



Isotopic evaluation of the National Water Model reveals missing agricultural irrigation contributions to streamflow across the western United States

Annie L. Putman¹, Patrick C. Longley², Morgan C. McDonnell¹, James Reddy³, Michelle Katoski⁴, Olivia L. Miller¹, and J. Renée Brooks⁵

¹US Geological Survey Utah Water Science Center

²US Geological Survey Colorado Water Science Center

³US Geological Survey New York Water Science Center

⁴US Geological Survey Maryland-Delaware Water Science Center

⁵US Environmental Protection Agency, Pacific Ecological Systems Division

Correspondence: Annie L. Putman (aputman@usgs.gov)

Abstract. The National Water Model (NWM) provides critical analyses and projections of streamflow that support water management decisions. However, the NWM performs poorly in lower elevation rivers of the western United States (US). The accuracy of the NWM depends on the fidelity of the model inputs and the representation and calibration of model processes and water sources. To evaluate the NWM, we performed a water isotope ($\delta^{18}O$ and δ^2H) mass balance using long term mean summer hydrologic fluxes between 2000 and 2019, and gridded precipitation and groundwater isotope ratios. We compared the NWM-flux-estimated ('model') river reach isotope ratios to 4503 in-stream water isotope observations in 877 reaches across 5 basins in the western US. A simple regression between observed and mass balance estimated isotope ratios explained 57.9% ($\delta^{18}O$) and 67.1% (δ^2H) of variance, though observations were 0.5‰ ($\delta^{18}O$) and 4.8‰ (δ^2H) higher, on average, than mass balance estimates. The unexplained variance suggest that the NWM does not include all relevant water fluxes to rivers. To infer possible missing water fluxes, we evaluated patterns in observation-model differences using $\delta^{18}O_{diff}$ ($\delta^{18}O_{obs} - \delta^{18}O_{mod}$) and d_{diff} ($\delta^2H_{diff} - 8 * \delta^{18}O_{diff}$). We detected evapoconcentration of observations relative to model estimates (negative d_{diff} and positive $\delta^{18}O_{diff}$) at lower elevation, higher stream order, arid sites. The catchment actual evaporation to precipitation ratio, the fraction of streamflow estimated to be derived from agricultural irrigation, and whether a site was reservoir-affected were all significant predictors of d_{diff} in a linear mixed effects model, with up to 15.1% of variance explained by fixed effects. This finding is supported by patterns in groundwater levels and groundwater isotope ratios, and suggests the importance of including irrigation return flows to rivers, especially in lower elevation, higher stream order, arid rivers of the Western US.

1 Introduction

The western United States (US) is experiencing multidecadal drought (Williams et al., 2022) and declining streamflows (Milly and Dunne, 2020). Major rivers are running dry (Kornfield, 2022), lakes are shrinking (Ramirez, 2022; Fergus et al., 2020, 2022), and water users are experiencing shortages and cuts (Bureau of Reclamation, Department of the Interior, 2022). These de-



creases in streamflow and groundwater fluxes are projected to continue in coming years (Miller et al., 2021b, a), with projected decreases in snowpack (Mote et al., 2021; Siirila-Woodburn et al., 2021) and increases in temperatures (Hicke et al., 2022). Under drought and snow drought stress, as well as changing wintertime precipitation patterns, river flows may become more difficult to forecast (Hammond and Kampf, 2020; Siirila-Woodburn et al., 2021). Yet, with decreasing water availability, water managers and other stakeholders tasked with managing and responding to current and future water supply increasingly depend on accurate streamflow predictions.

Fully routed, high spatial and temporal resolution streamflow models, like the National Oceanic and Atmospheric Administration's National Water Model (NWM) which utilizes Weather Research and Forecasting (WRF) Hydro model data (Gochis et al., 2018), provide short and medium term streamflow prediction in the United States, as well as analyses of past stream discharge at ungaged locations. The accurate, detailed, frequent results from the National Water Model may be used by emergency managers, reservoir operators, floodplain managers, and farmers to aid in water use decision making and flood or pollution risk evaluation. The accuracy of predictions and current snapshots produced by the model depend on 1) inclusion and faithful representation of relevant water sources and hydrologic processes 2) appropriate calibration of parameter estimations and 2) the fidelity of the model inputs.

With respect to the faithful representation of water sources, the major water sources to streams in the mountainous west include two broad water flux categories: runoff (also called quickflow) and groundwater discharge (also called baseflow). Runoff during the summer comes from late season snowmelt, rain, and irrigation water. Groundwater discharge comes from shallow or deep in-ground water, typically recharged at high elevation by snowmelt. Rivers in the west derive the majority of their water from springtime melt of high elevation wintertime snowpack (Li et al., 2017; Hammond et al., 2023) and little water is contributed to streams at lower elevations where there is minimal snowpack (Miller et al., 2021b). Some of the melt water enters streams as runoff during late spring and summer, while the remainder recharges shallow and deep groundwater and later in the season or in subsequent years enters the stream as groundwater discharge (Barnhart et al., 2016; Miller et al., 2021a; Brooks et al., 2021; Wolf et al., 2023). Rain contributes runoff to streamflow, but even in areas receiving a substantial proportion of their total annual precipitation during summer in association with the North American Monsoon, only a small proportion of the total precipitation makes it to the stream (Solder and Beisner, 2020; Tulley-Cordova et al., 2021); the remainder is evaporated from soils or transpired by plants (Milly and Dunne, 2020). Thus, lower elevation streams, particularly later in the summer, depend heavily on groundwater discharge from higher elevations to sustain their flows (Miller et al., 2016) and the majority of streams in lower elevation arid areas are likely to lose water to shallow groundwater recharge (Jasechko et al., 2021).

Within this hydrologic framework, human water use and management introduces complexity via reservoirs and managed release schedules, trans- and interbasin transfers, conveyances, and surface and groundwater withdrawals, as well as irrigation for agricultural crop or turf grass growth. Turf irrigation in cities composes the majority of household water use in most municipalities and agricultural irrigation can comprise up to 80% of total statewide water use in Western US states (Dieter et al., 2018). Water used for agricultural crop or turf grass growth locally intensifies water balance fluxes, through increases in both water application and evapotranspiration in these select tracts of land. Depending on the method, both agriculture and turf



grass irrigation can contribute to local groundwater recharge (Grafton et al., 2018), with greater recharge coming from flood irrigation compared to sprinkler or drip irrigation methods. Water for irrigation can come from either surface or groundwater withdrawals. The irrigation water source may have both direct and indirect influences on streamflows particularly during low flow seasons (Essaid and Caldwell, 2017; Condon and Maxwell, 2019). However, these processes and fluxes are not currently explicitly included in the NWM.

Past NWM evaluations have leveraged streamgauge measurements (Hansen et al., 2019; Towler et al., 2023; Seo et al., 2021) and model evaluation using streamgauge measurements is included in the NWM WRF-Hydro workflow (Gochis et al., 2018). Using measured discharge to evaluate the NWM is useful because the data are publicly available at high spatial and temporal resolution (e.g., dataset used in (Towler et al., 2023)). However, evaluation of streamflows with measured discharge 1) may allow modelers to get the correct total streamflow values and temporal patterns at a reach for the wrong process reasons or 2) may suggest that the model could be improved due to mismatches between measured and modeled data, but cannot provide information on the specific process(es) or sources responsible for the errors.

Among the climatic regions covered by the NWM, model streamflow evaluation metrics perform the most poorly in the Western US in lower elevation reaches. Metrics like the Kling–Gupta efficiency (KGE) indicate pervasive mismatches between measured and modeled streamflows and percent bias (PBIAS) results showed that simulated streamflow volumes tend to be overestimated in the west (Towler et al., 2023). Similarly, Hansen et al. (2019) found that the NWM has difficulty estimating flows during drought or low flow years in the Colorado River Basin. In the low elevation stream reaches of the Western US, disagreement between the NWM flows and observations within anthropogenically-altered reaches may come from incomplete representation of anthropogenic water sources or processes in the NWM.

In the western US, low elevation waterways have moderate to high potential for anthropogenic alteration (Fergus et al., 2021). For example, rivers and surface water supplies are managed by dams, and a large proportion of total water use is allocated to irrigating agriculture (Dieter et al., 2018). However, the NWM does not explicitly include surface water removal for agricultural irrigation nor subsurface return flows from irrigation in its streamflow computations. Likewise, the NWM represents inflow and outflow of lakes and reservoirs as passive storage and releases, with no active reservoir management. Both of these omissions may be contributors to the large errors observed in the NWM in lower elevation areas where land use includes large amounts of along-river agriculture and streamflow is heavily managed through reservoir operations. Unfortunately, the effects of contributions of these two water sources on streamflow are difficult to identify and quantify through evaluations of streamflow records alone.

Tracers, including the stable isotopes of water (^{16}O , ^{18}O and ^1H , ^2H), can be used to track the contributions of specific water sources (e.g., groundwater fluxes, runoff fluxes) that may be impossible to disentangle from comparisons with total or instantaneous discharge. Stable isotopes of oxygen and hydrogen (expressed in delta notation, $\delta = 1000 * (\frac{R_{\text{sample}} - R_{\text{standard}}}{R_{\text{standard}}})$), where R is the ratio of concentrations of the heavy (^2H and ^{18}O) and light (^1H and ^{16}O) isotopes) have diagnosed process limitations in other modeling contexts (Nusbaumer et al., 2017; Putman et al., 2019), leveraging the fractionation occurring for stable isotopes of waters during phase changes (Bowen et al., 2019). The secondary parameter, deuterium excess (defined as $d = \delta^2\text{H} - 8 * \delta^{18}\text{O}$), is a reliable metric for detecting evaporation of precipitation and surface waters, evaporation under



a vapor pressure gradient or non-equilibrium condensation processes, like snow formation in mixed phase clouds or isotopic fractionation during the melting of snow (Putman et al., 2019; Bowen et al., 2018; Ala-aho et al., 2017).

Because hydrologic processes including groundwater recharge, discharge, and precipitation runoff do not cause isotopic fractionation, we can use water fluxes from hydrologic models with estimates of the isotope ratios of those fluxes on the appropriate timescales to produce river water isotope estimates. This works well because the groundwater and runoff fluxes to summertime streamflow in the Western US have different, and often distinct, stable isotope ratios due to seasonal and spatial controls on precipitation isotope ratios. The signatures of groundwater inflow and snowmelt tend to have the lowest isotope ratios of the water sources in the hydrologic system and tend to be relatively temporally invariant (Bowen, 2008; Feng et al., 2009; Jasechko et al., 2014; Solder and Beisner, 2020; Tulley-Cordova et al., 2021). In contrast, summer precipitation, which contributes runoff to streams, tends to have higher isotope ratios than groundwater (Jasechko et al., 2014; Tulley-Cordova et al., 2021).

