017; Visser et al., 2019); (b) interbasin flow from headwater catchments to lower elevation streams and aquifers may be common (e.g., Ameli et al., 2018; Genereux et al., 2005; Welch et al., 2012); and (c) active circulation depths of 100–200 m are to be expected in crystalline rocks (e.g., Manning & Caine, 2007; Markovich et al., 2019; Welch & Allen, 2014). Despite these findings, data collection efforts and numerical hydrologic model representations for mountain watersheds are still largely limited to depths <10 m (Condon et al., 2020). A central objective of studying mountain watershed systems is to develop numerical models that can reliably predict future changes in runoff, stream water chemistry, and biogeochemical processes resulting from changing climatic and land use conditions, as these could critically impact downstream water resources (e.g., Hubbard et al., 2018). Although it is advantageous to minimize the depth of such models to maximize

computational efficiency, models that fail to include deeper groundwater (assume deep flow rates are negligible) without sound field-based justification run a considerable risk of generating inaccurate forecasts of future stream flows and solute loads.

Despite the importance of delineating the active groundwater circulation depth in mountain watersheds, relatively few studies have done so either directly or implicitly (Condon et al., 2020; Markovich et al., 2019). Studies that provide some constraint on the depth of active circulation are reviewed by Manning and Caine (2007), Welch and Allen (2014), and Markovich et al. (2019), and can be broadly categorized into three types: (1) studies using hydraulic, chemical, and environmental tracer (mainly groundwater age) data from tunnels and mines (e.g., Marechal and Etchevery, 2003; Mayo et al., 2003; Tomanaga et al., 2017); (2) studies using geothermometry and temperature data from open boreholes and springs, along with coupled heat and fluid flow modeling (e.g., Frisbee et al., 2017; Manning & Caine, 2007; Saar & Manga, 2004; Sanz & Yelamos, 1998); and (3) studies using surface geophysical surveys, in some cases coupled with borehole geophysical and fracture data (e.g., Flinchum et al., 2018, 2019; McClymont et al., 2012; Robinson et al., 1974; St. Clair et al., 2015). These previous studies have contributed valuable insights into possible active circulation depths in different geologic settings, but the reliability of reported or inferred active circulation depths are limited by several factors. Tunnels and open boreholes in areas of mountain topography can disturb natural groundwater flow regimes by acting as drains and vertical conduits, artificially inducing the flow of shallow young groundwater to depths far greater than in the natural system. Calculating circulation depth from geothermometers and spring discharge temperatures requires significant assumptions regarding subsurface thermodynamic conditions and local geothermal gradients, introducing considerable uncertainty. Geophysical methods can indicate the colluvial/bedrock contact and the transition from weathered highly fractured bedrock to unweathered bedrock with fewer fractures, but the hydrogeologic significance of these transitions in terms of groundwater flux and the degree to which they may correspond to the top of the inactive zone remains speculative. Finally, the extent of "modern" groundwater (recharged after the early 1950s) is often equated with the extent of active groundwater flow, yet the age of active versus inactive groundwater can vary substantially based on aquifer configuration, porosity, and recharge distribution.

A clear need exists to directly observe and more rigorously constrain the active circulation depth in mountain watersheds by measuring hydraulic conductivity (K) and groundwater-flux-dependent variables, such as age and temperature (T), at discrete levels in the subsurface sufficiently deep to penetrate the inactive zone. An increasing number of hillslope nested wells are being installed at Critical Zone Observatories and other research watersheds (e.g., Salve et al., 2012; Tokunaga et al., 2019; White et al., 2019), and these have provided important information on the flow at different depths within the active system and the relative importance of different subsurface water stores in streamflow generation. However, to date these lack sufficient depth and/or adequate groundwater-flux-dependent data collection and interpretation to permit direct, robust delineation of the active circulation depth. Prior studies of mountain groundwater flow have also largely focused on volcanic and crystalline rock terrains, meaning even less is known about circulation depths in watersheds underlain by sedimentary rocks. In this study we present discrete-depth K, hydraulic head (H), T, chemical, noble gas, and groundwater age data from two boreholes 46 and 81 m deep in Redwell Basin, Colorado, an alpine watershed underlain by subhorizontally bedded sedimentary rocks. These borehole data are combined with distributed chemical and environmental tracer data from shallow groundwater and surface water, along with a schematic 2-D hillslope coupled heat and fluid flow model, to characterize the fractured bedrock aquifer and determine the depth of active groundwater circulation. Our study approach maximizes the value of individual deep boreholes for achieving these fundamental groundwater characterization objectives, a necessity given that few such boreholes typically can be installed due to their expense and terrain access challenges.

1.1. Site Description

Redwell Basin is a 2.1-km^2 headwater catchment in the Elk Mountains near Crested Butte, Colorado, located within the Upper Colorado River Basin (Figure 1). Elevations range from 2,860 to 3,770 masl and land cover is predominantly barren alpine and conifer forests, with smaller zones of meadow, shrubs, and aspen. Average annual precipitation based on meteorological data from nearby locations at similar elevation is ~1,300 mm/yr, with ~70% falling as snow (Carroll et al., 2019; U.S. Department of Agriculture, Natural

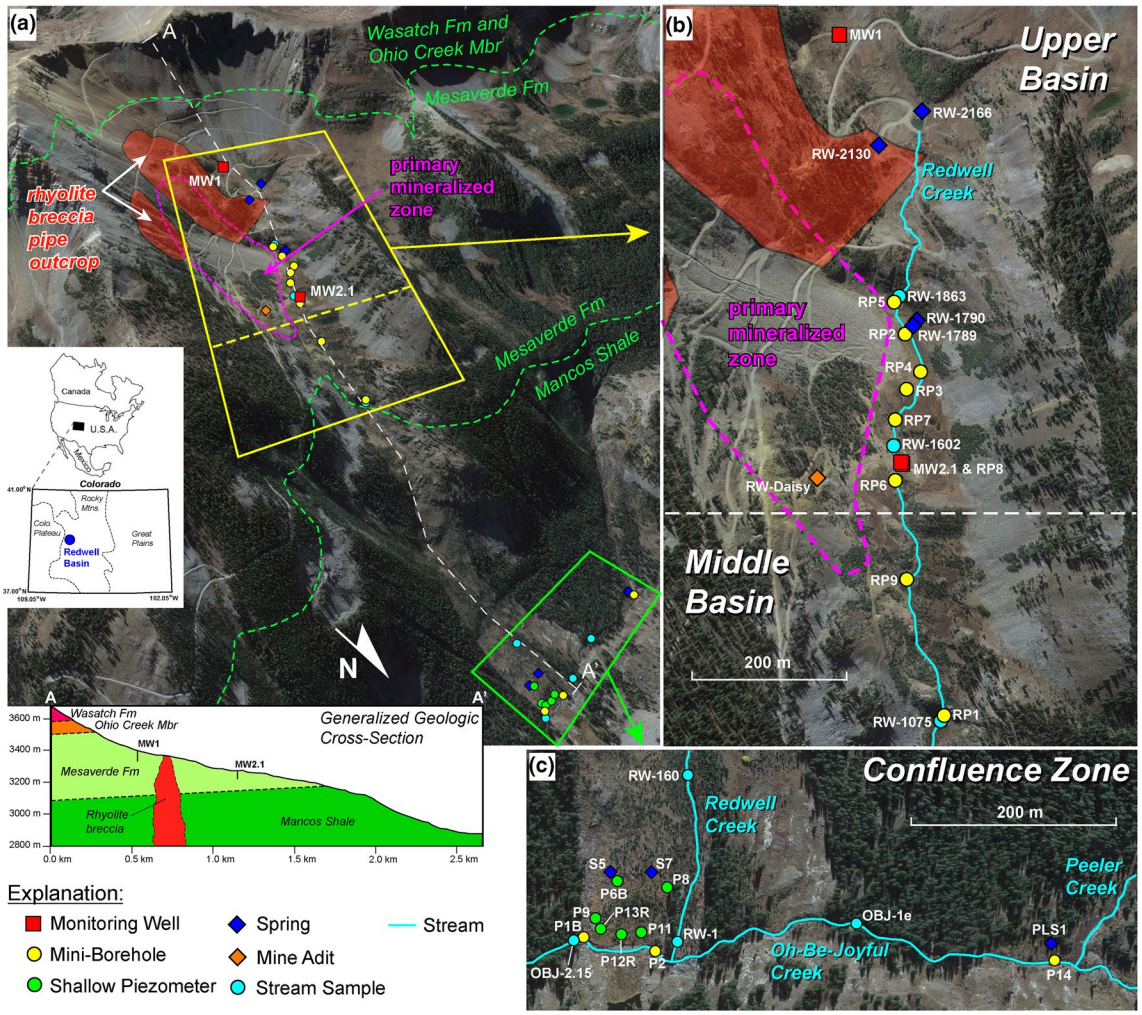

Figure 1. (a) Oblique aerial image of Redwell Basin looking southwest showing general location of sample sites in upper basin, middle basin, and confluence zones, as well as important geologic features and a generalized geologic cross-section adapted from Gaskill et al. (1967). Primary mineralized zone is based on Charnock et al. (2019). Expanded plan-view aerial images of upper and middle basin zones (b) and confluence zone (c) show more detailed sample site locations. Ohio Creek Mbr = Ohio Creek Member (of Mesaverde Formation). Aerial images from Google Earth.

Resources Conservation Service, https://wcc.sc.egov.usda.gov/nwcc/site?sitenum=737). Redwell Creek is a high-gradient (0.26) first-order perennial stream running along the valley axis. The stream is un-gaged, but similar streams in the area are snowmelt dominated with peak flows occurring in June and receding through the summer and fall (Carroll et al., 2019). A stream tracer dilution study performed in late August 2001 (near average water year) by Kimball et al. (2010) reported a baseflow discharge rate of 48 L/s at the bottom of Redwell Creek immediately above its confluence with Oh-Be-Joyful (OBJ) Creek.

The basin is underlain by a thick sequence of Cretaceous to Tertiary sedimentary rocks that dip <15° to the south-southwest. These include the marine Mancos Shale in the lower basin and dominantly marginal-marine and fluviolacustrine mixed sandstone and shale of the Mesaverde Formation in the middle and upper basin (Figure 1; Gaskill et al., 1967). The headwall of the upper basin is composed predominantly of sandstones of the Ohio Creek Member of the Mesaverde Formation and lower Wasatch Formation. Although shales are in general expected to have low permeability, shales in Redwell Basin have been sufficiently altered, tectonized, and weathered as a result of local magmatism and topographic uplift to potentially increase their fracture permeability substantially. The sedimentary sequence is intruded by 18-16 Ma granodiorite and quartz monzonite porphyry dikes and sills, as well as a rhyolitic (felsite) breccia pipe exposed in the upper basin (Figure 1). Several map-scale faults with trace lengths of 0.25–2 km are present, these

being generally steeply dipping, northeast- and northwest-striking, often vein-filled, and of relatively small displacement (<100 m; Caine et al., 2010; Charnock et al., 2019; Gaskill et al., 1967). Smaller faults, fracture networks, and veins occur more pervasively and have more variable orientation.

Local magmatism and associated hydrothermal fluids produced polymetallic vein mineralization (mainly pyrite, sphalerite, and galena) and strong phyllic alteration (quartz, sericite, and pyrite) near the breccia pipe in the upper basin. Recent detailed structural and hydrothermal alteration mapping by Charnock et al. (2019) indicates that vein mineralization is primarily located within a zone extending from the breccia pipe's northeast side ("primary mineralized zone" shown in Figure 1). Hydrothermal fluids also produced less intense vein-hosted and disseminated propylitic alteration (chlorite, epidote, and minor pyrite) throughout much of the rest of the basin, as well as widespread silicification that substantially reduced primary rock porosity. The vein systems were mined historically for Ag, Cu, Pb, and Zn at the Daisy Mine in Redwell Basin (Figure 1) and other mines in nearby watersheds, but these mines are now inactive (Berger et al., 2001). Weathering of the abundant sulfides in the upper basin in the vicinity of the breccia pipe naturally generates groundwater with low pH and elevated metal concentrations (acid-rock drainage, or ARD). Discharge of the ARD produces extensive deposits of Fe-oxide cemented colluvium, or ferricrete, near Redwell Creek between sites RP5 and RP6 (Figure 1), and, along with drainage from the Daisy Mine, causes stream metal concentrations to exceed aquatic-life standards (Kimball et al., 2010).

2. Methods

2.1. Installation of Monitoring Wells, Mini-Boreholes, and Shallow Piezometers

Depths and completion information for all groundwater sample points are provided in Table S1 (supplementary information), and details regarding drilling and installation of monitoring wells beyond the summary below are provided in Ball et al. (2020). Boreholes for monitoring wells MW1 and MW2.1 were drilled to a total depth of 81.3 and 45.9 m below ground surface (BGS), respectively, using an HQ (9.6 cm hole diameter) wireline coring system with nearly complete core recovery (Figure 1). Drilling of the MW2.1 borehole had to be discontinued when artesian flow of ~20 L/min was encountered. The borehole for the shallow bedrock well RP8 was drilled approximately 3 m from MW2.1 using the same methods and is not artesian. Nested 1.9-cm-diameter monitoring wells with screen lengths of 3–9 m were completed at four different depths in the MW1 borehole (MW1A-D) and three different depths in the MW2.1 borehole (MW2.1A-C; Table S1). Screen depths were selected to target relatively high-yield zones based on hydraulic packer test results, fracture abundance observed in drill core, and borehole geophysical data (Section 2.2). The one exception is MW2.1A, which has no screen and is open through the entire 13-m zone of primary artesian flow production, isolated below a dedicated packer to allow completion of overlying nested wells. Mini-boreholes (MBs) 41 mm in diameter and generally 1–2 m deep were drilled adjacent to Redwell and OBJ Creeks using a portable, handheld, rock coring drill manufactured by Shaw Tool Company (Figure 1). Mini-boreholes were drilled either in ferricrete or in bedrock intersecting weeping steeply dipping fractures apparently hosting groundwater discharge, and were left open. Shallow piezometers (henceforth piezometers) with screens 20–30 cm long were pounded by hand to depths of 0.5–2.5 m in colluvium in a broad zone of groundwater discharge (springs and small wetlands) at the base of Redwell Basin in the confluence zone (Figure 1).

2.2. Hydrogeologic, Chemical, and Environmental Tracer Data Collection

The extensive set of hydrogeologic and geophysical data collected from boreholes at the site and related method descriptions are provided in Ball et al. (2020). Long-interval (10 m) constant head step injection tests were performed during drilling of the two deep boreholes to determine in situ K using a single-packer apparatus and standard methods (Bliss & Rushton, 1984; Quinn et al., 2012). Water level and T data were monitored from September 2018 to September 2019 in all wells using dedicated pressure transducers. Transducers in MW2.1A-C could not be lowered to the screen depth due to strong artesian flow and had to be placed at the top of the casing, hence T data from the screen depth were not collected. Therefore, only the bottom-hole T from borehole geophysical logging (shallower measurements were influenced by strong upflow) and transducer T data from RP8 are presented for this location. Hydraulic head elevation (H) was

computed assuming a single estimated mean annual barometric pressure for each well location, and periods immediately following deployment or sampling when the water level was clearly disturbed were excluded.

Most water samples were collected in September and October 2017, when the majority of MBs and piezometers were installed and the most complete sample set was collected from the site. Exceptions include: (1) samples from RP9, P14, PLS1, and RW-1789, collected in September 2018 because they were either not yet installed or not sampled for noble gases in 2017; and (2) samples from the wells, collected in September 2019 allowing at least one full year between well installation and sampling to minimize potential contamination by residual drilling fluid (stream water). Samples were collected from wells MW1A-D and RP8 using a submersible double-valve pump manufactured by Solinst Canada. Wells MW2.1A-C required no pumping given their natural artesian pressure. Samples were collected from other sites using a peristaltic pump. Complete method descriptions for measurement of field parameters and collection, treatment, preservation, and analysis of samples for major ion and trace element concentrations are presented in Johnson et al. (2019). Details regarding collection and analysis of environmental tracer samples are provided in Manning et al. (2020). Tritium analyses were performed by either electrolytic enrichment and liquid scintillation counting or by ^{3}He ingrowth. Noble gas samples were collected in clamped copper tubes as described by Stute and Schlosser (2000), maintaining back-pressure within the sample line to prevent bubble formation and gas loss by exsolution. Samples were analyzed for He, Ne, Ar, Kr, and Xe concentration and for ^{3}He/^{4}He ratio by cryogenic gas separation and mass spectroscopy methods as detailed in Hunt (2015).

2.3. Recharge Parameter and Groundwater Age Calculations

Measured Ne, Ar, Kr, and Xe concentrations were used to compute recharge temperature and excess air parameters, assuming the widely applied closed-system equilibrium (CE) model of excess air formation (Aeschbach-Hertig et al., 2000; Jung et al., 2013). The noble gas recharge temperature (NGT) is the temperature at the water table at the recharge location, and excess air is the component of atmospheric gas dissolved from air bubbles trapped below the water table when it rises during recharge events (Kipfer et al., 2002). Specifically, A_e is the initial concentration of trapped air and F describes the degree of gas fractionation during subsequent dissolution, which depends on the fraction of trapped air that dissolves (Aeschbach-Hertig et al., 2000; Kipfer et al., 2002). Recharge parameters were derived using a computer code that employs standard inverse techniques to minimize the error-weighted misfit (χ^2) between measured and modeled gas concentrations, similar to that described by Aeschbach-Hertig et al. (1999) and Jung and Aeschbach (2018). Excess air levels are indicated by ΔNe, the excess-air component of Ne expressed as a percentage of the solubility component. Recharge elevation (elevation of the water table at the recharge location) was assumed to be the approximate mean land surface elevation of the watershed area above the sample point, except for samples in the confluence zone where a local elevation was assumed given the very young computed ages (see Section 3.3). Nitrogen was substituted for Xe in the recharge parameter derivation for all well samples except MW1D because Xe was found to be substantially depleted (8%–34%) relative to other gases (see supplementary information). Additionally, the prescribed oxygen saturation at the water table at the recharge location was reduced from 100% for samples from MW1A, MW1B, and RP1 to achieve an acceptably low χ^2 sum ($p > 0.10$).

Groundwater ^{3}H/^{3}He ages were computed from measured ^{3}H concentrations and tritiogenic ^{3}He (^{3}He*) concentrations calculated from the recharge parameters (Schlosser et al., 1988, 1989) assuming both piston-flow (PF) and exponential model (EM) age distributions (Cook & Bohlke, 2000). The PF age assumes the age of sampled water is uniform and is most appropriate for samples from short-screened wells outside of discharge zones. The EM mean age assumes sampled water is of mixed age and is most appropriate for samples of groundwater discharging from a simple aquifer having uniform porosity, thickness, and recharge rate. The EM age distribution has been widely assumed for stream baseflow in catchment travel time studies (e.g., McGuire & McDonnell, 2006). A local precipitation ^{3}H record was assumed based on Michel et al. (2018) and Jurgens (2018). The concentration of terrigenic He (He_{terr}) was also computed, a component of He produced in the subsurface by radioactive decay of U and Th in aquifer solids that can provide additional groundwater age information (Solomon, 2000). The He_{terr} concentration is expressed as a percentage of the solubility He component using the term ΔHe_{terr}. Uncertainties for all computed parameters and ages were determined using the Monte Carlo method with 3,000 realizations.

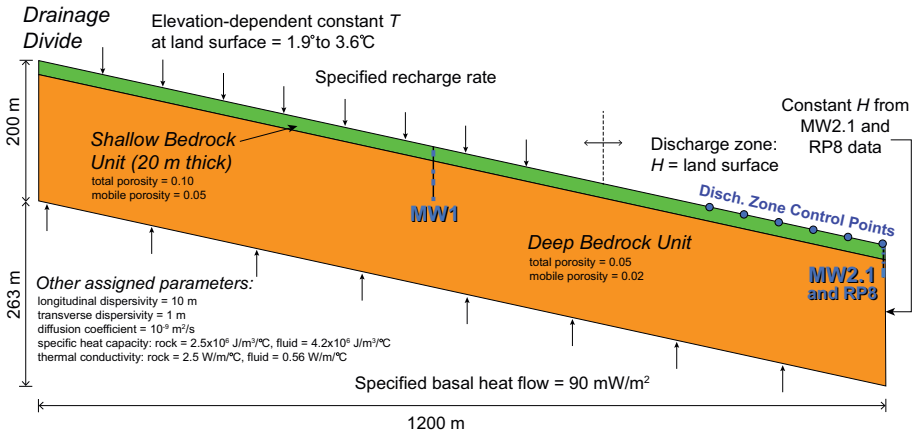

Figure 2. Diagram showing main components of 2-D numerical coupled groundwater flow and heat transport model. Blue segments in wells indicate screened zones.

2.4. Numerical Groundwater Flow and Heat Transport Modeling

A simple 2-D coupled groundwater flow and heat transport model was constructed to quantify rates of groundwater flow at different depths generally consistent with observed ages and subsurface temperatures. Steady-state variably saturated flow was simulated within a hillslope-parallel transect extending from MW2.1 upslope through MW1 to the drainage divide, along A-A' in Figure 1, using the finite-element code FEFLOW (Diersch, 2014, Figure 2). Unsaturated zone flow was simulated to enable estimation of the water table elevation, using Richard's equation and assuming simple linear relationships between capillary pressure, saturation, and *K*. Note that accurate representation of these relationships and resulting saturation profiles within the vadose zone are unimportant for this modeling exercise, having no effect on total groundwater flux or the water table position. Mean groundwater age was computed using the age-mass approach of Goode (1996), employing FEFLOW's mass transport package and including both advection and dispersion.

The box model is 1,200 m long by 200 m deep, and the top boundary is a linear approximation of the ground surface with a slope equal to that between MW1 and MW2.1 (Figure 2). Only flow within the bedrock is simulated, and the bedrock is represented as an equivalent porous medium. The model space is divided into a shallow bedrock unit 0–20 m deep and a deep bedrock unit 20–200 m deep, with mesh node spacing of 1–3 m and 5–6 m, respectively. Thermal boundary conditions include a basal heat flow of 90 mW/m^2 (Blackwell et al., 2011) and elevation-dependent constant *T* along the top equal to the mean annual ground *T* (T_g) estimated by interpolation of *T* data from MW1 and MW2.1 (see Section 3.1). Hydraulic boundary conditions include: (1) constant recharge along the top (varied in different model runs), which switches to a constant *H* condition equal to the land surface where the water table intersects the land surface to allow discharge in the lower part of the slope; and (2) depth-dependent constant *H* along the right-side vertical boundary at MW2.1 based on a curve fit to *H* data from MW2.1 and RP8 (Figure S1, supplementary information), which allows water to exit the model laterally.

Total and mobile porosities used to compute *T* and age distributions (Figure 2) are based on the water content measured from the nuclear magnetic resonance log from the MW2.1 borehole, as well as laboratory He porosity measurements on 13 drill core samples from representative lithologies in the two deep boreholes (Ball et al., 2020). Longitudinal and transverse dispersivities are typical values for flow systems with length scales of hundreds of meters (Dong et al., 2018; Gelhar et al., 1992). Specific heat capacities and thermal conductivities for fluid and solids are reasonable literature-based values for shales and sandstones (Clauser & Huenges, 1995; Waples & Waples, 2004). Shallow and deep bedrock *K* are isotropic unless otherwise noted.