Anthropogenic modifiers of streamflow that are not included explicitly in the NWM may be expected to alter the isotopic signature of streamflow downstream of the river water source areas in the headwaters. Agricultural irrigation can contribute both runoff to streams and recharge groundwater (Essaid and Caldwell, 2017; Gochis et al., 2018). This specific water source becomes isotopically enriched during conveyance and application both due to direct evaporation during conveyance and soil evaporation after application (Craig and Gordon, 1965; Yang et al., 2019). The evaporated isotope signatures of agricultural soil waters are also evident in fruit water (Oerter et al., 2017). Thus, irrigation return flows (runoff or ground) associated with irrigation are much more evaporatively enriched, with lower d , than naturally recharged groundwater or precipitation runoff. The effects of evaporation on the isotope ratios of the return flows are expected to be greater in arid areas with higher summer temperatures and higher vapor pressure deficits. Although lakes can be evapoconcentrated relative to other surface waters (Bowen et al., 2018), we do not expect similar signals of evaporative enrichment from reservoirs. Relative to natural lakes across the US, evaporation rates from western lakes are low relative to inflow (Brooks et al., 2014). Instead, reservoirs may alter the isotope ratios of streamflow through retention of and later discharge of spring snowmelt. Thus, reservoir outflow may have lower isotope ratios and higher d than the upstream rivers during the summer months.

In this study, we performed an isotope mass balance using the water fluxes of the NWM and gridded water stable isotope datasets. We compared the isotope mass balance-derived long term mean streamflow isotope ratios directly with stream water isotope observations. If the NWM constrains all water sources affecting streamflow, we expect the differences between the isotope mass balance results and isotopic observations (observation-model differences) will be small and be uniformly positive or negative throughout each basin. If we observe spatial and/or seasonal variability and structured patterns in observation-model differences within basins (i.e., patterns with elevation, stream order, or aridity), particularly with respect to the sign of the difference, we may infer that the NWM is incorrectly partitioning runoff and groundwater fluxes, or missing important water sources. We hypothesize that if we observe spatial variability and structured patterns in our observation-model difference data, we will observe higher isotope ratios and lower d in more arid reaches reflecting the influence of evapoconcentrated irrigation return flows on streamflow as compared to higher elevation, humid or seasonally snowy reaches with minimal anthropogenic influence.



2 Methods

Using an isotope mass balance approach, we calculated the long term mean stable isotope ratios of river reaches in the western US (Figure 1). Water fluxes were supplied by NWM simulations of groundwater and runoff fluxes (National Oceanographic and Atmospheric Administration, 2022) and isotope ratios came from gridded groundwater and precipitation stable isotope products (Bowen, 2022b; Bowen et al., 2022). We compared calculated river isotope ratios to observations of isotope ratios. The data compiled and calculated data presented and analyzed in this manuscript are publicly available (Reddy et al., 2023). Differences between observations and modeled data were compared in an error-partitioning framework. Where applicable, results were interpreted in terms of the estimates of the influence of agricultural water use in the upstream area as well as the presence of impoundments on the river. A groundwater isotope ratio dataset and a well water surface elevation relative to river surface elevation dataset from Jasechko et al. (2021) were used as independent lines of evidence supporting our analysis of observation- mass balance estimate differences.

2.1 Temporal domain

Our analysis was constrained to summer months (June, July, August) between 2000 and 2019. The specific months chosen reflect those with greatest evapotranspiration and thus consumptive water use and correspond to the season with the largest number of spatially distributed river water isotope observations.

2.2 Spatial domain

We selected 5 HUC2 basins (U.S. Geological Survey, National Geospatial Technical Operations Center, 2023) in the Western US to compose our study area: the Upper Colorado (14), Lower Colorado (15), Great Basin (16), Pacific Northwest (17), and California (18). All basins were characterized by rivers sustained by wintertime snowpack mediated by groundwater infiltration and discharge. All basins also included water management through impoundments and substantial water use for agriculture. In a simplified Köppen climate classification (Rubel and Kottek, 2010), the southern and central portions of the study area were characterized as arid, whereas much of the northern and mountainous portions of the study area was classified as warm temperate or seasonally snowy.

2.3 Data assimilation framework

Various data types used in this analysis (point, raster, vector) were spatially joined to coarsened catchments (n=15787) derived from the National Hydrography Dataset Plus (NHDPlus, (U.S. Geological Survey, 2019), see Text S1 for details on catchment coarsening). The full NHDPlus network was clipped to the spatial domain of our study. Each catchment contained one reach. Coarsened catchments had a median size of 51 km² and a mean size of 221 km², and flowlines had a median length of 20 km² and a mean length of 32 km². We also utilized some attributes provided with the NHDPlus, including the catchment area, Strahler stream order, the reach length, minimum and maximum catchment elevation, and the feature code, which denoted the flowline path type.

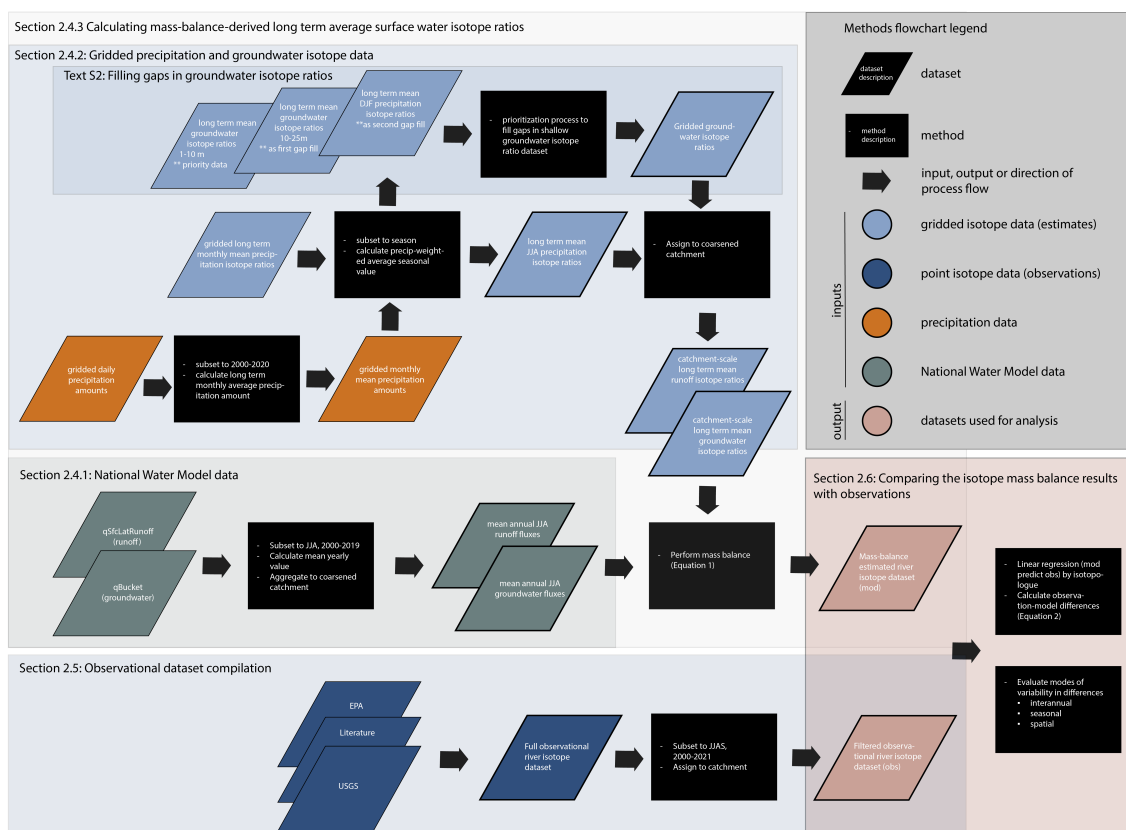


Figure 1. Diagram showing methods and datasets as described in Sections 2.4-2.6. Four datastreams were used to formulate the long isotope mass balance estimates of river isotope ratios: gridded precipitation isotope estimates Bowen (2022b), gridded groundwater isotope estimates (Bowen et al., 2022), gridded precipitation data (University of East Anglia Climatic Research Unit and Harris, I.C. and Jones, P.D. and Osborn, T., 2021) and NWM data (National Oceanographic and Atmospheric Administration, 2022). Three data categories contributed to the observational river isotope dataset: USGS (U.S. Geological Survey, 2022), EPA (U.S. Environmental Protection Agency, 2016b, 2020), and literature datasets accessed from the waterisotopesdb.org (Putman and Bowen, 2019).



2.4 Performing the isotope mass balance with NWM water fluxes and gridded water isotope ratios

Using estimates of long term mean groundwater and precipitation isotope ratios (Bowen, 2022b; Bowen et al., 2022), we applied an isotope mass-balance to the NWM groundwater and runoff fluxes to streams (Figure 1). The operational hydrologic model is based on inputs from WRF-Hydro (Gochis et al., 2020b, a) and simulates and forecasts major water components (e.g., evapotranspiration, snow, soil moisture, groundwater, surface inundation, reservoirs, and streamflow) in real time across the CONUS, Hawaii, Puerto Rico, and the US Virgin Islands. In the NWM framework, surface and soil evaporation are wrapped into the evapotranspiration flux variable, and direct evaporation from rivers and reservoirs are not considered in the NWM surface water balance. Thus, we did not apply any additional isotopic fractionation to the groundwater and runoff isotopic fluxes. This approach produced the ‘characteristic’ long term mean isotope ratio expected for river reaches in the western US. These estimates were directly comparable to river water isotope observations.

2.4.1 National Water Model data

We accessed the groundwater ‘bucket’ (q_{Bucket} , $\text{m}^3 \text{s}^{-1}$) runoff ($q_{\text{SfcLatRunoff}}$, $\text{m}^3 \text{s}^{-1}$), and streamflow (streamflow, $\text{m}^3 \text{s}^{-1}$) fluxes from the NWM v 2.1 Analysis Assimilation dataset (National Oceanographic and Atmospheric Administration, 2022). We used runoff and groundwater fluxes for our mass balance estimates (Figure 2), and retained streamflow as a reach scale quantity to be included in data analyses. These variables are available at the NHDPlus reach scale on an hourly timestep between 2000 and 2019. We subset these variables to the summer months (June, July, and August) and calculated the mean water fluxes to each reach for the summer season of each year. The interannual variability in the summer fluxes was leveraged as an estimate of the uncertainty of the long term mean summer water fluxes.

2.4.2 Gridded precipitation and groundwater isotope data

The precipitation and groundwater stable isotope ratios ($\delta^2\text{H}$, $\delta^{18}\text{O}$) that we used to perform the isotope mass balance came from two different publicly available gridded products. Both represent long term means or climatologies and provide estimates of uncertainty.

We obtained monthly precipitation isotope ratio climatological predictions and uncertainty estimates (1 standard deviation) for both H and O from Bowen (2022b). The monthly USA grids were available at 1 km, and were produced with the OIPC v3.2 database (Bowen, 2022a) following methods described in Bowen et al. (2005). Monthly grids have been adjusted for consistency with annual values (see version notes for OIPC2.0 (Bowen, 2006)). We calculated the precipitation-weighted long term mean summer (June, July, August) and winter (December, January, February) seasonal isotope ratio climatologies with long term monthly mean precipitation climatologies calculated from the Climatic Research Unit (CRU) mean monthly precipitation amounts (Harris et al., 2020; University of East Anglia Climatic Research Unit and Harris, I.C. and Jones, P.D. and Osborn, T., 2021) for the period 2000-2020. The precipitation weighted mean seasonal climatology error was calculated analytically from the timeseries.