The modeling process involved comparison of simulated *H*, *T*, and mean age under different model scenarios with observed values at well screen locations, focusing on MW1 in the recharge zone. In addition,

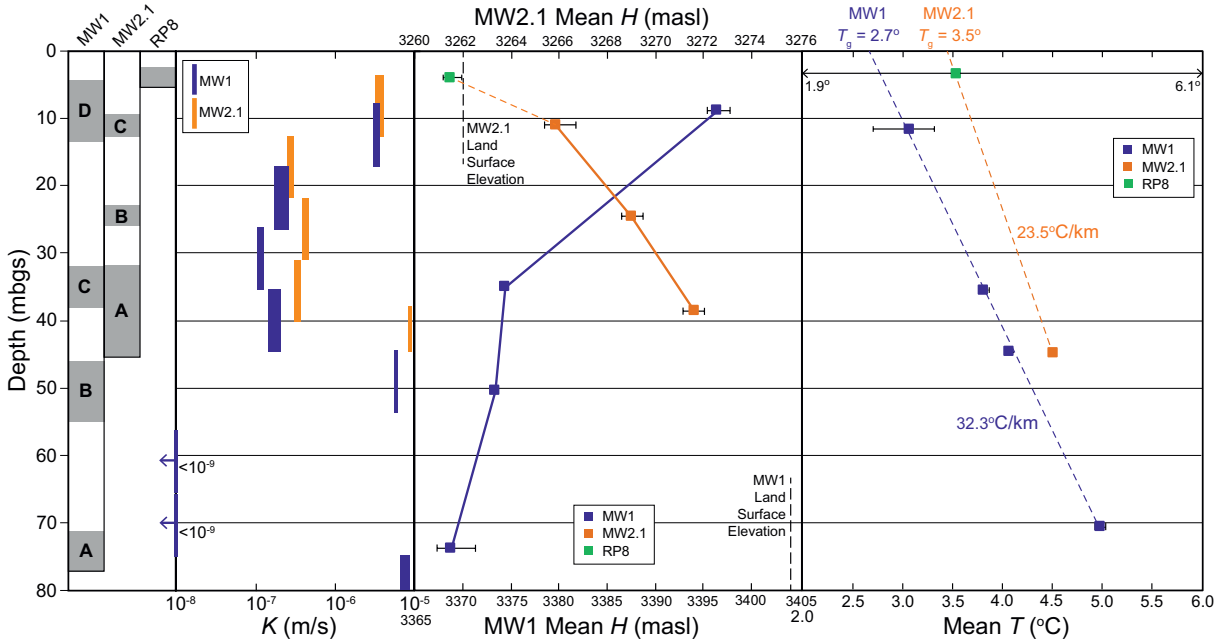

Figure 3. Well-screen intervals (shaded areas), hydraulic conductivities (K) measured in aquifer tests, observed mean hydraulic heads (H), and observed mean temperatures (T) in MW1, MW2.1, and RP8. For bars indicating measured K, length is packer test interval and width is uncertainty range. For mean H and mean T, error bars indicate range of annual variation. T_g = mean annual surface ground temperature, masl, meters above mean seal level; mbgs, meters below ground surface.

modeled median age at six evenly spaced control points located along the surface 0–250 m upslope from MW2.1 (Figure 2) were compared to observed ages in upper basin MBs, located within this same zone. Instead of comparing ages at specific MB locations, the overall range and median of modeled and observed ages within the upper basin groundwater discharge zone were compared as a general consistency test, an approach we believe more appropriate given the highly schematic nature of the model. An initial model scenario (Model 1A) was developed with a reasonable expected recharge rate of 400 mm/yr for alpine watersheds in the western U.S. (e.g., Carroll et al., 2019; Flint et al., 2013; Knowles et al., 2015), and the same K for shallow and deep bedrock units determined by matching the observed water table elevation in MW1D. Successive scenarios were then tested in which recharge, shallow K, and deep K were progressively adjusted to reduce discrepancies between observed and modeled H, T, and mean age.

3. Results and Discussion

3.1. Hydraulic and Temperature Data

Hydraulic conductivity values typically associated with the active zone (10^{-7} to 10^{-5} m/s; Markovich et al., 2019) predominate in both boreholes (Figure 3; Table S2; see Ball et al., 2020, for all packer test, H, and T data). Generally, both holes display a shallow high-K zone (on the order of 10^{-6} m/s) at roughly 0–15 m depth overlying an intermediate-K zone (on the order of 10^{-7} m/s) at roughly 15–40 m, in turn overlying a deeper zone of highly variable K ($<10^{-9}$ to 10^{-5} m/s) at roughly 40–80 m. The vertical H gradient is negative in MW1 and positive in MW2.1/RP8, as expected for a recharge zone and discharge zone, respectively (Figures 3, S2, and S3, Table S3 supplementary information). However, H gradient absolute values are unexpectedly high (average of 0.4), suggesting a generally low subvertical K component, particularly between MW1D and C where the H gradient is −0.8. The T gradient in MW1 is nearly linear, and gradients in both wells of 23.5°C and 32.3°C/km are typical for conductive continental geotherms (20°C–30°C/km; Lowell et al., 2014, Figures 3, S2, and S3, Table S3, supplementary information). Surface projections of observed T gradients indicate T_g values of 2.7° for MW1 and 3.5° for MW2.1, and a lapse rate of 0.6°C per 100 m of elevation gain similar to other mountain settings in the western U.S. (e.g., Powell et al., 1988).

Measured K values suggest that active groundwater flow may extend to a depth 80 m or more. However, high vertical H gradients suggest that subvertical K may be substantially less than aquifer test K values, possibly indicating a strongly anisotropic K enhanced in the subhorizontal (bedding-parallel) direction. The dramatic difference in H between MW1D and underlying wells MW1A-C (22–28 m) suggests limited hydraulic connection between groundwater at depths >30 m and shallow groundwater near the water table, and this shallow aquifer may be seasonally or perennially perched. Furthermore, observed T profiles are inconsistent with typical recharge and discharge zone thermal profiles, which tend to be nonlinear and noticeably steeper (recharge) or shallower (discharge) than a conductive geotherm due to vertical groundwater flow and resulting advective heat transport. A linear conductive geotherm like that in MW1 likely cannot be sustained in mountain settings unless vertical Darcy groundwater flow velocity (q) is roughly 1 cm/yr or less (Forster & Smith, 1989; Manning & Caine, 2007), a flow rate more consistent with the inactive zone given that recharge rates are typically tens of centimeters per year. Together, observed K, H, and T data therefore suggest relatively high subhorizontal K but relatively low subvertical K, and appear inconsistent with active groundwater circulation below the zone separating MW1D and C, or a depth of ~20 m.

3.2. Water Chemistry

Three distinct end-member types of sampled water are apparent from field parameter, major ion, and trace element data (Figure 4; Table S4; see Johnson et al., 2019 for all water chemistry data). Type 1 samples are dilute (SC < 200 μS/cm) with circumneutral pH (5.5–8.0), low SO_4 and Ca (<30 mg/L), and low Zn (<1,000 μg/L). Note that for samples in this data set, Zn concentrations are generally representative of other metals primarily derived from sulfide weathering, such as Fe. Type 1 samples are mainly from the confluence zone, including the OBJ Creek samples and most of the shallow groundwater samples. Type 1 samples also include the highest elevation Redwell Creek sample RW-1863, located upstream of the primary mineralized zone, and two upper basin spring samples RW-2166 and RW-1790 with upslope apparent catchment areas on the west side of the watershed (opposite the breccia pipe and primary mineralized zone). Type 2 samples are typical ARD, with elevated SC (>300 μS/cm), low pH (3–4), high SO_4 (>100 mg/L), low Ca (<20 mg/L), and high Zn (>3,000 μg/L). These include all upper basin shallow groundwater samples located within or immediately downslope of the primary mineralized zone. Type 3 samples are also evolved with elevated SC (>400 μS/cm) and high SO_4 (>120 mg/L), but with circumneutral pH (6.5–8.0), high Ca (>60 mg/L), and low Zn (<1,000 μg/L). Type 3 samples include all of the well samples, except MW1D. Most of the remaining samples plot between Types 1 and 2 and include: (1) all Redwell Creek samples, except the highest elevation one; (2) upper basin shallow groundwater samples RP5 and RW-2130 and well sample MW1D, all located outside the primary mineralized zone, but with apparent catchment areas near the breccia pipe; and (3) a few of the confluence zone shallow groundwater samples. The two middle basin shallow groundwater samples also plot outside of end-member zones. Sample RP1 on the west side of the creek plots between Types 1 and 3, and RP9 on the east side plots alternately as a Type 2 sample or between Types 2 and 3.

The water chemistry data suggest that upper basin groundwater deeper than MW1D (~20 m, henceforth "deep groundwater") is well evolved and has a chemistry (Type 3) clearly distinct from that of shallower waters in the basin. All sampled stream water and shallow groundwater are Types 1, 2, or intermediate between these two types. The chemistry of the intermediate samples may be explained by either: (a) mixing of Type 1 and 2 waters; (b) an intermediate degree of influence of sulfide weathering near the breccia pipe; or (c) an unsampled end-member water with low pH and SC, such as soil water. Importantly, the stream and shallow groundwater samples show no indication of significant mixing with or evolution toward deep Type 3 groundwater. Sample RP1 in the middle basin is a notable exception (perhaps RP9 as well), suggesting some deep groundwater discharge in this location. However, stream sample RW-1075 co-located with RP1 shows no influence of Type 3 water, suggesting that this discharge is volumetrically minor compared to stream flow. Sample RP8 is another exception, and is discussed below in Section 3.3 with the age results. The feasibility of shallow Type 1–2 water, such as in MW1D, evolving while flowing to depth to become Type 3 water, such as in MW1A and B, was evaluated using forward geochemical modeling with PHREEQC (Parkhurst & Appelo, 1999). The modeling indicated that this is indeed possible if the Type 1–2 water reacts with carbonate minerals. Carbonate minerals were observed in the MW1 core, but only in minor amounts at discrete locations, calling into question the likelihood of this chemical evolution path being

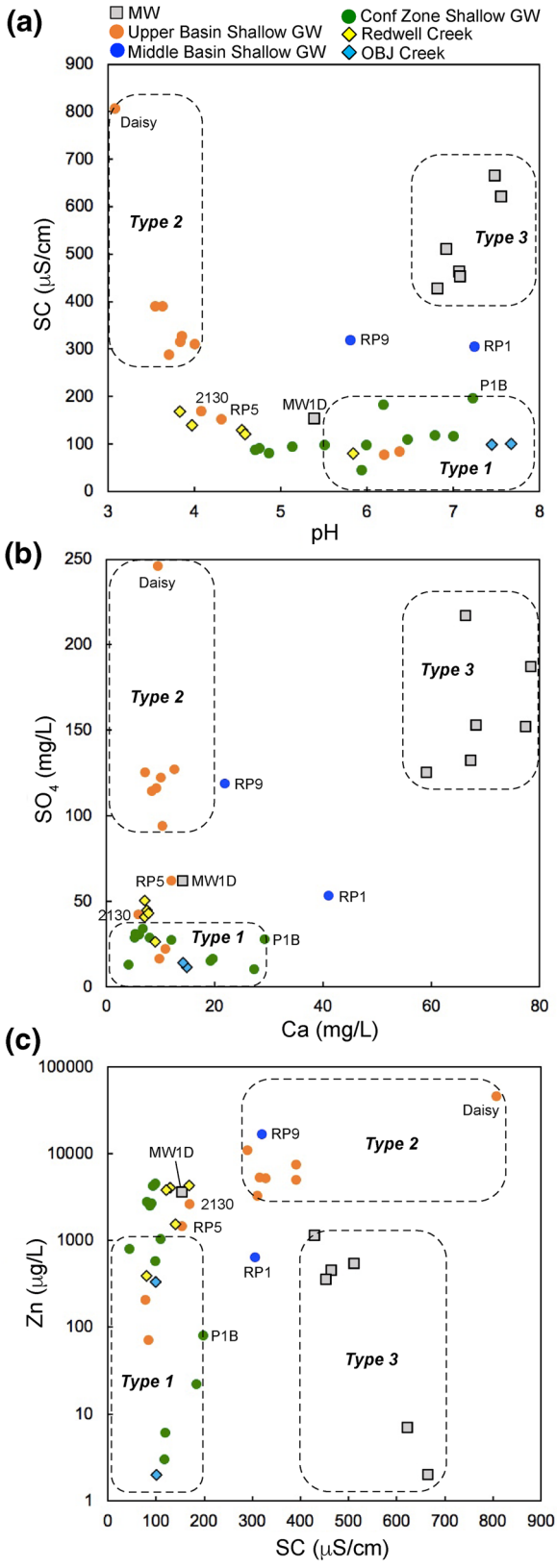

Figure 4. Plots of (a) specific conductance (SC) versus pH, (b) SO_4 versus Ca, and (c) Zn versus SC. GW, groundwater; MW, monitoring well; OBJ Creek, Oh-Be-Joyful Creek. Dashed lines are interpretive delineations of zones including chemically distinct end-member water types.

common throughout the watershed. Although the distinctly different chemistry of shallow waters and deep groundwater cannot rule out substantial downward flow to depth in recharge zones, it does argue against substantial upward flow from deep to shallow levels and to streams in discharge zones. The chemistry data are therefore consistent with deep groundwater being hydraulically relatively isolated from the stream and shallow groundwater, and limited active groundwater circulation below a depth of ~20 m.

Five of the confluence zone shallow groundwater samples are intermediate between Types 1 and 2, with a pH of 4.70–5.50 and Zn concentration of 2,490–4,460 μg/L. The fact that the chemistry of these samples trends toward lower Redwell Creek water (pH of ~4.6, Zn concentration of ~4,000 μg/L), combined with the fact that the three spring/piezometer samples located closest to the creek (S7, P8, and P11; Figure 1) are part of this group, suggest that the lower basin shallow groundwater system contains water recharged by stream loss from Redwell Creek. However, another possible explanation is that the lower basin shallow groundwater contains some upper basin shallow groundwater, which would imply long stream-parallel groundwater flowpaths from the upper basin to the lower basin. Although the location of these samples is more consistent with the former explanation, the latter cannot be ruled out based on the chemistry data alone.

3.3. Recharge Parameters and Groundwater Age

Noble gas recharge temperatures are generally near or below T_g for the sample site elevation, as observed at other snowmelt-dominated mountain sites (e.g., Manning & Solomon, 2003, Figure 5a; see Manning et al., 2020, for all environmental tracer data). Exceptions include four shallow groundwater samples in the confluence zone (P14, S5, S7, and PLS1) with *NGT*s >1°C above the local T_g of 5.8°C, suggesting that these samples contain months-old water recharged during the prior late spring/summer when shallow ground temperatures were above T_g (Masbruch et al., 2012). Another exception is well sample MW1A, with an *NGT* of 4.7°C compared to a local T_g of 2.7°C. Given the considerable depth of this sample, the warm *NGT* in this case likely reflects the water table in the recharge zone being deep enough to have a temperature above T_g due to the geothermal gradient (Heilweil et al., 2012), roughly 70 m based on the observed MW1 *T* profile (Figure 3). Excess air concentration (ΔNe) is in general negatively correlated with *NGT*, and different site types display distinctive ΔNe and *NGT* characteristics. Upper basin shallow groundwater samples have relatively high ΔNe and low *NGT* (means of 43% and 1.3°C), confluence zone shallow groundwater samples have relatively low ΔNe and high *NGT* (means of 5% and 5.1°C), and middle basin shallow groundwater samples have intermediate ΔNe and *NGT* values. Most deep groundwater samples from the wells have ΔNe values of 100%–120%, substantially higher than the shallow samples. Observed trends in ΔNe and *NGT* thus suggest that water table oscillations are in general positively correlated with recharge elevation, consistent with the finding of Manning and Caine (2007) that the largest excess air concentrations occurred in groundwater that apparently recharged high in the watershed where the largest water tables oscillations were observed. The recharge parameters also suggest that the confluence zone shallow groundwater contains nearly all local recharge, with little or no high-ΔNe deep groundwater from the upper basin, consistent with the water chemistry data.

Sample initial ^{3}H concentrations (measured ^{3}H + computed ^{3}He*) were plotted versus recharge year as calculated from the ^{3}H/^{3}He piston-flow (PF) age, and compared to the precipitation ^{3}H record to determine if samples contain a substantial fraction of water recharged prior to the early 1950s ("premodern" water; Figure 5b). Modern PF samples should plot near the precipitation ^{3}H line (PF line) on Figure 5b. A dashed line indicating where EM samples would plot (EM line) is also shown. All shallow groundwater samples and well sample MW1D plot above the PF line and relatively close to the EM line (for samples <1 year old, EM and PF lines overlap). Better agreement with the EM age distribution is consistent with most of the shallow groundwater samples being in discharge locations relatively close to the creek. Well samples RP8, MW2.1C, and MW1A plot well below both lines, suggesting that they are "mixed-modern" and contain large fractions of premodern water. The other well samples are not plotted because they have ^{3}H concentrations <0.5 TU indicating that they are premodern.

The two computed independent age indicators, ^{3}H/^{3}He EM mean age (henceforth "EM age") and $\Delta He_{terr,}$ are compared in Figure 5c, and their spatial distributions are shown in Figure 6. Different sample types have notably different age characteristics. The upper basin shallow groundwater samples have consistent EM

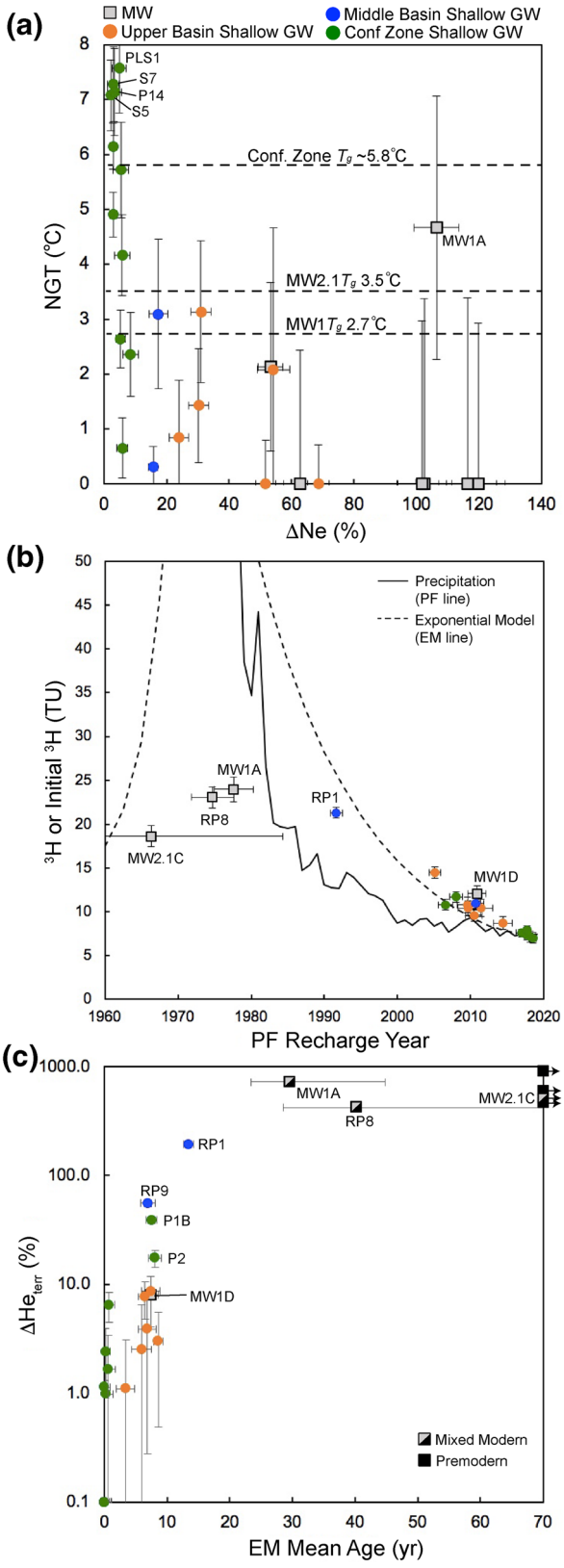

Figure 5. Plots of (a) noble gas recharge temperature (NGT) versus ΔNe, (b) ^{3}H or initial ^{3}H versus piston-flow (PF) ^{3}H/^{3}He recharge year, and (c) ΔHe_{terr} verus ^{3}H/^{3}He exponential model (EM) mean age. GW, groundwater; MW, monitoring well; T_g, mean annual ground surface temperature.

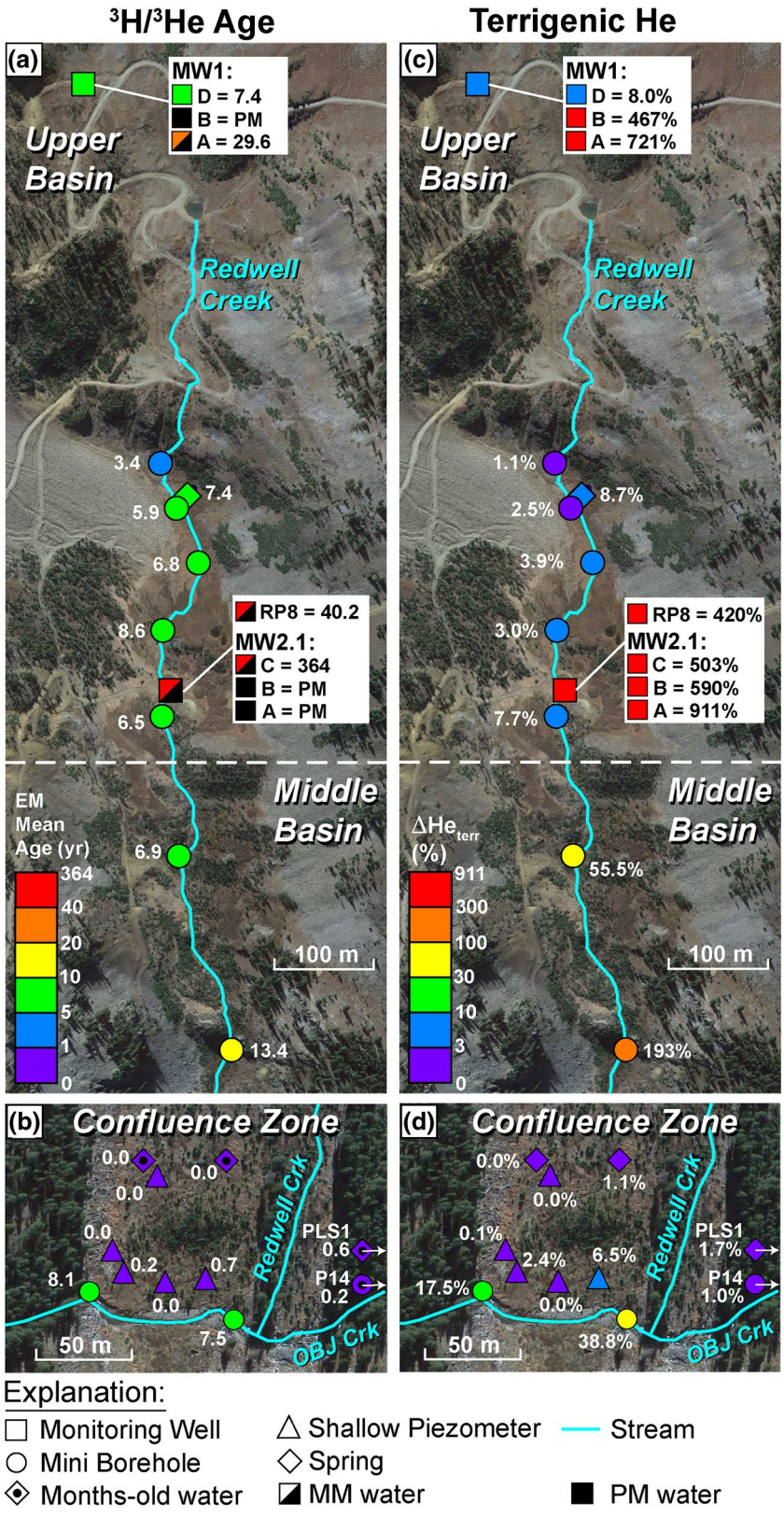

Figure 6. Spatial distribution of ^{3}H/^{3}He exponential model (EM) mean age in (a) upper and middle basin zones, and (b) confluence zone, along with spatial distribution of terrigenic He (ΔHe_{terr}) in (c) upper and middle basins, and (d) confluence zone. MM, mixed-modern; PM, premodern; OBJ, Oh-Be-Joyful.

ages of 6–9 years and ΔHe_{terr} of 3%–9%, the one exception being the slightly younger most-upstream sample RP5 at 3 years and 1%. This consistency in age of discharging groundwater suggests a fairly uniform active circulation depth in the upper basin (Haitjema, 1995). In the middle basin, the more upstream shallow groundwater sample RP9 has an EM age in the same range as the upper basin shallow groundwater samples (6.9 years), suggesting a continuation of the same aquifer thickness into the middle basin. However, the more downstream sample RP1 is clearly older with an EM age of 13.4 years and ΔHe_{terr} of 193%. The greater age of RP1 is consistent with its distinct chemistry, intermediate between Type 1 and Type 3 waters (Figure 4). Possible reasons for older groundwater discharge at RP1 include: (1) it is located on a fault or discrete fracture zone that may enhance upward flow from depth; and (2) it is also located immediately above the contact between the Mesaverde Formation and underlying Mancos Shale (Figure 1), the latter perhaps acting as a barrier to deep stream-parallel groundwater flow due its lack of sandstone beds. The increase in ΔHe_{terr} from upper to middle basin zones is considerably more pronounced than the increase in EM age, being 10–40 times higher (compared to 0–2 times) and apparent in both middle basin samples (not just RP1). This suggests that discharge of old high-He_{terr} groundwater increases downstream throughout the middle zone, but remains volumetrically small above the location of RP1such that EM age is unimpacted.