The groundwater isoscapes used in this analysis were produced by Bowen et al. (2022) for 7 depth intervals ranging from 1 to 1000m. The groundwater isoscapes were not temporally resolved. Because this project focuses on groundwater discharge to streams, we preferentially utilized the 1-10m depth interval. However, this layer contained some data gaps where insufficient well data were present to perform an estimate. Where available, we filled these data gaps using method outlined in Text S2. The groundwater isotope ratio data included estimates of uncertainty, which were retained for the characterization of uncertainty around the mass balance isotope ratio estimates.

The gridded precipitation and groundwater isotope datasets and their uncertainties were assimilated to the NHDPlus spatial framework. Because the raster data grid sizes were larger than the catchment sizes we employed a distance minimization approach using the centroid of the catchment and the centroids of the grid cells.

2.4.3 Calculating mass-balance-derived long term mean surface water isotope ratios

To estimate the long term mean surface water isotope ratio ($R_{sw,r}$) at each reach (r) in the spatial domain (Equation 1), we accumulated the groundwater (gw) and runoff (ro) isotope fluxes (i.e., the isotope ratio multiplied by the water flux, $R * F$) for all reaches (i) from the headwaters downstream to the reach. The isotope ratio for runoff (R_{ro}) came from the summer mean gridded precipitation isotope ratios and the isotope ratio for the groundwater flux (R_{gw}) came from the gridded groundwater isotope ratios (see Section 2.4.2). The summed isotope fluxes were divided by the summed runoff and groundwater fluxes.

$$R_{sw,r} = \frac{\sum_{i=0}^r R_{gw,i} * F_{gw,i} + R_{ro,i} * F_{ro,i}}{\sum_{i=0}^r F_{gw,i} + F_{ro,i}} \quad (1)$$

Because volumetric contributions of groundwater and runoff to streamflow in each reach varied with year, and because the long term mean estimates of the groundwater and precipitation isotope ratios include estimates of uncertainty, we evaluated the potential for uncertainty in our estimates of $R_{sw,r}$. We used the interannual variability in the mean summer water fluxes as characteristic of the uncertainty in our estimate of our long term mean water fluxes, and used the uncertainty estimates provided with the isotope ratio products as representative of the uncertainty in the isotope ratios. We performed 10 random draws of the isotope ratio distributions for each of the 20 years of record, which gave a total of 10 estimates of R_{sw} per year over 20 years of record, or 200 total estimates of R_{sw} per reach. Joint distributions (of either H and O, or isotopes with water fluxes), were not used because information about how the isotope ratios might covary was not available from the gridded isotope datasets and no assumptions were made about how the isotopes might vary with interannual variability in climatic conditions. Similarly, no assumptions were made that the precipitation and groundwater isotope ratios covaried in time. These 200 estimates were used to calculate a long-term mean estimated isotope ratio for river water in each reach of the network and to evaluate uncertainty in our estimates.

2.5 Compilation of river isotope observations

The results of the mass balance were compared with observations of stable water isotope ratios from rivers collected between 2000 and 2021, during the growing season months of June, July, August and September. We included two additional years (2020 and 2021) as well as data from the month of September beyond the temporal constrains of the NWM model domain in



220 our set of observations. This decision as made to maximize the amount of data and number of unique river reaches in the spatial domain that are available for analysis, and reflects the assumption that the long term mean river isotope ratios calculated from the mass balance approach will be insensitive to inclusion or exclusion of a small number of additional years or an additional growing season month.

We compiled surface water stable isotope ($\delta^2\text{H}$, $\delta^{18}\text{O}$) measurements from various sources including the Environmental Protection Agency (EPA), the United States Geological Survey (USGS) National Water Information System (NWIS, U.S. Geological Survey (2022)), and published datasets assimilated in the Water Isotopes Database (Putman and Bowen, 2019). Not all reaches had one or more stable water isotope observations, and river reaches with multiple stable water isotope ratio observations were sometimes, but not always, from the same sampling site within the catchment.

The EPA surface water stable isotope data came from the National Rivers and Streams Assessments (NRSA, U.S. Environmental Protection Agency (2016b, 2020)) and the National Lakes Assessment (NLA, U.S. Environmental Protection Agency (2009, 2016a)). These data were collected once or twice per summer on a five year rotating basis as part of routine sampling campaigns. Over the time period of our analysis, we obtained three collections of NRSA samples (2008-2009, 2013-2014, and 2018-2019). Sites were sometimes but not always resampled among the campaigns. Sampling was stratified based on Strahler stream order and by state ensuring that all orders were sampled within each state in the assessments (U.S. Environmental Protection Agency, 2016b, 2020). This means that higher order reaches are less frequently sampled than medium or low order reaches.

The USGS surface water stable isotope data for rivers were downloaded through the NWIS API (U.S. Geological Survey, 2022) and the literature data came from published and unpublished sources that are publicly available through the Water Isotopes Database (Putman and Bowen, 2019). Stable isotope collections are not part of routine measurements for the USGS, but rather are collected by specific USGS projects. Thus, stable isotope data collections from the USGS and literature datasets tended to be spatially and temporally clustered.

2.6 Comparing the isotope mass balance results with observations

The relationship of the NWM isotope mass balance (modeled) to the river isotope observations were evaluated using correlation and simple regression analyses, where the modeled isotope ratio (either $\delta^2\text{H}$ or $\delta^{18}\text{O}$) values are used to predict the observed isotope ratios. We evaluated the results with all unaveraged observations and mean isotope ratio at river reaches with multiple observations. A Pearson correlation analysis was performed using the 'corr()' function of python's 'pandas' package (Wes McKinney, 2010; The pandas development team, 2020). Regression analysis was performed using the ordinary least squares (OLS) function in the python 'statsmodels' package (Seabold and Perktold, 2010).

We calculated the likelihood that an observation and the model result came from the same distribution, based on the variance in the model estimate, and variance associated with river water isotope observations (Text S3) using a two-tailed t-test. We reported p-values, where $p < 0.1$ indicated that the isotope mass balance estimate was statistically different from the observed surface water isotope ratio for the specific element (H or O).



2.6.1 Calculating observation-model differences

We calculated the observation (*obs*)-model estimate (*mod*) differences in both $\delta^{18}O$ and δ^2H , by subtracting the model estimate from the observation ($\delta^{18}O_{diff} = \delta^{18}O_{obs} - \delta^{18}O_{mod}$; $\delta^2H_{diff} = \delta^2H_{obs} - \delta^2H_{mod}$). Using both isotope systems, we established a framework for interpretation of our results (Figure 2) that utilizes movement along or deviation from the global mean $\delta^2H : \delta^{18}O$ ratio of 8 that is used to represent fractionation that occurs at equilibrium and defines the slope of the Global Meteoric Water Line (GMWL Craig, 1961).

Observation-model differences may arise from either 1) incorrect model source representation (i.e., missing water sources or incorrect fluxes of established sources) or 2) errors in the isotope ratio datasets used for the isotope mass balance calculation. Thus, for positive or negative values of $\delta^{18}O_{diff}$ and δ^2H_{diff} that exhibit a $\delta^2H_{diff} : \delta^{18}O_{diff}$ ratio of 8, we infer either errors in NWM with respect to the proportions of runoff and groundwater contributed, or errors in the gridded isotope ratios (likely groundwater, due to its disproportionate contributions). For positive or negative $\delta^{18}O_{diff}$ and δ^2H_{diff} with $\delta^2H_{diff} : \delta^{18}O_{diff}$ ratios different from 8, we infer that the NWM is missing uncharacterized water sources with isotope values bearing a signature of non-equilibrium fractionation. We quantify differences of $\delta^2H_{diff} : \delta^{18}O_{diff}$ ratios from 8 using a metric of similar to d (Equation 2).

$$d_{diff} = \delta^2H_{diff} - 8 * \delta^{18}O_{diff} \quad (2)$$

We can interpret combinations of $\delta^{18}O_{diff}$ and d_{diff} together, as well as d_{diff} independently to infer the uncharacterized sources responsible for the observation-model difference. This framework is useful because the ratios of δ^2H to $\delta^{18}O$ of the isotopic inputs to the isotope mass balance tend to be close to 8 (Bowen, 2022b; Bowen et al., 2022) whereas those from the observations more often differ from 8 (U.S. Environmental Protection Agency, 2016b, 2020). This means that all non-zero d_{diff} values can be used to identify omitted water sources and where they are important to streamflow.

2.7 Evaluating modes of variability in observation-model differences

We evaluated interannual, seasonal, and spatial modes of variability in our observation-model differences. The interannual variability was assessed to ensure that patterns in the other modes of variability did not arise due to the timescale difference between our isotope mass balance estimates (long term mean) and observations (instantaneous). The seasonal and spatial variability were investigated to understand the reasons for observation-model differences in our dataset.

2.7.1 Evaluating interannual variability in observation-model differences

To assess the influence of interannual variability on unexplained variance in our observation-model comparison, we utilized the years (2008, 2009, 2013, 2014, 2018, 2019) where sampling was conducted by the EPA NRSA, because these years are likely to have the most representative and consistent spatial distribution of samples. Every NRSA-sampled year had over 250 observations and represented more than 100 river reaches. No other years exhibited this spatial representation. We calculated