Shallow groundwater samples in the confluence zone are mostly younger than those in the upper and middle basin zones, having EM ages of <1 year and ΔHe_{terr} of 0%–7% (mostly <3%). Samples from stream-side bedrock MBs P1B and P2 at the base of Redwell Basin are two important exceptions, having EM ages of 8.1 and 7.5 and ΔHe_{terr} of 17.5% and 38.8%, respectively (Figure 6). These values are in the same range as upper basin shallow groundwater samples and RP9 in the middle basin, suggesting a similar fractured bedrock aquifer system exists in the confluence zone. The other shallow groundwater samples in lower Redwell Basin are from the colluvium, and their very young age (the most uphill samples containing months-old water based on their *NGT*s) suggests they primarily contain discharge from a locally recharged, highly transmissive colluvial aquifer rather than bedrock discharge. These young ages are therefore inconsistent with the possibility raised by the chemistry data that the colluvial aquifer contains some shallow groundwater from the upper basin. Significant stream-parallel groundwater flow from upper to lower portions of Redwell Basin should produce groundwater discharge in the confluence zone older than that in the middle and upper basin zones, given the required longer flowpath lengths. The age data thus provide further evidence, in addition to the chemistry and noble gas data, of an absence of such flow occurring along either shallow or deep pathways. Similar chemistry, noble gas, and age signatures of samples at the base of Peeler Basin immediately west of Redwell (PLS1 and P14) suggest that such flow may be minimal in the adjacent basin as well.

Except for MW1D, well samples are considerably older than the shallow groundwater samples, being either premodern or mixed-modern with EM ages of 30 to >1,000 years and ΔHe_{terr} of 420%–911%. The EM ages provide some indication of relative age for these samples, but, because the actual distribution of ages within the samples is unknown (PF and EM models are unlikely based on Figure 5b), the true mean ages of these older samples are uncertain. However, as with premodern samples (mean age >65 years), the ^3H and $^3He^*$ concentrations can be used to estimate a minimum mean age for mixed-modern samples. Assuming the sample is a bimodal mixture of modern water (having an age indicated by $^3H/^3He^*$ ratio) and premodern water, the relative vertical distance of its plotted location between the x-axis and the PF or EM line on Figure 5b roughly approximates the fraction of modern water in the sample (e.g., Manning et al., 2005). Using this fraction and assuming a minimum age of 65 years for the premodern fraction yields a minimum mean age of >56 years for the mixed-modern samples. Assuming the premodern component has an only slightly older minimum age of 79 years yields minimum sample mean ages of >65 years.

Terrigenic He concentrations provide additional insight into the approximate mean age of premodern and mixed-modern well samples. A groundwater age was estimated from the He_{terr} concentration using the TracerLPM calculator of Jergens et al. (2012), which assumes He is released to pore fluid at the same rate it is produced by radioactive decay of U and Th in the aquifer solids (Andrews & Lee, 1979). This calculation requires several assumptions and is thus highly approximate, but nonetheless provides a useful order-of-magnitude estimate. Uranium and Th concentrations of 3 and 10 ppm, respectively, were assumed based on unpublished whole-rock chemistry data from MW1 core samples, along with a porosity of 0.02. Contribution from upward crustal He flux was assumed to be zero, which is reasonable for a sampling

depth of <100 m, particularly for MW1 where groundwater flow is downward (Stute et al., 1992). Computed ages are 1,700–3,600 years, consistent with the fact that the He_{terr} levels observed in the premodern and mixed-modern samples are typically found in groundwater >1,000 years old (Solomon, 2000). Together, the ^3H and He isotopic data therefore strongly suggest that the mean age of both mixed-modern and premodern samples is >65 years, this being roughly an order of magnitude greater than the mean age of the shallow groundwater samples.

The EM age and ΔHe_{terr} of well sample MW1D are similar to other upper basin shallow groundwater samples. However, ages of samples from the same shallow interval <20 m deep at MW2.1 (MW2.1C and RP8) are mixed-modern and clearly older, and these samples have Type 3 water chemistry. This could indicate local discharge of deep groundwater, but the lack of a discernible deep groundwater influence on the age or chemistry of nearby MB sample RP6 and the adjacent stream segment suggests that a shallower local transition from active to inactive groundwater at a depth of ~5 m is more likely. Well samples from depths >20 m are all either premodern (3 of 4) or mixed-modern, and have the highest ΔHe_{terr} values in the data set, suggestive of inactive groundwater at this depth. The mixed-modern age of MW1A is interesting, given that the overlying shallower sample MW1B is premodern and thus apparently older. This implies that some amount of modern water reaches depths >20 m in the upper basin, probably via discrete steeply dipping high-*K* fractures/faults. A modern component in MW1A is consistent with MW1A having a larger seasonal *H* variation than in the other deep wells (Table S3, Figures S2 and S3, supplementary information), indicating a greater hydraulic connection to the shallow system. However, as discussed above, the ^3H and He isotope data suggest that a minority of the water in sample MW1A is modern, and the other three deep samples contain essentially no modern water, so the total volume of modern water in the deep system appears to be relatively small. In summary, age data for the well samples suggest that groundwater below a depth of ~20 m in the upper basin is dominantly premodern with mean ages generally an order of magnitude (or more) greater than shallow groundwater, and is thus likely inactive.

3.4. Numerical Modeling Results

The *H*, *T*, chemistry, noble gas, and age results together all support a hypothesis of an active circulation depth of ~20 m in Redwell Basin. A simple 2-D hillslope numerical model was constructed to more rigorously and quantitatively test this hypothesis. The model was used to determine the groundwater flow velocities at different depths generally compatible with the *H*, *T*, and age data, with a focus on determining the maximum allowable *q* at depths >20 m in MW1. As discussed above, conductive geotherms and premodern ages have been used as likely indictors of inactive groundwater, but this association has not been quantitatively demonstrated for a mountain watershed, and these characteristics provide no guarantee that the deep groundwater flux is a sufficiently small fraction of recharge to be considered inactive. The purpose of our model is therefore to aid in the development of a well-supported watershed conceptual model upon which a more comprehensive numerical model eventually could be based. One shortcoming of the 2-D model is that it does not include the component of groundwater flow perpendicular to Redwell Creek in the lower part of the section (discharge zone shown in Figure 2). However, the focus of the modeling is on reproducing observations in MW1 in the recharge zone. Furthermore, if the flow system is relatively uniform throughout the upper basin, as suggested by the consistency of near-stream shallow groundwater ages, the mean age of discharging stream-perpendicular flow should be approximately the same as the stream-parallel flow within the 2-D model section (Haitjema, 1995). In comparing measured and modeled values of *T*, mean age, and *H*, modeled values were considered acceptable if they met the following criteria (Table 1): *T* for MW1A-D within 0.2° of the measured value (twice the measurement uncertainty); age for MW1D of 5–10 years (EM age ± twice the 1-σ uncertainty); age for MW1A and B of >65 years; median age for discharge zone control points of 3–9 years (range observed in upper basin MBs); *H* for MW1A-D within 1 m of measured value (reasonable target given manual model calibration and mesh node spacing).

The objective of the first modeling step (Models 1A and 1B, Table 1) was to establish an acceptable bedrock recharge rate. Shallow and deep units were assumed to have the same *K*, the value of which was determined by matching the water table elevation in MW1D. In Model 1A, a reasonable expected recharge rate of 400 mm/yr was applied as a starting point. Resulting simulated *T* and age at all depths in MW1 are clearly too low, and *H* in MW1A-C is too high (Figure 7). Simulated ages for MW2.1 are also too young (18–20 years

Water Resources Research

10.1029/2020WR028548

Table 1
Summary of Model Characteristics and Results

Model	Objective	Calibration points	Recharge rate (mm/yr)	Shallow bedrock K (m/s)	Deep bedrock K (m/s)	Deep bedrock flow fraction (% of total)	Modeled versus Observed H, T, and age			
							Acceptable T in MW1A,B,C?	Acceptable Age in MW1A,B?	Acceptable Disch Zone Median Age?	Acceptable H in MW1A,B,C?
1A	Simulate T and age for typical recharge rate and homogeneous K	H at MW1D	400	2.9E-07	2.9E-07	NA	No	No	Yes	No
1B	Determine more likely recharge rate	T, age, and H at MW1D	80	5.9E-08	5.9E-08	NA	No	No	No	No
2A	Reduce deep BR K to better match T and age in MW1A,B,C	SAA	80	3.1E-07	2.9E-08	47	Yes	No	No	No
2B	SAA	SAA	80	4.8E-07	1.0E-08	16	Yes	A yes, B no	No	No
2C	SAA	SAA	80	5.2E-07	5.0E-09	8	Yes	Yes	No	No
3A	Include distributed anisotropy to better match H in MW1A,B,C	SAA, and H in MW1A	80	5.0E-07	surface-parallel K = 1.1E-8 aniotropy ratio = 0.043	12	Yes	Yes	Yes	A yes, B/C no
3B	Include two aquitards to better match H in MW1A,B,C	SAA, and H in MW1A,B,C	80	5.1E-07	aquifers = 1.0E-8 upper quitard = 1.2E-10 loweraquitard = 5.0E-10	11	Yes	Yes	Yes	Yes

anisotropy ratio, ratio of surface-perpendicular K to surface-parallel K; BR, bedrock; Disch, discharge; H, hydraulic head; K, hydraulic conductivity; SAA, same as above; T, temperature.