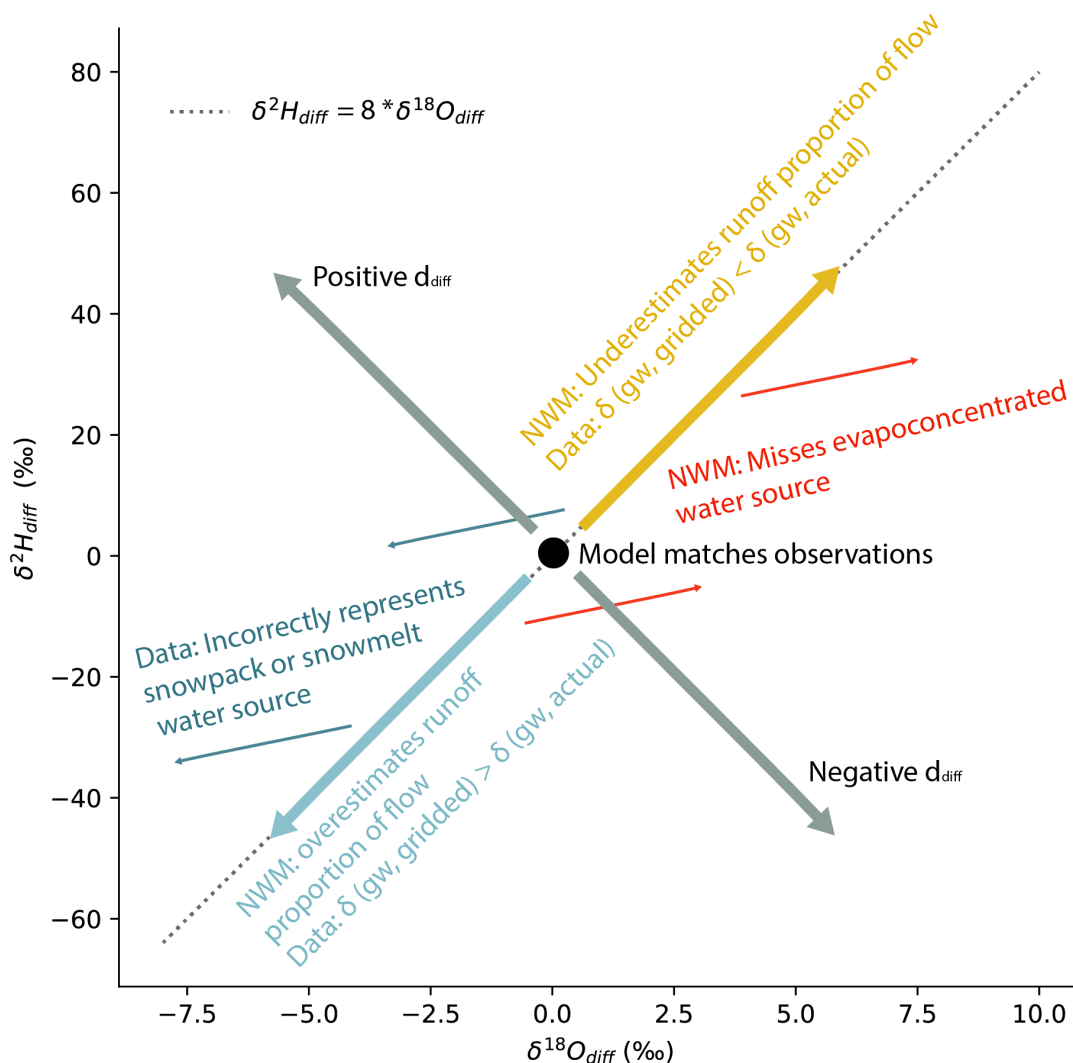


Figure 2. Schematic for interpretations of observation model differences utilizing dual isotope space and assumptions about the expected relationships between $\delta^{18}O_{diff}$ and δ^2H_{diff} . The annotations associated with ‘NWM’ specify the sort of hydrologic model error (i.e. water source apportionment) that could produce the observation-model comparison result, if all isotope data supplied to the isotope mass balance are correct. The annotations associated with ‘Data’ specify the sort of error in the gridded isotope datasets that could produce the observation-model result if all NWM water source contributions are assumed to be correct.



the mean and standard deviation of the observation-model differences, and performed a regression on the differences (e.g., as described for all data in Section 2.6.1). The EPA method is sufficiently standardized (Theobald et al., 2007) that the regression results for each year are unlikely to primarily reflect spatial variability in sampling locations.

2.7.2 Evaluating variability in observation-model differences across the growing season

To assess how the observation model difference may change over the growing season, we obtained all sites-year combinations where there were at least three observations during at least three of the four months (Jun-Sep) of the growing season. We required one of the months be the month of June to form the basis for comparison. From the June value(s) of $\delta^{18}O_{diff}$ and d_{diff} for a site-year combination, we subtracted the $\delta^{18}O_{diff}$ and d_{diff} values calculated for other months at the same site and from the same year. We evaluated the distribution of the aggregate results, as well as the distributions at the HUC2 basin scale by comparing their means and interquartile ranges.

2.7.3 Evaluating sources of spatial variability in observation-model differences

Spatial structure in the observation-model differences were evaluated graphically by comparison of $\delta^{18}O_{diff}$ and d_{diff} with catchment mean elevation, Strahler stream order, and Köppen climate class (Rubel and Kottek, 2010). The former two variables were retained from the NHDPlus catchment dataset (U.S. Geological Survey, 2019). The former was joined to the spatial framework as described in Text S4.

Spatial structure in the observation-model differences were evaluated statistically with linear mixed effects modeling using the basin (HUC2) as a random variable using the python ‘statsmodels’ module and the ‘mixedlm()’ function (Seabold and Perktold, 2010). Linear mixed effects modeling with basin as the random (grouping) variable was selected for the analysis method because water in streams at low elevations is likely to be more isotopically similar to water in the basin headwaters than a nearby stream in a different basin with different water source regions. Thus, we assume the groups are likely to have different mean values reflecting their hydrologic and climatic differences. Although we also expect that the relationship of the response variable d_{diff} to the explanatory variables may differ among basins, both our response and explanatory variables contain substantial scatter as well as small numbers of high leverage points in each basin, such that a more nuanced analysis would be likely to produce misleading results.

Using the linear mixed effects approach, we tested the statistical relationship between d_{diff} and the ratio of actual evaporation to precipitation ($\frac{ET_a}{P}$, Text S4), catchment mean elevation, fraction of streamflow estimated to come from agricultural return flows (Text S5), and a categorical variable indicating influence of large reservoirs (capacity >50,000 acre-feet, Text S5.2). We performed statistical analysis on all sites on streams not categorized as intermittent, ditches, or canals.

2.8 Evaluation of independent lines of evidence supporting signature of agricultural water use in rivers

Because it is difficult to disentangle the effects of elevation and aridity from the effects of human water use and management due to their spatial covariance, we utilized analyses of independent datasets to support the results of our statistical inference.



The analyses evaluated relationships between land use or cover and groundwater isotope ratios and the fraction of well water
315 levels that are below the nearby river level in catchments across the western US.

2.8.1 Associating groundwater stable isotope observations with land use / land cover types

Estimates of the isotopic evapoconcentration of groundwater associated with different land use and land cover classes supports
our inferences from observation-model differences. We made the associations between groundwater isotope ratios and land use
classes at a HUC12 scale (U.S. Geological Survey, National Geospatial Technical Operations Center, 2023).

320 We considered five land use type categories that were aggregations of two or more National Land Cover Database (De-
witz and U.S. Geological Survey, 2021) categories. The ‘desert’ category was composed of barren land (NLCD code=31),
shrub/scrub (52), and grasslands/herbaceous (71) land classes. The ‘forest’ category was composed of evergreen, deciduous
and mixed forests (41-43). The ‘developed’ category was composed of all the ‘developed’ classes, including open (21-24). The
‘agriculture’ category was composed of pasture/hay (81) and cultivated crops (82). The final category, ‘water and wetlands’
325 included all other land types, which include open water (11), perennial ice/snow (12), woody wetlands (90) and emergent
herbaceous wetlands (95). We assigned the dominant land use/land cover category for each HUC12 using data based on the
land use type with the greatest fractional coverage.

We compiled groundwater stable isotope ($\delta^{18}O$, δ^2H) measurements from the USGS NWIS (U.S. Geological Survey, 2022),
and published datasets assimilated in the Water Isotopes Database (Putman and Bowen, 2019). We did not place temporal or
330 well depth constraints on the samples used in our analysis. Not imposing well depth constraints may contribute to scatter associ-
ated with differences in water sources recharging shallow groundwater compared to deeper confined aquifers. The groundwater
isotope ratio observations were spatially joined to the hydrologic units.

2.8.2 Evaluation of NWM groundwater discharge with well level fractions

The Jasechko et al. (2021) dataset compared river surface elevations with river-side well water elevations within catchments.
335 The approach produced the fraction of wells in a catchment whose water surface levels were lower than the water surface level
of the nearby river. In catchments where most well water levels are below the river water level (scores close to 1), we expect
the river to lose water to shallow groundwater recharge under the right geologic conditions (e.g., permeability). In contrast, in
catchments where most well water levels are above the river water level (scores close to 0), we expect groundwater discharge
to streams.

340 We predicted the long term mean summer NWM ‘qBucket’ magnitude using the Jasechko et al. (2021) dataset using a simple
linear regression. This approach tests the hypothesis that if NWM accurately represents groundwater discharge to streams, the
relationship of well water elevations to river surface elevation would predict the summer mean NWM groundwater discharge
flux (assuming a linear relationship between the two quantities), with some scatter to account for subsurface permeability and
spatial variability in groundwater discharge rates. We then evaluated the effect of agricultural irrigation in a catchment on the
345 relationship between NWM ‘qBucket’ (binned by to the 0-20th, 20th-40th, 40th-60th, 60th-80th, and 80th-100th percentiles)
and the Jasechko et al. (2021) dataset. The evaluation was split into reaches influenced by irrigation sourced from groundwater



and irrigation sourced from surface water, as well as reaches uninfluenced by irrigation water. Irrigation contributions and irrigation water sources were determined using the methods for estimating irrigation water use described in Text S5.1.

3 Results and discussion

350 3.1 Evaluation of the isotope mass balance approach for estimating surface water isotope ratios

The analysis evaluated 4503 stream stable isotope observations in 877 unique river reaches across the western United States relative to NWM-driven isotope mass-balance-derived estimates (hereafter, ‘modeled’) of the river isotope ratios. Of these, 448 reaches had more than one observation (often all at the same sampling site in the catchment, but sometimes at multiple sites, Figure S1), and up to 571 observations in a catchment (Figure S1 and S2). On average, across all data, the observations
355 were significantly greater than the modeled values by $0.537 \pm 0.033 \text{ ‰}$ and $4.81 \pm 0.222 \text{ ‰}$ for $\delta^{18}\text{O}$ and $\delta^2\text{H}$, respectively (Figure 3). For $\delta^{18}\text{O}$ we observed a standard deviation of 3.16‰ for the observed data and 2.96‰ for the modeled data (for all data averaged by catchment). For $\delta^2\text{H}$ we observed a sample standard deviation of 25.4‰ for the observed data and 24.4‰ for the modeled data (for all data averaged by catchment, Figure 3).

The modeled isotope ratios and observed isotope ratios were well correlated (Table 1, Figures S4-7), with correlation co-
360 efficients between 0.761 and 0.866, depending on the isotopologue and whether individual observations or catchment means were considered. These correlations translated to statistically significant simple linear regressions where the modeled isotope ratios were used to explain the observed isotope ratios (Table 1). Depending on the isotopologue and whether individual observations or means were considered, the model explained between $\sim 58\%$ and 75% of the variance in the observations. The model explained more variance for $\delta^2\text{H}$ than $\delta^{18}\text{O}$, and more variance for catchment mean values relative to individual
365 observations. For all regressions, the slopes ranged from 0.879 to 0.937, with catchment mean slopes tending to be lower than slopes calculated from all observations. Intercepts for all regressions were close to, but less than 0, with lower intercepts associated with regressions calculated from catchment mean values, relative to regressions calculated from all observations. The statistically significant slopes of less than 1 and statistically significant intercepts arise in all observation-model comparison
370 regressions because the observations tended to exhibit higher isotope ratios than the model estimated at the lower end of the isotopic distribution (Figures S4-7). Many of the catchments characterized by this pattern were in more arid locations.