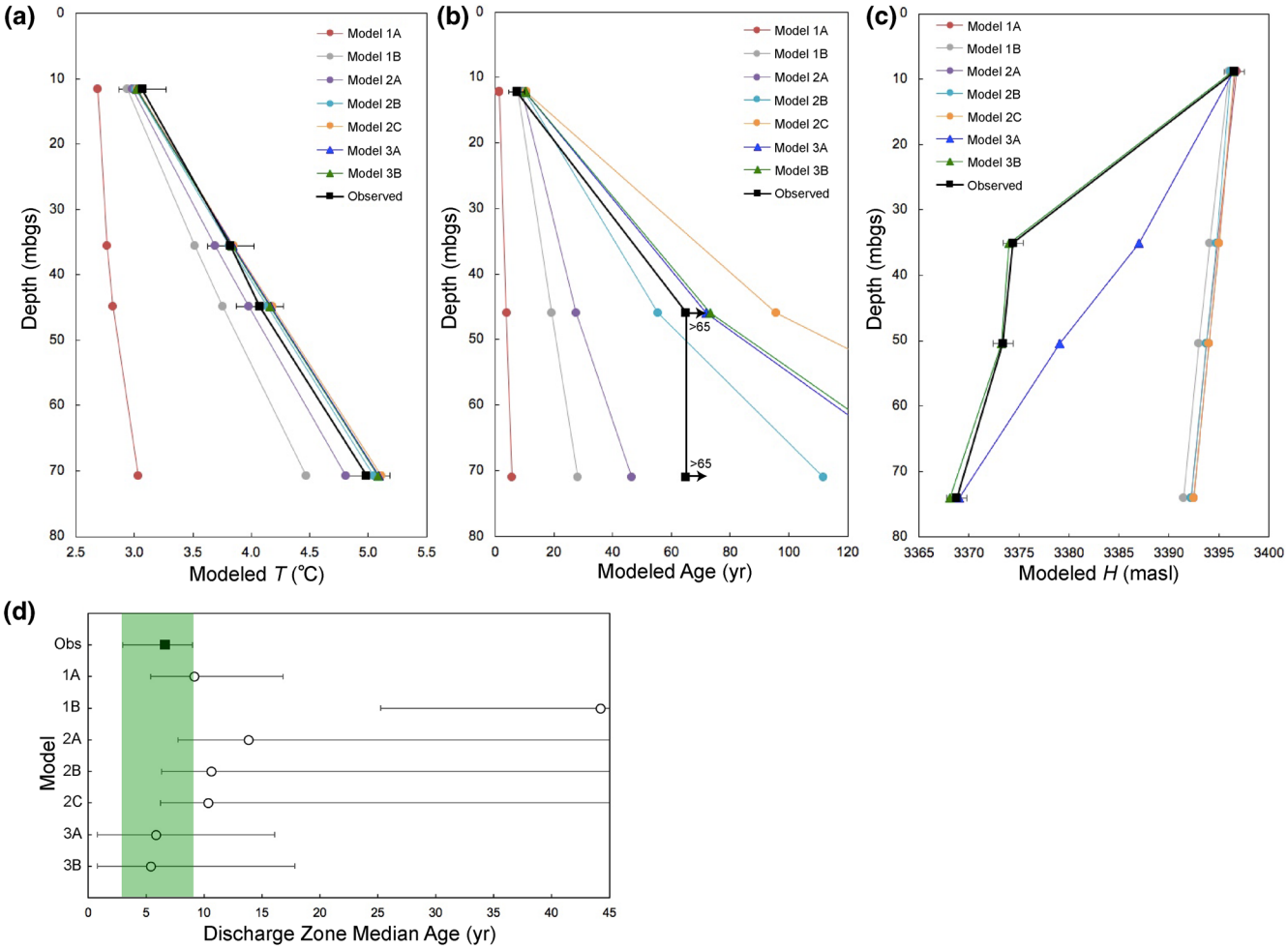

Figure 7. Plots of modeled (a) temperature (*T*), (b) mean age, and (c) hydraulic head (*H*) versus depth in monitoring well MW1, with observed values also shown for comparison. Error bars for observed values indicate acceptable range for modeled values. Modeled discharge zone median age is shown in (d), along with the median of observed mean ages for upper basin discharge zone shallow groundwater samples. Error bars indicate range of modeled and observed discharge zone ages, with the observed range further highlighted by the green zone. masl, meters above mean sea level; mbgs, meters below ground surface; Obs, observed.

compared to >65 years) and bottom-hole *T* is too warm (5.1° compared to 4.5°). The recharge rate then was reduced until simulated age and *T* matched observed values in MW1D, this being in the shallow unit in the recharge zone where, unlike in the deep unit and discharge zone, age and *T* are relatively insensitive to *K* at depth. The resulting lower recharge rate of 80 mm/yr still yields unacceptably low *T*, low age, and high *H* in MW1A-C, as well as unacceptably old ages in the discharge zone, indicating excessive flow in the deep unit. MW2.1 ages become acceptable (90–97 years) but bottom-hole *T* remains too warm (6.4°). Simulated values for MW2.1 are not discussed further for the following two reasons. First, ages are acceptable in all following model runs (177–1,260 years) and *T* is unacceptably high (5.6°–6.4°), meaning they do not contribute to discerning acceptable versus unacceptable deep *q* values. Second, in some models the specified *H* at the downhill model boundary (Figure S1) had to be changed to no-flow at some depths to prevent flow into the model. This adjustment limits the ability of water to exit laterally and likely increases upflow at the model boundary, potentially producing artificially high ages and *T* at MW2.1.

The second modeling step (Models 2A–2C) was performed to determine the maximum groundwater flow in the deep unit compatible with observed *T* and age in MW1A-C at a recharge rate of 80 mm/yr. In each model, the deep unit *K* was reduced by a factor of 2 or 3 from the prior model and the shallow unit *K* was increased as required to maintain acceptable *H* in MW1D (*T* and age remained acceptable). Simulated *T* becomes acceptable in Model 2A, in which flow through the deep unit is 47% of total flow. Simulated age

approaches acceptability in Model 2B, with a deep unit flow fraction of 16%, but does not become acceptable in both MW1A and B until Model 2C, with a deep unit flow fraction of 8%. The discrepancy between simulated and observed discharge zone median age improves considerably in the step 2 models, becoming nearly acceptable in Model 2B and C at 10–11 years. Simulated *H* remains unacceptably high in all step 2 models.

In the third modeling step (Models 3A and 3B), *K* anisotropy orthogonal to the ground surface was introduced in an effort to reduce the discrepancy between measured and modeled *H*. In Model 3A, distributed anisotropy was included in the deep unit, applying the same recharge rate of 80 mm/yr. The ground-parallel *K* component was maximized and the anisotropy ratio (ratio of ground-perpendicular to ground-parallel components) was adjusted in an effort to match *H* in MW1A-C, while maintaining acceptable *T* and keeping age in MW1B close to the 65 years minimum (MW1A age always exceeded MW1B). The shallow unit *K* was simultaneously adjusted to maintain acceptable *T*, age, and *H* in MW1D. An acceptable *H* was simulated for MW1A, but *H* for MW1B and C remained unacceptably high due to the approximately linear *H* gradient. In model 3B, discrete rather than distributed anisotropy was included by adding two 10-m thick ground-parallel aquitards, an upper one between MW1D and C, and a lower one between MW1B and A (Figure S4). The upper aquitard was also positioned closer to the surface near MW2.1 in an effort to better match observed ages in shallow groundwater in this location; in all prior models, simulated ages in RP8 and the overlying surface control point equivalent to RP6 were similar, unlike the disparate observed mean ages of >65 years (mixed-modern) for RP8 and 7 years for RP6. As with Model 3A, the deep aquifer *K* was maximized and *K* for the aquitards and shallow unit adjusted while keeping age in MW1B close to the 65 years minimum and acceptable *T*, age, and *H* for all MW1 wells. Resulting *T*, age, and *H* at all levels in MW1 are all acceptable. Interestingly, the water table becomes perched in a ~150 m zone immediately upslope of MW1. In addition to more closely matching observed *H*, Models 3A and B produce acceptable discharge zone median ages, unlike the step 2 models, and simulated mean ages in Model 3B for RP8 and the overlying control point (>65 years and 18 years, respectively) more closely approximate observed values. Deep unit flow fractions remain low for both step 3 models at 11% and 12%.

Overall, the modeling results suggest that at most ~10% of bedrock recharge (<1 cm/yr) flows through the deep bedrock unit in the upper basin, further supporting the hypothesis of an active circulation depth of ~20 m. The improved match between observed and modeled *H* and discharge zone ages in the step 3 models provides additional evidence that the bedrock *K* is in general highly anisotropic, with subvertical *K* substantially smaller than subhorizontal *K* (ratio of 0.043 in Model 3A). Model 3B suggests that this anisotropy is due at least in part to discrete low-*K* layers acting as aquitards within the sedimentary-rock sequence, and it may cause local perched groundwater conditions.

A circulation depth of ~20 m is not unprecedented (Flinchum et al., 2018; St. Clair et al., 2015), but was somewhat unexpected given that it is relatively shallow compared to depths of 100–200 m typically reported for other sites (Manning & Caine, 2007; Markovich et al., 2019; Welch & Allen, 2014). The shallower circulation in Redwell Basin is likely caused by the *K* anisotropy in the underlying sedimentary rock sequence being greater than in the crystalline rocks predominantly underlying other study sites, resulting in a relatively low overall subvertical *K* at depths >20 m (10^{-10} to 10^{-9} m/s in step 3 models). However, a circulation depth of ~20 m is also shallow compared to the two prior estimates of which we are aware from mountain sites underlain by gently dipping clastic sedimentary rocks: ~150 m for an area of the Wasatch Plateau in Utah (Mayo et al., 2003) and ~180 m for a site only ~10 km east of Redwell Basin in the Upper East River watershed (Tokunaga et al., 2019). The former area is composed of mixed shale and sandstone, and the latter is composed entirely of shale. We can see two possible reasons why our estimated circulation depth may be shallower than estimates from these other two studies. First, the high degree of hydrothermal alteration in Redwell Basin, absent at the other sites, may have significantly reduced K_v through widespread silicification and/or mineral precipitation within subvertical fractures. Second, the circulation depth at the other two sites may be overestimated. Mayo et al. (2003) rely mainly on groundwater age and hydraulic observations from mine tunnels, which, as noted in the Introduction, can significantly overestimate circulation depth. Tokunaga et al. (2019) rely on *K* measurements from depths <10 m obtained using a method that, as acknowledged by the authors, likely underestimates *K* in the shallow weathered bedrock zone, thereby producing an overestimation of circulation depth. Regardless, our results raise the possibility that circulation depths commonly may be less than 100–200 m in mountain watersheds underlain by subhorizontal clastic sedimentary rock sequences.

The simulated bedrock recharge rate of 80 mm/yr is lower than reported estimates for similar mountain watersheds, most notably rates of 300–600 mm/yr estimated by Carroll et al. (2019) and Tokunaga et al. (2019) for two different sites in the nearby Upper East River watershed. Our lower bedrock recharge rate is probably due to the shallow bedrock unit in our model not including an uppermost highly weathered zone, which commonly have K values up to 10^{-5} m/s (Tokunaga et al., 2019; Welch & Allen, 2014). Inclusion of such a shallow high-K zone in our model was not justified given our lack of discrete data at depths <5 m outside of areas of groundwater discharge, but if included would presumably allow considerably higher bedrock recharge rates (though flow at depths >5 m would remain the same). There is no reason to suspect an unusually low total recharge rate (to the soil zone) for the basin, and in fact the baseflow discharge for Redwell Creek estimated by Kimball et al. (2010) requires a total recharge rate of 720 mm/yr. This total recharge rate implies that baseflow is supported mainly by interflow from the colluvium and/or discharge from a very shallow highly weathered bedrock zone, neither of which were included in our model. Importantly, a total recharge rate of this magnitude also implies that deep groundwater accounts for <2% of baseflow and thus likely has a minimal influence on the quantity and quality of stream water.

Finally, the modeling suggests that the discrete-depth T, age, and H data from MW1 provide a relatively robust level of constraint on bedrock recharge, K, and q at different depths near this well. The age data provide the greatest constraint on the relative fraction of deep versus shallow groundwater flux. However, the T data still provide useful independent support for deep groundwater flow likely being <10% of total recharge. The T data probably would be more valuable in wells closer to the drainage divide having a larger proportion of vertical flow. Obviously, more wells would be desirable to assess spatial variability of the circulation depth, but the modeling results illustrate that the general well completion and data collection approach taken in this study is highly effective for constraining circulation depth in borehole locations, maximizing the value of the inevitably small number of deep boreholes that feasibly can be drilled in a mountain watershed.

4. Summary and Conclusions

The depth of active groundwater circulation directly controls a watershed's capacity to store and transmit solutes and water, and its proper characterization is essential for accurately predicting surface water quality and quantity under varying climate and land use conditions. Obtaining a robust field-based estimate of active circulation depth is a fundamental challenge of mountain hydrology, and this study demonstrates a tractable approach for addressing this challenge. We present a highly unique data set including hydraulic conductivity (K), hydraulic head, temperature, chemistry, noble gas, and $^3H/^3He$ groundwater age data from discrete-depth bedrock wells up to 81 m deep in Redwell Basin, an alpine watershed underlain by subhorizontal clastic sedimentary rocks. These data are combined with chemical and age data from shallow groundwater discharge, along with a simple 2-D numerical hillslope model, to constrain the depth of active groundwater circulation.

Steep vertical hydraulic head gradients (average of 0.4) and linear thermal profiles consistent with typical conductive continental geotherms in wells MW1 and MW2.1 suggest relatively low subvertical K and sufficiently low vertical groundwater flow rates below a depth of ~20 m to be potentially considered inactive. Groundwater below ~20 m ("deep groundwater") in MW1 and MW2.1 is chemically distinct from sampled creek water and shallow groundwater throughout the basin. Tritium and He isotope data indicate that the deep groundwater samples are dominantly premodern with sample mean ages likely >65 years. In contrast, shallow groundwater samples throughout the basin, most being near-stream groundwater discharge, predominantly have mean ages <9 years. Mean ages of near-stream shallow groundwater samples in the upper and middle basin zones are generally 6–9 years, and this consistency suggests a relatively uniform flow system and active circulation depth. Shallow groundwater samples in the confluence zone have noble gas recharge temperature and excess air signatures distinct from the upper basin samples (particularly deep groundwater), and have mean ages younger than or similar to shallow groundwater in the upper and middle basin zones. These chemistry and age patterns suggest that: (a) deep groundwater flow rates are relatively slow, and deep groundwater contributions to the surface water system are minimal; and (b) stream-parallel groundwater flow from upper to lower Redwell Basin following either deep or shallow pathways is negligible. Together, the above findings support a hypothesis of a relatively shallow active groundwater circulation depth of ~20 m.