We calculated surface water lines (SWL, similar to a meteoric water line (MWL), but instead of being calculated from precipitation, they are calculated with isotope ratios from surface water samples in an area) for both the modeled and observed results using all available data (Figure 3). The observations yielded a surface water line with a slope of $7.570 (\pm 0.023)$, and intercept of $1.2301 (\pm 0.320)$, which was significantly different from the slope of the GMWL slope of 8 and intercept of 10
375 and was within the range of local MWLs (LMWL) slopes for western North America (6.5-8) (Putman et al., 2019), reported in Table 2. The model results yielded a surface water line $8.12 (\pm 0.010)$ and an intercept of $8.06 (\pm 0.14)$ which was more similar to, but still statistically different from the GMWL and differed from LMWLs for the region (Table 2). Comparison of the observation and modeled data distributions and water lines reveals evidence for evapoconcentration of surface waters in the observations but not in the isotope mass balance results (Figure 3). This is because the primary source of streamflow in

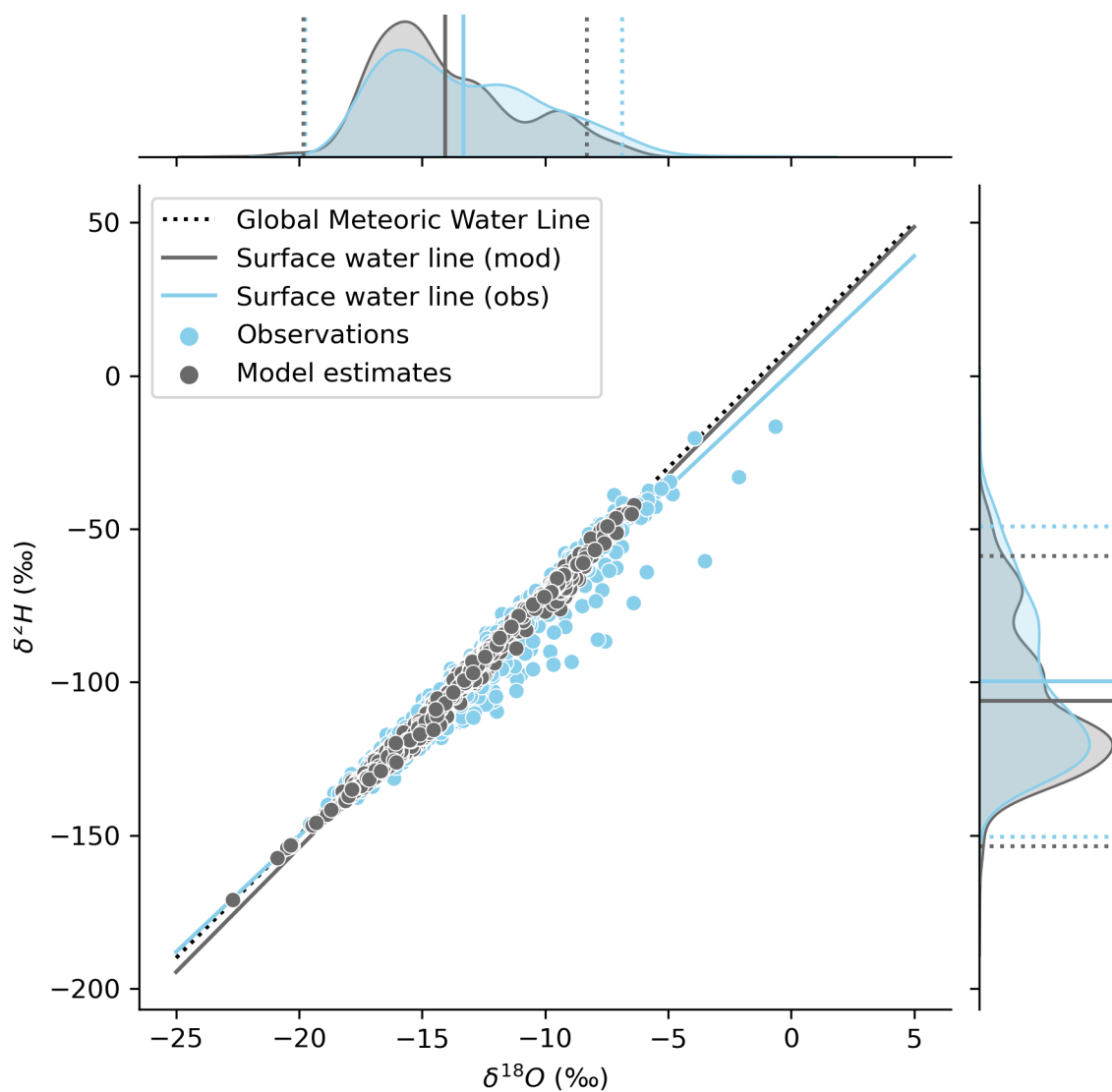


Figure 3. The distribution of the catchment mean observation (obs, blue) and isotope mass balance estimates (mod, gray) ($n=448$) with the Global Meteoric Water line (dotted) and the two datasets surface water lines (solid lines). Data distributions, including mean and two standard deviations of each data type (dotted lines), are shown in the plot margins. Observations plotting below the GMWL indicate evapoc-
oncentration (Putman et al., 2019).



| Statistical model | n | Corr. coef | β (\pm s.e.) | I (\pm s.e.) | R ² |
|--|------|---------------|-----------------------|------------------------|----------------|
| $\delta^{18}O_{obs} \sim \delta^{18}O_{mod} + I$ | 4503 | 0.761 | 0.917 (\pm 0.012)* | -0.645 (\pm 0.168)* | 57.9% |
| $\delta^{18}O_{obs,avg} \sim \delta^{18}O_{mod,avg} + I$ | 448 | 0.820 | 0.879 (\pm 0.029)* | -0.891 (\pm 0.414)* | 67.3% |
| $\delta^2H_{obs} \sim \delta^2H_{mod} + I$ | 4503 | 0.819 | 0.937 (\pm 0.010)* | -1.90 (\pm 1.06)* | 67.1% |
| $\delta^2H_{obs,avg} \sim \delta^2H_{mod,avg} + I$ | 448 | 0.866 | 0.905 (\pm 0.025)* | -3.10 (\pm 2.66) | 75.1% |
| $\delta^2H_{diff} \sim \delta^{18}O_{diff} + I$ | 4503 | 0.959 | 6.54 (\pm 0.029)* | 1.30 (\pm 0.065)* | 91.9% |
| $\delta^2H_{avg,diff} \sim \delta^{18}O_{avg,diff} + I$ | 448 | 0.958 | 6.70 (\pm 0.094)* | 1.46 (\pm 0.190)* | 91.9% |

Table 1. Correlation and regression results for observation-model comparisons. Regressions were performed on all data (n = 4503), as well as on the mean values in a subset of the reaches with more than one observation (n=448). An asterisk (*) indicates the coefficient is significant at $p < 0.1$.

380 the modeling framework, high elevation groundwater discharge, does not bear an evapoconcentrated isotopic signature in our input datasets, and lower elevation water sources (groundwater or runoff) that could bear a signature of evapoconcentration, depending on the region, are considered by the model to be minor contributors to streamflow over the timescale integrated by our study.

3.2 Model-observation differences

385 Of 4503 observations, 1763 $\delta^{18}O$ and 3306 δ^2H observations were significantly different from the long term mean isotope mass balance NWM estimate at $p < 0.1$. Of these, 1756 observations indicated significant differences in both isotopologues. This corresponded to a median absolute difference of 2.2‰ for $\delta^{18}O$ and 9.7‰ for δ^2H . For both, a larger proportion of the distribution indicated positive significant differences and those differences tended to be greater in absolute magnitude than the negative significant differences.

390 We used an observation-model difference interpretation framework (Figure 2) to interpret differences between the observations and the isotope mass balance estimates. The differences may contain process information that can be used to improve our understanding of terrestrial water balance and process inclusion in the NWM. The observation-model differences in $\delta^{18}O$ and δ^2H were correlated (Figure 4) and yielded similar results for analyses performed with all data as compared to means of reaches with multiple observations (Table 1). Correlations between $\delta^{18}O_{diff}$ and δ^2H_{diff} were about 0.96. Simple linear regressions, where variance in $\delta^{18}O_{diff}$ explained variance in δ^2H_{diff} , with all data and catchment mean data both explained about 92% of the variance, were significant and exhibited slopes of less than 8 (Table 1), suggesting the presence of errors arising from NWM omission of water sources that bear signatures of non-equilibrium processes.

In our dataset, model estimates do not deviate much from the GMWL, and deviate less than the observations (Figure 3). This information is quantified using d_{diff} (Figure 2). Positive values of $\delta^{18}O_{diff}$ tended to be associated with negative values of d_{diff} (Figure S8). The shape of the relationship between the two quantities is non-linear, with a stronger relationship between $\delta^{18}O_{diff}$ and d_{diff} among data from arid reaches compared to humid reaches.

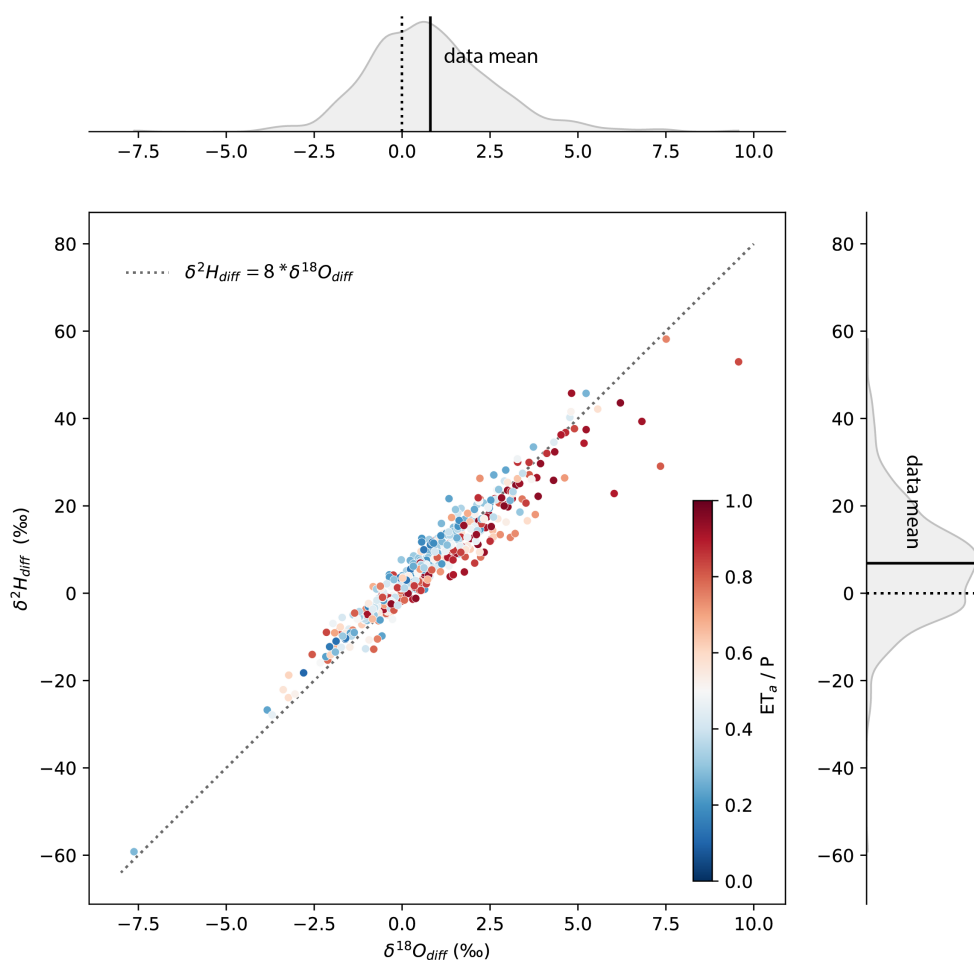


Figure 4. The relationship of observation (obs) - isotope mass balance (mod) estimation differences for $\delta^{18}O$ and δ^2H . The catchment mean value is plotted, and only sites with at least two observations are shown (n=448). The equilibrium line with slope 8 is plotted for context (dotted line), and data are colored by their site's the ratio of actual evaporation to precipitation. Data distributions are shown for both $\delta^{18}O_{diff}$ and δ^2H_{diff} in the margins, along with the mean differences indicated as a solid line. No difference (0) is marked with a dotted line for reference.



| Surface water lines: | β (\pm s.e.) | I (\pm s.e.) |
|---------------------------------------|--|----------------------------------|
| Model-derived | 8.12 (\pm 0.010) | 8.06 (\pm 0.14) |
| Observations | 7.57 (\pm 0.02) | 1.23 (\pm 0.32) |
| Meteoric water lines: | β_{min}, β_{max} (β_{avg}) | I_{min}, I_{max} (I_{avg}) |
| Global Meteoric Water Line | 8 | 10 |
| Arid and Temperate dry summer LMWLs | 6.56, 8.02 (7.57) | -10.5, 9.85 (3.02) |
| Temperate humid and Continental LMWLs | 7.34, 7.64 (7.49) | -3.82, 3.31 (0.62) |

Table 2. Surface water line slopes and intercepts ($\delta^2H = \delta^{18}O + I$) compared to the Global Meteoric Water Line and published precipitation water line ranges (LMWLs) from different climate classifications in North America (data from Putman et al. (2019)).

The relationship between $\delta^{18}O_{diff}$ and d_{diff} , as well as our regression (Table 1) and surface water line analyses (Table 2) indicate that the modeling approach for estimating long term isotope ratios of rivers produce results that are similar to, but on average, lower and exhibit less variability than observations. The strongest signal in our data is that of evaporation, evidenced by combinations of positive $\delta^{18}O_{diff}$ and negative d_{diff} in arid regions. We also observe evidence of non-equilibrium condensation processes in reaches characterized by negative $\delta^{18}O_{diff}$ and positive d_{diff} .

We suggest that patterns in $\delta^{18}O_{diff}$ and d_{diff} contain useful model diagnostic information that can be useful for improving the NWM and our understanding of the terrestrial water balance. The modes of variability we evaluate include 1) interannual 2) monthly and 3) spatial. The first is to evaluate the effect of comparing long term mean mass balance estimates with sample collection. The second and third modes of variability provide information about the presence of possible missing water sources to the model. Additional sources of variability are discussed in Text S6.

3.3 Interannual variability in observation-model differences

The annual mean $\delta^{18}O_{diff}$ and d_{diff} values for NRSA collection years were statistically indistinguishable from the mean differences calculated with the whole dataset (Figure 5). Though all observation-model difference regression slopes were less than 8, and all intercepts were greater than 0, the observation-model difference regression slopes and intercepts were significantly different from one another and the regression calculated from all available data.

We interpret the year to year similarity in the regression slopes and intercepts as indicating the pervasive presence of evapoconcentrated observations in the region. We interpret the variability in the slopes and intercepts as arising from interannual differences in climate. Such interannual differences may include variability in snowpack isotope ratios and the proportion of runoff from rain or snowmelt in the stream relative to groundwater. We suggest that the year to year variability in slopes and intercepts is more likely to derive from climatic differences than differences in sampling distribution.

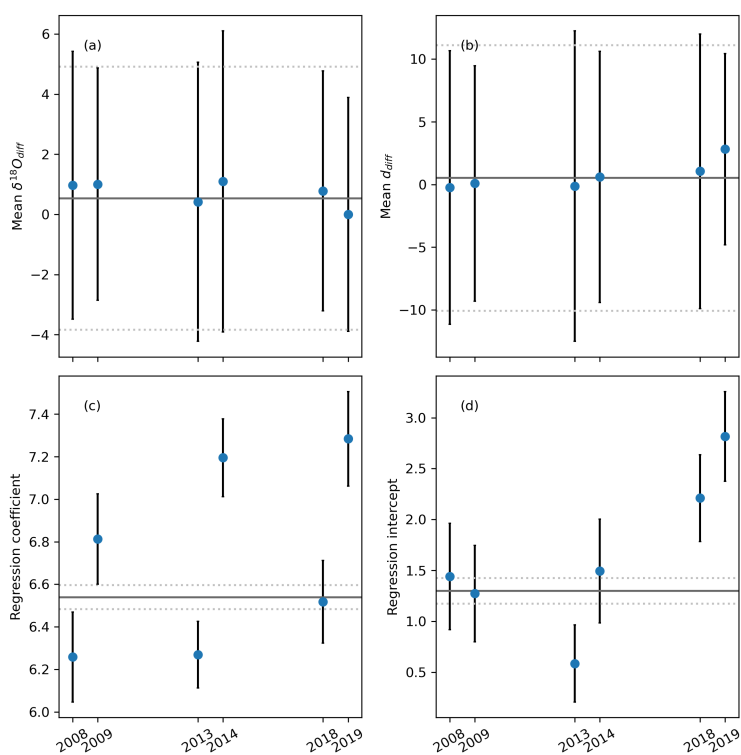


Figure 5. Interannual variability in observation-model comparisons using data from years with at least 250 samples in 100 different reaches (NRSA years). (a) The annual mean (blue points) and 95% confidence intervals (solid vertical lines) of $\delta^{18}O_{diff}$ compared to the mean (solid horizontal line) and 95% confidence intervals (dotted horizontal lines) of all data. (b) The annual mean and 95% confidence intervals of d_{diff} compared to the mean and 95% confidence intervals of all data. (c) The slope of $\delta^2 H_{diff} \sim \delta^{18}O_{diff}$ and 95% confidence intervals compared to the slope and 95% confidence intervals of calculated from all data. (d) The regression intercept of $\delta^2 H_{diff} \sim \delta^{18}O_{diff}$ and 95% confidence intervals compared to the regression intercept and 95% confidence intervals of calculated from all data. In all subplots, the all-data mean and 95% confidence interval are plotted for reference



3.4 Seasonal variability in observation-model differences

There are systematic patterns in $\delta^{18}O_{diff}$ and d_{diff} when examined across the growing season. For example, $\delta^{18}O_{diff}$ tends to be greater during the latter months of the growing season relative to the mean $\delta^{18}O_{diff}$ value for the month of June for that site and year (Figure 6 a) in most basins and months. The pattern is especially evident in the Great Basin and California. Likewise, d_{diff} , is lower in July, August, September relative to June (Figure 6 b), but only in the Great Basin and California. The contrast between basins with both increased $\delta^{18}O_{diff}$ and decreased d_{diff} (Great Basin and California) and those with only increased $\delta^{18}O_{diff}$ and little change in d_{diff} (Upper and Lower Colorado and Pacific Northwest) suggests that two different mechanisms may drive isotopic change during the growing season.

In California and the Great Basin, which are characterized by larger ^{18}O -enrichment and d decrease over the growing season, we suggest increased contributions of evapoconcentrated waters to rivers later in the growing season. In California, this may reflect the water use in the Central Valley.

In the Upper and Lower Colorado, and Pacific Northwest, where we observe small ^{18}O enrichment in the absence of a notable change in d the second case, we suggest sustained dependence on the same water source regions throughout the growing season. In the Pacific Northwest and parts of the Upper Colorado, the relative invariability may indicate the sustained dependence on groundwater discharge from high elevations to streamflow during the growing season (Miller et al., 2016; McGill et al., 2021; Windler et al., 2021). In lower parts of the Upper Colorado and the Lower Colorado, where rivers are characterized by large reservoirs, the seasonal invariance may reflect that the primary ‘water source’ regions for these reaches are reservoirs, which retain snowmelt from early in the season and discharge it later in the season. In this way, the primary ‘water source’ for the streams in many areas of these basins may remain constant throughout the growing season.

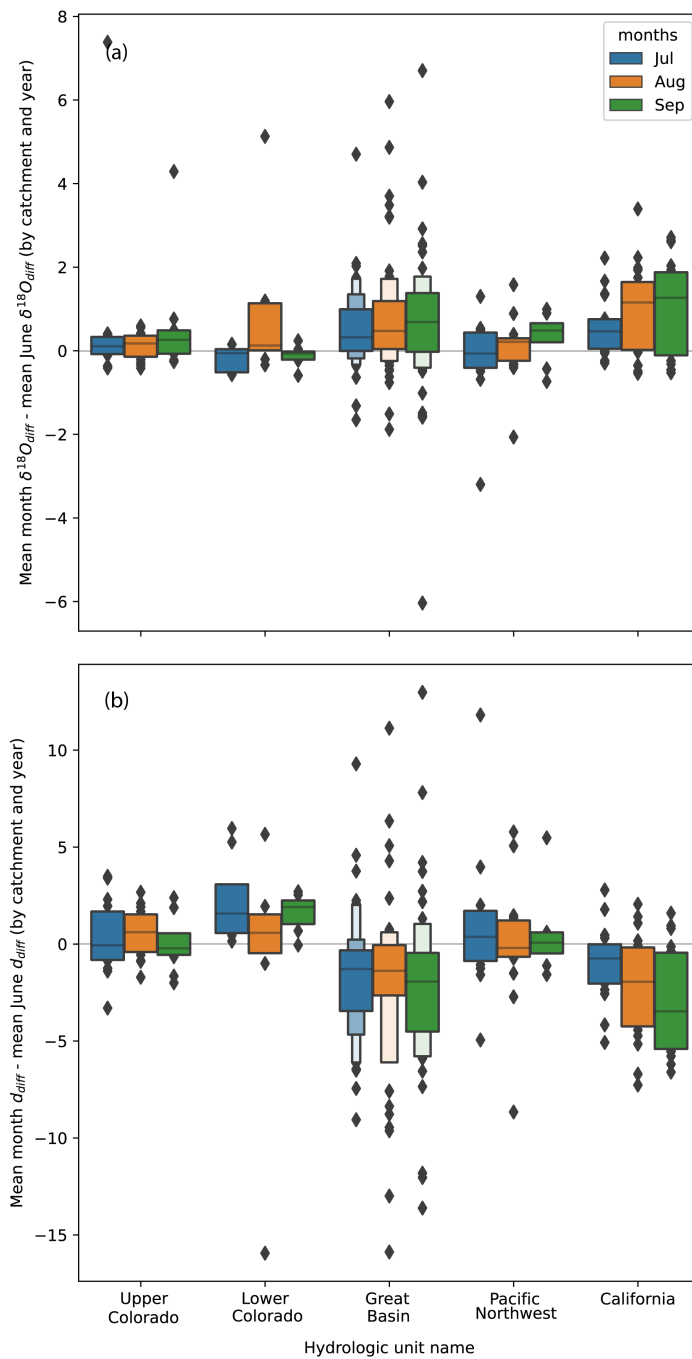


Figure 6. Evaluation of seasonal variability in observation-model comparisons. Data include all reaches and years with collections in the month of June as well as two of the 3 other months of the summer season. (a) The distribution of month-specific differences from June $\delta^{18}O_{diff}$ by basin. (b) The distribution of month-specific differences from June d_{diff} .



3.5 Spatial distribution of observation - model differences

If the NWM fully constrained all relevant water sources, we expect to observe similar values of $\delta^{18}O_{diff}$ and d_{diff} throughout each basin, irrespective of the location of the observation in the basin. Instead, we observed spatial variability (Figures 7 and S9), where smaller magnitude $\delta^{18}O_{diff}$ values occurred in the highest elevation, lowest stream order, and least arid reaches, and larger magnitude, often positive $\delta^{18}O_{diff}$, values occurred in lower elevation, arid or intermittent flow reaches (Figure S10). d_{diff} tended to exhibit higher values in higher elevation, lower stream order reaches, and lower values in lower elevation, more arid, higher stream order reaches (Figure 8). We observed a greater range in the absolute magnitudes of $\delta^{18}O_{diff}$ and d_{diff} in higher order, lower elevation reaches (Figures 8 and S10). Notably, the pattern was similar across basins, suggesting the importance of within-basin processes in determining $\delta^{18}O_{diff}$ and d_{diff} , as opposed to absolute relationships of $\delta^{18}O_{diff}$ and d_{diff} to elevation, stream order, or climate classification.