A simple 2-D numerical fluid flow and heat transport model was constructed representing a hillslope transect in the upper basin, including MW1 and MW2.1, to more rigorously test the above hypothesis and further aid in conceptual model development. The modeling suggests that at most ~10% of bedrock recharge reaches depths >20 m. The simulated deep bedrock flow rate combined with a previous baseflow discharge estimate by Kimball et al. (2010) suggests that deep groundwater accounts for <2% of baseflow, a small enough fraction to likely have little influence on stream water quality or quantity and thus be considered inactive. The modeling also suggests strong K anisotropy in the sedimentary rock sequence underlying the basin, probably attributable to the occurrence of discrete low-K layers and producing an overall low subvertical K at depths >20 m. Modeling results thus provided further support for an active circulation depth of ~20 m in Redwell Basin. This circulation depth is relatively shallow compared to previously reported values but could be typical for mountain watersheds underlain by subhorizontally bedded sedimentary sequences containing shale units (even if highly tectonized) for which circulation depth has rarely been examined.

This study demonstrates the considerable value of discrete-depth K, hydraulic head, temperature, chemistry, and age data collected from a limited number of bedrock boreholes deep enough to penetrate the inactive zone for mountain watershed studies aiming to constrain active circulation depth. More broadly, our finding of a circulation depth of only ~20 m underscores the large potential variability of this critical watershed parameter, its strong dependence on bedrock geology, and the clear need for more direct observations at depth in mountain watersheds.

Data Availability Statement

All data are accessible online in electronic format in cited U.S. Geological Survey data releases. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Acknowledgments
This study was funded by U.S. Department of Energy grant DE-SC0016250 and the U.S. Geological Survey's Mineral Resources Program. Drilling costs were supported by the Lawrence Berkeley National Laboratory's Watershed Function Scientific Focus Area, with funding from the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research under contract DE-AC02-05CH11231 (Lawrence Berkeley National Laboratory; operated by the University of California). We are thankful for constructive reviews provided by Jonathan Caine and two anonymous reviewers that helped improve the manuscript.

References

Aeschbach-Hertig, W., Peeters, F., Beyerle, U., & Kipfer, R. (1999). Interpretation of dissolved atmospheric noble gases in natural waters. *Water Resources Research*, *35*, 2779–2792. https://doi.org/10.1029/1999WR900130

Aeschbach-Hertig, W., Peeters, F., Beyerle, U., & Kipfer, R. (2000). Paleotemperature reconstruction from noble gases in ground water taking into account equilibration with entrapped air. *Nature*, *405*, 1040–1043. https://doi.org/10.1038/35016542

Ameli, A. A., Gabrielli, C., Morgenstern, U., & McDonnell, J. J. (2018). Groundwater subsidy from headwaters to their parent water watershed: A combined field-modeling approach. *Water Resources Research*, *54*, 5110–5125. https://doi.org/10.1029/2017WR022356

Andrews, J. N., & Lee, D. J. (1979). Inert gases in groundwater from the Bunter Sandstone of England as indicators of age and paleoclimatic trends. *Journal of Hydrology*, *41*, 233–252. https://doi.org/10.1016/0022-1694(79)90064-7

Ball, L. B., Manning, A. H., Carr, B. J., Williams, K. H., Burton, B. L., & Martinez, L. (2020). *Hydrologic and geophysical data from high-elevation boreholes in Redwell Basin near Crested Butte*. CO. U.S. Geological Survey Data Release. https://doi.org/10.5066/P900URV6

Berger, B. R., Wanty, R. B., & Tuttle, M. L. (2001). Scale versus detail in water–rock investigations 2: Field-scale models of fracture networks in mineral deposits. In R. Cidu (Ed.), *Proceedings of the 10th International symposium on water-rock Interaction* (pp. 137–140). New York, NY: A. A. Balkema.

Blackwell, D., Richards, M., Frone, Z., Batir, J., Ruzo, A., Dingwall, R., & Williams, M., (2011). *Temperature at depth maps for the conterminous US and geothermal resource estimates, Geothermal Resources Council Transactions, v. 35 (GRC1029452)*. Available at: https://www.smu.edu/geothermal

Bliss, J. C., & Rushton, K. R. (1984). The reliability of packer tests for estimating the hydraulic conductivity of aquifers. *Quarterly Journal of Engineering Geology and Hydrogeology*, *17*, 81–91. https://doi.org/10.1144/GSL.QJEG.1984.017.01.10

Caine, J. S., Manning, A. H., Berger, B. R., Kremer, Y., Guzman, M. A., Eberl, D. D., & Schuller, K. (2010). *Characterization of geologic structures and host rock properties relevant to the hydrogeology of the Standard Mine in Elk Basin*. Gunnison County, CO: U.S Geological Survey Open-File Report 2010–1008 56 p. https://pubs.usgs.gov/of/2010/1008/

Carroll, R. W. H., Deems, J. S., Niswonger, R., Schumer, R., & Williams, K. H. (2019). The importance of interflow to groundwater recharge in a snowmelt-dominated headwater basin. *Geophysical Research Letters*, *46*, 5899–5908. https://doi.org/10.1029/2019GL082447

Charnock, R., Scherschel, T., Barron, M., Mauk, J. L., Kuiper, Y. D., & Monecke, T. (2019). *Structural and alteration framework of a base metal mineralized quartz vein system that overlies a Climax-type porphyry Mo deposit at Crested Butte, Colorado, USA*. Poster presented at 2019 Geological Society of America Annual Meeting, Sept. 22–25, Phoenix, Arizona, USA. https://gsa.confex.com/gsa/2019AM/meetingapp.cgi/Paper/339592

Clauser, C., & Huenges, E. (1995). Thermal conductivity of rocks and minerals. In T. J. Ahrens (Ed.), Vol. 3 (pp. 105–126). *Rock physics & phase relations: A handbook of physical constants*. Washington, DC: American Geophysical Union. https://doi.org/10.1029/RF003p0105

Condon, L. E., Markovich, K. H., Kelleher, C. A., McDonnell, J. J., Ferguson, G., & McIntosh, J. C. (2020). Where is the bottom of a watershed?. *Water Resources Research*, *56*. e2019WR026010. https://doi.org/10.1029/2019WR026010

Cook, P. G., & Bohlke, J. K. (2000). Determining timescales for groundwater flow and solute transport. In P. G. Cook, & A. L. Herczeg (Eds.), *Environmental tracers in subsurface hydrology* (pp. 1–30). Dordrecht, Holland: Kluwer Academic Publishers.

Diersch, H. J. G. (2014). *FEFLOW: Finite element modeling of flow, mass and heat transport in porous and fractured media*. NY: Springer Publishing Co.

Dong, S., Dai, Z., Li, J., & Zhou, W. (2018). The scale dependence of dispersivity in multi-facies heterogeneous formations. *Carbonates and Evaporites*, *33*, 161–165. https://doi.org/10.1007/s13146-018-0421-6

Flinchum, B. A., Holbrook, W. S., Parsekian, A. D., & Carr, B. J. (2019). Characterizing the critical zone using borehole and surface nuclear magnetic resonance. *Vadose Zone Journal*, *18*(1), 180209. https://doi.org/10.2136/vzj2018.12.0209

Flinchum, B. A., Holbrook, W. S., Rempe, D., Moon, S., Riebe, C. S., Carr, B. J., Hayes, J. L., et al. (2018). Critical zone structure under a granite ridge inferred from drilling and three-dimensional seismic refraction data. *Journal of Geophysical Research: Earth Surface*, *123*, 1317–1343. https://doi.org/10.1029/2017JF004280

Flint, L. E., Flint, A. L., Thorne, J. H., & Boynton, R. (2013). Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. *Ecological Processes*, *2*, 25. https://doi.org/10.1186/2192-1709-2-25

Forster, C., & Smith, L. (1989). The influence of groundwater flow on thermal regimes in mountainous terrain: A model study. *Journal of Geophysical Research*, *94*, 9439–9451. https://doi.org/10.1029/JB094iB07p09439

Frisbee, M. D., Phillips, F. M., Campbell, A. R., Liu, F., & Sanchez, S. A. (2011). Streamflow generation in a large, alpine watershed in the southern Rocky Mountains of Colorado: Is streamflow generation simply the aggregation of hillslope runoff responses?. *Water Resources Research*, *47*, W06512. https://doi.org/10.1029/2010WR009391

Frisbee, M. D., Tolley, D. G., & Wilson, J. L. (2017). Field estimates of groundwater circulation depths in two mountainous watersheds in the western U.S. and the effect of deep circulation on solute concentrations in streamflow. *Water Resources Research*, *53*, 2693–2715. https://doi.org/10.1002/2016WR019553

Gaskill, D. L., Godwin, L. H., & Mutschler, F. E. (1967). *Geologic map of the Oh-Be-Joyful quadrangle*, Gunnison County, CO: Geological Survey Geologic Quadrangle Map GQ–578. https://doi.org/10.3133/gq578

Gelhar, L. W., Welty, C., & Rehfeldt, K. R. (1992). A critical review of data on field-scale dispersion in aquifers. *Water Resources Research*, *28*, 1955–1974. https://doi.org/10.1029/92WR00607

Genereux, D. P., Jordan, M. T., & Carbonell, D. (2005). A paired-watershed budget study to quantify interbasin groundwater flow in a lowland rain forest, Costa Rica. *Water Resources Research*, *41*, W04011. https://doi.org/10.1029/2004WR003635

Gleeson, T., & Manning, A. H. (2008). Regional groundwater flow in mountainous terrain: Three-dimensional simulations of topographic and hydrogeologic controls. *Water Resources Research*, *44*, W10403. https://doi.org/10.1029/2008WR006848

Goode, D. J. (1996). Direct simulation of groundwater age. *Water Resources Research*, *32*, 289–296. https://doi.org/10.1029/95WR03401

Haitjema, H. M. (1995). On the residence time distribution in idealized groundwatersheds. *Journal of Hydrology*, *172*, 127–146. https://doi.org/10.1016/0022-1694(95)02732-5

Hale, V. C., & McDonnell, J. J. (2016). Effect of bedrock permeability on stream base flow mean transit time scaling relations: 1. A multiscale catchment intercomparison. *Water Resources Research*, *52*, 1358–1374. https://doi.org/10.1002/2014WR016124

Heilweil, V. M., Healy, R. W., & Harris, R. N. (2012). Noble gases and coupled heat/fluid flow modeling for evaluating hydrogeologic conditions of volcanic island aquifers. *Journal of Hydrology*, *464–465*, 309–327. https://doi.org/10.1016/j.jhydrol.2012.07.019

Hubbard, S. S., Williams, K. H., Agarwal, D., Banfield, J., Beller, H., Bouskill, N., et al. (2018). The East River, Colorado, Watershed: A mountainous community testbed for improving predictive understanding of multiscale hydrological–biogeochemical dynamics. *Vadose Zone Journal*, *17*, 18006. https://doi.org/10.2136/vzj2018.03.0061

Hunt, A. G. (2015). *Noble Gas Laboratory's standard operating procedures for the measurement of dissolved gas in water samples*. U.S. Geological Survey Techniques and Methods Book 5, Chapter A11, 22 p. http://dx.doi.org/10.3133/tm5A11

Jasechko, S., Kirchner, J. W., Welker, J. M., & McDonnell, J. J. (2016). Substantial proportion of global streamflow less than three months old. *Nature Geoscience*, *9*, 126–129. https://doi.org/10.1038/ngeo2636

Johnson, M. R., Wanty, R. B., Bembenek, A. J., Verplanck, P. L., & Manning, A. H. (2019). *Geochemical analyses of surface water, groundwater and springs surrounding Mount Emmons near Crested Butte, Colorado (ver. 2.0, September 2020)* U.S. Geological Survey Data Release. https://doi.org/10.5066/P9CQJ0XR

Jung, M., & Aeschbach, W. (2018). A new software tool for the analysis of noble gas data sets from (ground)water. *Environmental Modelling & Software*, *103*, 120–130. https://doi.org/10.1016/j.envsoft.2018.02.004

Jung, M., Wieser, M., von Oehsen, A., & Aeschbach-Hertig, W. (2013). Properties of the closed-system equilibration model for dissolved noble gases in groundwater. *Chemical Geology*, *339*, 291–300. https://doi.org/10.1016/j.chemgeo.2012.08.006

Jurgens, B. C. (2018). *Data for tritium deposition in precipitation in the United States, 1953–2012*. U.S. Geological Survey Data Release https://doi.org/10.5066/P92CEFXN

Jurgens, B. C., Böhlke, J. K., & Eberts, S. M. (2012). *TracerLPM (Version 1): An Excel® workbook for interpreting groundwater age distributions from environmental tracer data*. U.S. Geological Survey Techniques and Methods Report. https://pubs.usgs.gov/tm/4-f3

Kimball, B. A., Runkel, R. L., Wanty, R. B., & Verplanck, P. L. (2010). Reactive solute-transport simulation of pre-mining metal concentrations in mine-impacted catchments: Redwell Basin, Colorado, USA. *Chemical Geology*, *269*, 124–136. https://doi.org/10.1016/j.chemgeo.2009.05.024

Kipfer, R., Aeschbach-Hertig, W., Peeters, F., & Stute, M. (2002). Noble gases in lakes and ground waters. D. Porcelli, C. J. Ballentine, & R. Wieler (Eds.), Noble gases in geochemistry and cosmochemistry, reviews in mineralogy and geochemsitry. Vol. 47 (pp. 615–700). Chantilly, VA: Mineralogical Society of America.