The spatial pattern in d_{diff} (Figure 7) was similar to the pattern observed for the KGE and other metric evaluations of the NWM (Towler et al., 2023). Areas with negative d_{diff} tended to correspond to areas with poor NWM performance (Towler et al., 2023). However, the isotopic evaluation of NWM and the Towler et al. (2023) datasets could not be directly compared due to there being only a small number of reaches with both isotope observations and daily discharge measurements.

The spatial structure of our results was statistically well explained by the ratio of actual evaporation to precipitation ($\frac{ET_a}{P}$) in a linear mixed effects model with basin as the grouping variable (Table 3). Variability among basins explained 16.2% of the variance in d_{diff} , while the fixed effect of aridity explained 13.9% of the variability in the dataset. The regression slope associated with the fixed effects of aridity was negative (-7.87 ± 0.78) and significant ($p < 0.001$), indicating that sites with higher aridity indices tended to exhibit more negative d_{diff} . This regression was stronger than a linear mixed effects model with elevation predicting d_{diff} , where the fixed effects of elevation explained 4.7% of the variability in d_{diff} .

Analysis of the spatial variability in our results suggest that 1) higher elevation, lower stream order, perennial, warm temperate or seasonally snowy reaches had small $\delta^{18}O_{diff}$ and positive d_{diff} values and 2) lower elevation, higher stream order, arid and sometimes intermittent stream reaches had larger and more positive $\delta^{18}O_{diff}$ values and more negative d_{diff} values. The first point suggests errors associated with the challenges of providing input values at appropriate temporal resolutions, whereas the second point suggests the model is missing critical evapoconcentrated water sources in more arid, lower elevation areas of each basin.

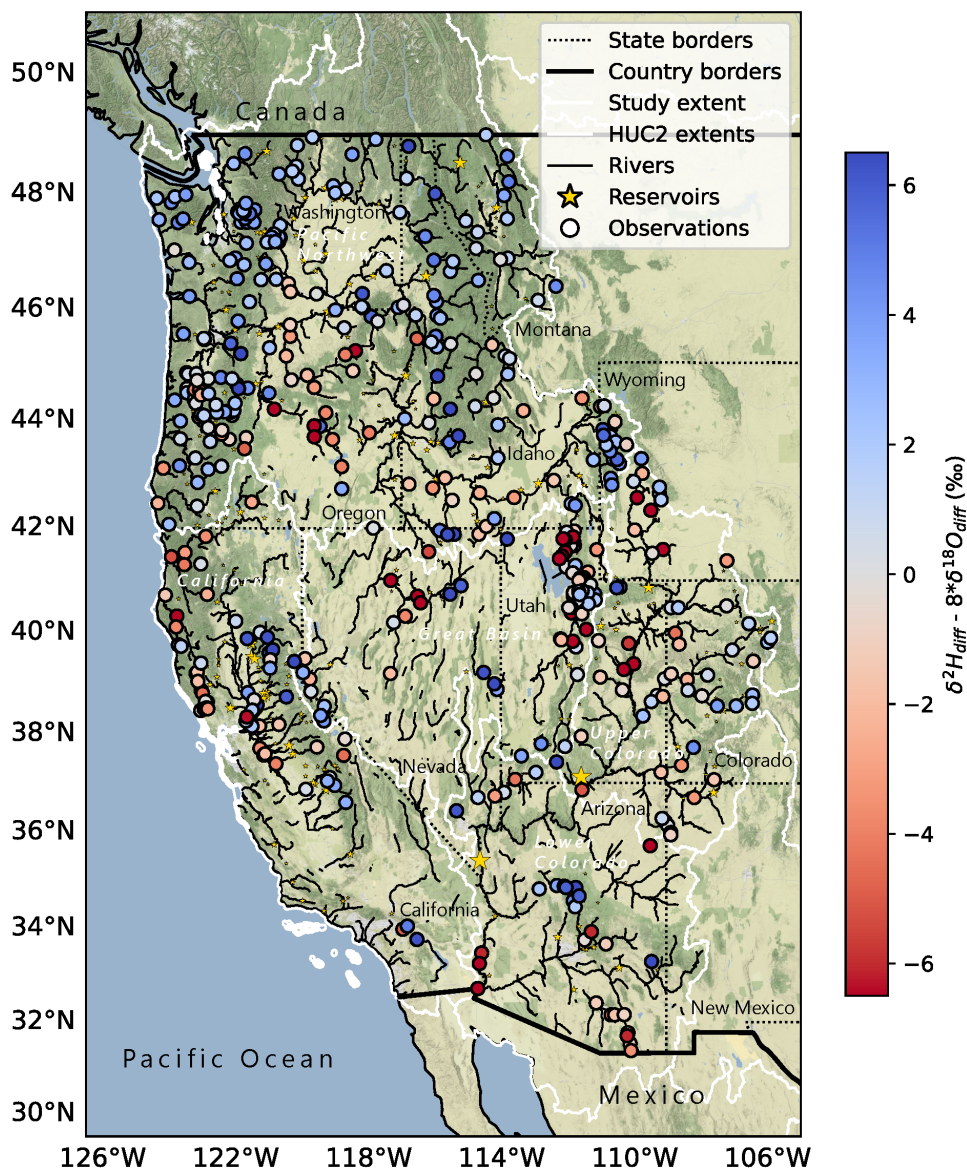


Figure 7. The spatial distribution of mean catchment d_{diff} ($\delta^2 H_{diff} - 8 * \delta^{18} O_{diff}$) in reaches with more than one observation (n=448). Locations of reservoirs are marked by yellow stars, with the star size proportional to the reservoir capacity. Redder colors indicate more evapoconcentrated waters than expected based on the model estimate. Map data is from ©OpenStreetMap contributors 2023, distributed under the Open Data Commons Open Database License (ODbL) v1.0, accessed through Stamen OpenSource Tools (<https://stamen.com/open-source/>). HUC2 basins come from WBD (U.S. Geological Survey, National Geospatial Technical Operations Center, 2023) and rivers are modified from the NHDPlus streamline network (U.S. Geological Survey, 2019).

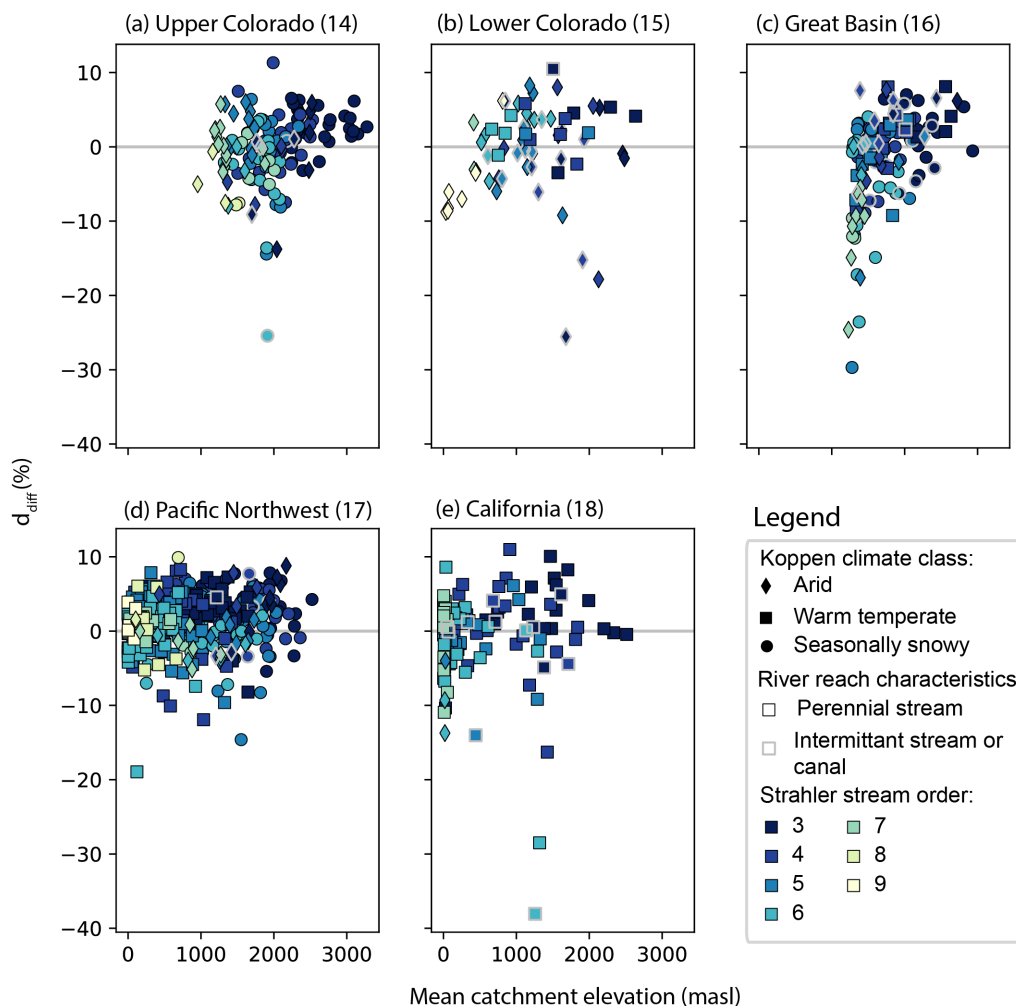


Figure 8. Relationship of elevation, Strahler stream order, and Köppen climate classification (Rubel and Kottek, 2010), and stream persistence to d_{diff} in each basin. We observe higher d_{diff} in perennial, lower order streams at middle and higher elevations in each basin. Lower d_{diff} is associated with higher order streams at lower elevations in each basin. This effect was greater in catchments classified as arid or seasonally snowy compared to those classified as warm temperate. This pattern was generally true in each basin, irrespective of the absolute elevation or stream order, suggesting the importance of accumulated effects within a basin on d_{diff} .



3.5.1 Observation-model differences in headwater reaches reflect groundwater isotope ratio estimates

We observe statistically significant reach-scale $\delta^{18}O_{diff}$ and d_{diff} in higher elevation, low stream order, low aridity, temperate or seasonally snowy reaches in our dataset (Figures 8, S10). These differences tend to be smaller than than the full dataset mean $\delta^{18}O_{diff}$ and d_{diff} . At most of these reaches we also observe positive d_{diff} values (Figures 7, 8).