Knowles, J. F., Harpold, A. A., Cowie, R., Zeliff, M., Barnard, H. R., Burns, S. P., Blanken, P. D., et al. (2015). The relative contributions of alpine and subalpine ecosystems to the water balance of a mountainous, headwater catchment. *Hydrological Processes*, *29*(22), 4794–4808. https://doi.org/10.1002/hyp.10526

Lowell, R. P., Kolandaivelu, K., & Rona, P. A. (2014). *Reference module in Earth systems and environmental sciences*. Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.09132-6

Manning, A. H., & Caine, J. S. (2007). Groundwater noble gas, age, and temperature signatures in an Alpine watershed: Valuable tools in conceptual model development. *Water Resources Research*, *43*, W04404. https://doi.org/10.1029/2006WR005349

Manning, A. H., & Solomon, D. K. (2003). Using noble gases to investigate mountain-front recharge. *Journal of Hydrology*, *275*(3–4), 194–207. https://doi.org/10.1016/S0022-1694(03)00043-X

Manning, A. H., Solomon, D. K., & Thiros, S. A. (2005). $^{3}H/^{3}He$ age data in assessing the susceptibility of wells to contamination. *Ground Water*, *43*, 353–367. https://doi.org/10.1111/j.1745-6584.2005.0028.x

Manning, A. H., Wanty, R. B., Ball, L. B., & Williams, K. H. (2020). *Environmental tracer data from surface water and groundwater samples collected in Redwell Basin near Crested Butte, Colorado, 2017-2019*. U.S. Geological Survey Data Release. https://doi.org/10.5066/P9W7JBFQ

Marechal, J. C., & Etcheverry, D. (2003). The use of ^{3}H and ^{18}O tracers to characterize water inflows in Alpine tunnels. *Applied Geochemistry*, *18*, 339–351. https://doi.org/10.1016/S0883-2927(02)00101-4

Markovich, K. H., Manning, A. H., Condon, L. E., & McIntosh, J. C. (2019). Mountain-block recharge: A review of current understanding. *Water Resources Research*, *55*, 8278–8304. https://doi.org/10.1029/2019WR025676

Masbruch, M. D., Chapman, D. S., & Solomon, D. K. (2012). Air, ground, and groundwater recharge temperatures in an alpine setting. *Water Resources Research*, *48*, W10530. https://doi.org/10.1029/2012WR012100

Mayo, A. L., Morris, T. H., Peltier, S., Petersen, E. C., Payne, K., Holman, L. S., et al. (2003). Active and inactive groundwater flow systems: Evidence from a stratified, mountainous terrain. *GSA Bulletin*, *115*(12), 1456–1472. https://doi.org/10.1130/B25145.1

McClymont, A. F., Hayashi, M., Bentley, L. R., & Liard, J. (2012). Locating and characterising groundwater storage areas within an alpine watershed using time-lapse gravity, GPR and seismic refraction methods. *Hydrological Processes*, *26*(12), 1792–1804. https://doi.org/10.1002/hyp.9316

McGuire, K. J., & McDonnell, J. J. (2006). A review and evaluation of catchment transit time modeling. *Journal of Hydrology*, *330*, 543–563. https://doi.org/10.1016/j.jhydrol.2006.04.020

Michel, R. L., Jurgens, B. C., & Young, M. B. (2018). Tritium deposition in precipitation in the United States, 1953–2012, U.S. *Geological Survey Scientific Investigations Report 2018*, *5086*, 11. https://doi.org/10.3133/sir20185086

Parkhurst, D. L., & Appelo, C. A. J. (1999). User's Guide to PHREEQC (version 2)—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. *U.S. Geological Survey Water-Resources Investigations Report*, *99–4259*, 312. https://doi.org/10.3133/wri994259

Powell, W. G., Chapman, D. S., Balling, N., & Beck, A. E. (1988). Continental heat-flow density. In R. Haenel, L. Rybach, & L. Stegena (Eds.), *Handbook of terrestrial heat-flow density determination* (pp. 167–222). Dordrecht, Holland: Kluwer Academic Publishers.

Quinn, P., Cherry, J. A., & Parker, B. L. (2012). Hydraulic testing using a versatile straddle packer system for improved transmissivity estimation in fractured-rock boreholes. *Hydrogeology Journal*, *20*, 1529–1547. https://doi.org/10.1007/s10040-012-0893-8

Robinson, C. S., Lee, F. T., Scott, J. H., Carroll, R. D., Hurr, R. T., Richards, D. B., et al. (1974). *Engineering geologic, geophysical, hydrologic and rock-mechanics investigations of the Straight Creek Tunnel site and pilot bore*, Colorado: U.S. Geological Survey Professional Paper. https://doi.org/10.3133/pp815

Rumsey, C. A., Miller, M. P., Schwarz, G. E., Hirsch, R. M., & Susong, D. D. (2017). The role of baseflow in dissolved solids delivery to streams in the Upper Colorado River Basin. *Hydrological Processes*, *31*, 4705–4718. https://doi.org/10.1002/hyp.11390

Saar, M. O., & Manga, M. (2004). Depth dependence of permeability in the Oregon Cascades inferred from hydrogeologic, thermal, seismic, and magmatic modeling constraints. *Journal of Geophysical Research*, *109*, B04204. https://doi.org/10.1029/2003JB002855

Salve, R., Rempe, D. M., & Dietrich, W. E. (2012). Rain, rock moisture dynamics, and the rapid response of perched groundwater in weathered, fractured argillite underlying a steep hillslope. *Water Resources Research*, *48*, W11528. https://doi.org/10.1029/2012WR012583

Sanz, E., & Yelamos, J. G. (1998). Methodology for the study of unexploited aquifers with thermal waters: Application to the aquifer of the Alhama de Aragon hot springs. *Ground Water*, *36*(6), 913–923. https://doi.org/10.1111/j.1745-6584.1998.tb02098.x

Schaller, M. F., & Fan, Y. (2009). River basins as groundwater exporters and importers: Implications for water cycle and climate modeling. *Journal of Geophysical Research*, *114*, D04103. https://doi.org/10.1029/2008JD010636

Schlosser, P., Stute, M., Dorr, H., Sonntag, C., & Munnich, K. O. (1988). Tritium/^{3}He dating of shallow groundwater. *Earth and Planetary Science Letters*, *89*, 353–362. https://doi.org/10.1016/0012-821X(88)90122-7

Schlosser, P., Stute, M., Sonntag, C., & Munnich, K. O. (1989). Tritiogenic 3He in shallow groundwater. *Earth and Planetary Science Letters*, *94*, 245–256. https://doi.org/10.1016/0012-821X(89)90144-1

Solomon, D. K. (2000). ^{4}He in groundwater. In P. G. Cook, & A. L. Herczeg (Eds.), *Environmental tracers in subsurface hydrology* (pp. 425–439). Dordrecht, Holland: Kluwer Academic Publishers.

St Clair, J., Moo, S., Holbrook, W. S., Perron, J. T., Riebe, C. S., Martel, S. J., et al. (2015). Geophysical imaging reveals topographic stress control of bedrock weathering. *Science*, *350*(6260), 534–539. https://doi.org/10.1126/science.aab2210

Stute, M., & Schlosser, P. (2000). Atmospheric noble gases. In P. G. Cook, & A. L. Herczeg (Eds.), *Environmental tracers in subsurface hydrology* (pp. 349–377). Dordrecht, Holland: Kluwer Academic Publishers.

Stute, M., Sonntag, C., Deak, J., & Schlosser, P. (1992). Helium in deep circulating groundwater in the Great Hungarian Plain: Flow dynamics and crustal and mantle helium fluxes. *Geochimica et Cosmochimica Acta*, *56*, 2051–2067. https://doi.org/10.1016/0016-7037(92)90329-H

Tokunaga, T.K., Wan, J., Williams, K.H., Brown, W., Henderson, A., Kim, Y., et al., (2019). Depth- and time-resolved distributions of snowmelt-driven hillslope subsurface flow and transport and their contributions to surface waters, *Water Resources Research*, *55*, 9474–9499. https://doi.org/10.1029/2019WR025093

Tomonaga, Y., Marzocchi, R., Pera, S., Pfeifer, H.-R., KipferDecrouy, R. L., & Vennemann, T. (2017). Using noble-gas and stable-isotope data to determine groundwater origin and flow regimes: Application to the Ceneri Base Tunnel (Switzerland). *Journal of Hydrology*, *545*, 395–409. https://doi.org/10.1016/j.jhydrol.2016.11.043

Visser, A., Thaw, M., Deinhart, A., Bibby, R., Safeeq, M., Conklin, M., et al. (2019). Cosmogenic isotopes unravel the hydrochronology and water storage dynamics of the Southern Sierra Critical Zone. *Water Resources Research*, *55*, 1429–1450. https://doi.org/10.1029/2018WR023665

Waples, D. W., & Waples, J. S. (2004). A review and evaluation of specific heat capacities of rocks, minerals, and subsurface fluids. Part 1: Minerals and nonporous rocks. *Natural Resources Research*, *13*, 97–122. https://doi.org/10.1023/B:NARR.0000032647.41046.e7

Welch, L. A., & Allen, D. M. (2014). Hydraulic conductivity characteristics in mountains and implications for conceptualizing bedrock groundwater flow. *Hydrogeology Journal*, *22*(5), 1003–1026. https://doi.org/10.1007/s10040-014-1121-5

Welch, L. A., Allen, D. M., & van Meerveld, H. J. (2012). Topographic controls on deep groundwater contributions to mountain headwater streams and sensitivity to available recharge. *Canadian Water Resources Journal*, *37*(4), 349–371. https://doi.org/10.4296/cwrj2011-907

White, A., Moravec, B., McIntosh, J., Olshansky, Y., Paras, B., Sanchez, R. A., et al. (2019). Distinct stores and the routing of water in the deep critical zone of a snow-dominated volcanic catchment. *Hydrology and Earth System Sciences*, *23*, 4661–4683. https://doi.org/10.5194/hess-23-4661-2019

References From the Supporting Information

Hall, C. M., Castro, M. C., Lohmann, K. C., & Ma, L. (2005). Noble gases and stable isotopes in a shallow aquifer in southern Michigan: Implications for noble gas paleotemperature reconstructions for cool climates. *Geophysical Research Letters*, *32*, L18404. https://doi.org/10.1029/2005GL023582

Paul, M. J., Biegalski, S. R., Haas, D. A., Jiang, H., Daigle, H., & Lowrey, J. D. (2018). Xenon adsorption on geological media and implications for radionuclide signatures. *Journal of Environmental Radioactivity*, *187*, 65–72. https://doi.org/10.1016/j.jenvrad.2018.01.029