The presence of both negative and positive values of $\delta^{18}O_{diff}$ likely reflect interannual variability in the isotope ratios of actual groundwater and snowmelt discharged to rivers in high elevation headwater areas. Although groundwater's contribution to streams is conceptualized in this study to be constant in magnitude and isotope ratio, in reality, the isotope ratios of both groundwater and snowmelt fluxes vary spatially and interannually. The groundwater flux magnitudes vary interannually based on variations in snowpack magnitudes, antecedent hydrologic conditions, (Brooks et al., 2021; Wolf et al., 2023) and hydrogeologic (Gentile et al., 2023) controls including hydrologic residence times. Snowpack isotope ratios vary in response to climate patterns and local conditions (Anderson et al., 2016). The observed variability of $\delta^{18}O_{diff}$, which does not exhibit a uniform tendency towards positive or negative values, suggests the mean groundwater isotope ratios used in this study are reasonably representative of the long term mean estimates of the isotope ratios of water contributed at high elevation water source areas by groundwater and snowmelt fluxes. Though improvements may be made by using interannually varying estimates of the isotope ratios of groundwater and snowmelt. However, the systematic positive d_{diff} result cannot be explained by the timescale of the isotope input.

Higher d streamflow relative to weighted mean precipitation values have been documented in other studies (Nickolas et al., 2017). This may be because higher d is associated with lower precipitation $\delta^{18}O$ that falls during the cold season in mid-latitude regions, particularly in areas near open water (Putman et al., 2019; Corcoran et al., 2019; Aemisegger and Sjolte, 2018). Secondly, high d in rivers relative to precipitation or groundwater may be attributed to fractionation occurring during melt. The snow melt process has been demonstrated to begin with preferential melt of water molecules bearing lighter isotopologues, and to exhibit higher d earlier in the melt season (Ala-aho et al., 2017; Beria et al., 2018; Carroll et al., 2022). The higher d of the snow and initial meltwater may be passed along to the rivers via direct runoff to streams or through shallow groundwater recharge and rapid discharge to streams (see the relatively higher upper bound on d values for forested land use types in Figure 9).

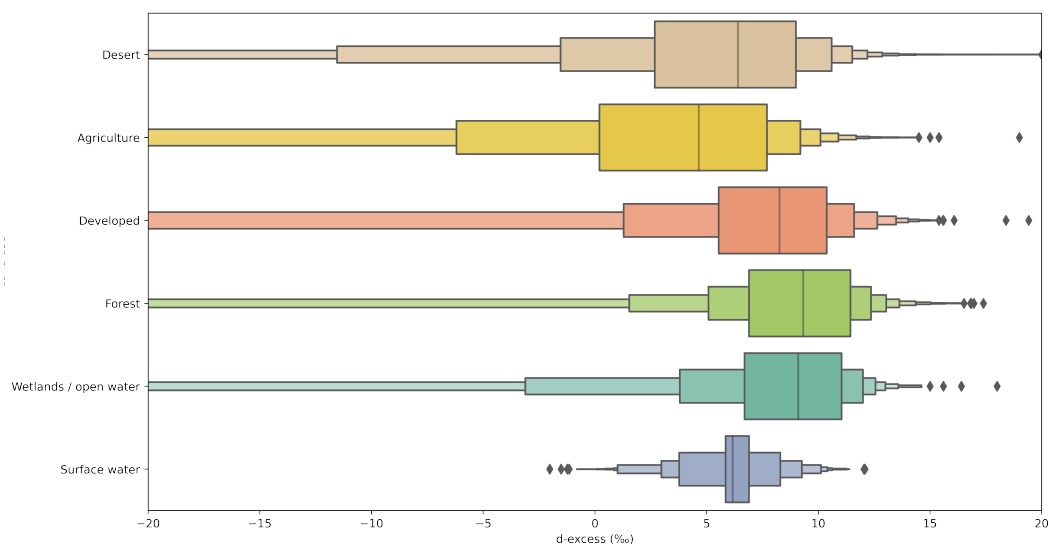


Figure 9. Distributions of groundwater d grouped by the dominant land type from NLCD (Dewitz and U.S. Geological Survey, 2021) in the HUC12 (U.S. Geological Survey, National Geospatial Technical Operations Center, 2023) of the observation. The desert land class includes barren land (often playas or dried lakebeds), shrub/scrub, grasslands/herbaceous. The agricultural land class includes pasture/hay and cultivated crops. The developed land class includes developed land of any intensity. Forest includes evergreen, deciduous and mixed forest. The wetlands/open water land class category any type of wetland as well as open water. The distribution of our 4303 river samples is also shown for context.

3.5.2 Evapoconcentration at low elevations suggests contributions of irrigation return flows to streamflow

Greater spatial and temporal variability in both $\delta^{18}O_{diff}$ and d_{diff} in lower elevation, higher stream order, arid reaches suggests the importance of various spatially and temporally heterogeneous processes and water sources that may alter streamflow isotope ratios relative to upstream values. Positive values of $\delta^{18}O_{diff}$ and negative values of d_{diff} in more arid regions of each basin suggests that evapoconcentrated waters compose a large fraction of streamflow in these areas (Figures 7, 8, S9 and S10), especially in the later part of the growing season (Figure 6) when streams depend more heavily on groundwater fluxes. We observed evidence of surface water evapoconcentration in all basins (Figure 8), though it was most apparent in Lower Colorado River Basin, lower elevation regions of the Upper Colorado River Basin, California's Central Valley, near Great Salt Lake in the Great Basin and throughout the Snake River Plain (Figures 7 and S9).

The isotope ratios and d we observe in low elevation, high stream order arid reaches are similar to those we would expect to observe in highly evaporative contexts, like within lakes (Bowen et al., 2018), intermittent flow rivers, or downstream of wetlands. Yet the majority of rivers in our study are perennial, and most are not characterized by substantial wetlands.

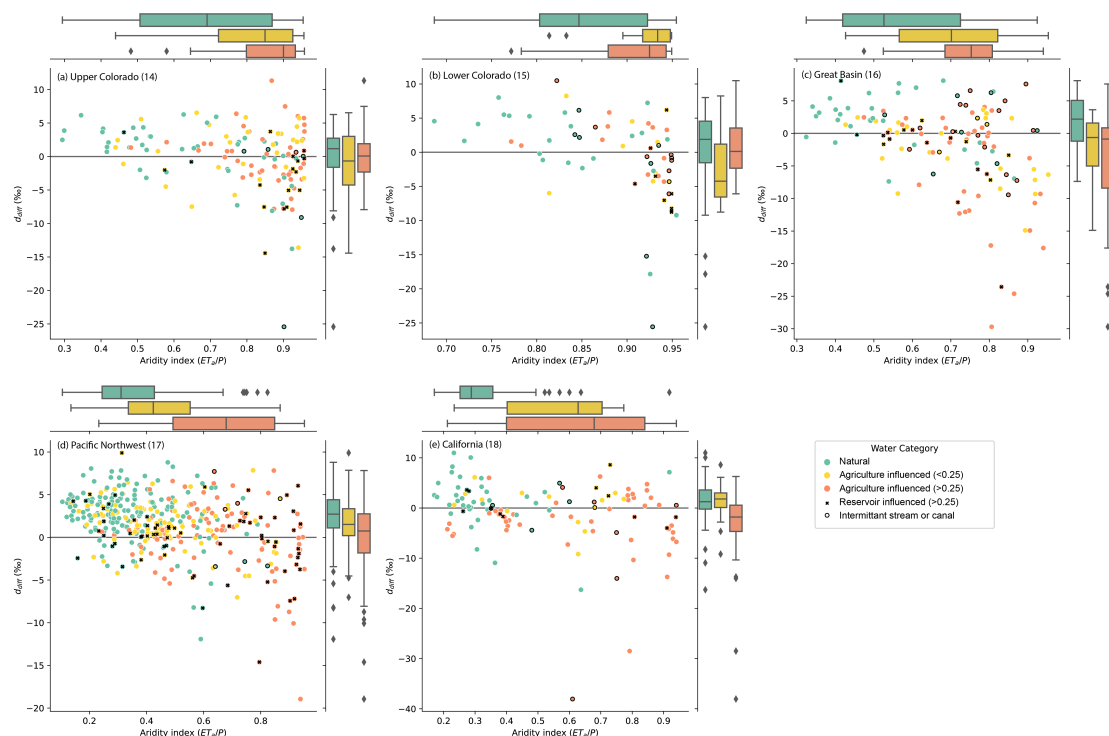


Figure 10. Relationship of aridity to d_{diff} , by water use categories and basins. Natural waters are not estimated to be influenced by agricultural irrigation. Fractions of agricultural irrigation contributing to streamflow are estimated using water use data and land cover data and do not account for losses to evapotranspiration. Reaches affected by large reservoirs and reaches categorized as intermittent or as canals or ditches are indicated.

The evapoconcentration in our dataset is unlikely to arise from river or reservoir evaporation because both evaporation of reservoirs and evaporation to inflow ratios in the region tend to be low, especially for deep man-made reservoirs (Brooks et al., 2014; Friedrich et al., 2018). Instead, isotopic evidence of evapoconcentration occurs in waterways likely to be affected by anthropogenic hydrologic alteration (Fergus et al., 2021) and characterized by larger fractions of ‘young water’ (Jasechko et al., 2014; Burt et al., 2023; Xia et al., 2023).

We tested the hypothesis that the spatial pattern in isotopically-inferred evapoconcentration could arise from contributions of irrigation return flows to streams and reservoir releases. Within each basin, on mean, d_{diff} was most negative at sites with the highest proportion of total inflows attributed to agricultural return flows and highest at sites with no apparent contributions of agricultural return flows (Figure 10). Reservoir influence was associated with low d_{diff} more often where dams are used for water management and water supply (e.g., Upper Colorado, Lower Colorado, Great Basin, and California) and were associated with high d_{diff} in the Pacific Northwest, where dams are more often used for hydropower. Intermittent streams and canals in arid regions were sometimes associated with low d_{diff} as well, even when no water was contributed by agricultural irrigation.



We demonstrated the relationships of agricultural and reservoir influence on d_{diff} statistically in a linear mixed effects model (Table 3). The fraction of streamflow estimated to come from agricultural irrigation return flows and a categorical variable delineating reservoir influence together explained 7.7% of the variance in d_{diff} , with the whole model (including random group effects) explaining 14.5% of the variance in the dataset. Both explanatory variables were significant ($p < 0.001$), and, as
520 expected, exhibited negative slopes indicating that greater agriculture and reservoir influences tended to produce lower, more evapoconcentrated d_{diff} values. When we also include the ratio of actual evaporation to precipitation with these explanatory variables, all three are significant ($p < 0.01$) and explain 15.1% of the variance through fixed effects, and 23.4% of the variance overall (fixed and random effects). Among the linear mixed effects models tested, it exhibited the highest log likelihood value, explained the greatest amount of variance using fixed effects, and reduced the amount of variance attributed to random within-
525 basin effects.

While this statistical model performance is not substantially better at explaining variance in d_{diff} than the model that uses aridity alone, the findings do suggest that both agricultural activity and reservoirs influence the isotope ratios of streamflows across the Western US. The low variance explained by these models is expected, due to the difficulty estimating a true long term mean agricultural return fluxes with the available data and the potential for isotopically heterogeneous reservoir effects, as well
530 as the covariance of both irrigation return flows and the presence of reservoirs with aridity and elevation. However, our findings are well supported by studies of irrigation-based contributions to streamflow, an analysis of groundwater evapoconcentration by land use, comparison of the Jasechko et al. (2021) dataset with NWM groundwater contributions to streams in the context of patterns in agricultural irrigation.



| Statistical model | β (\pm s.e.) | I (\pm s.e.) | HUC2 Var | Cond. R ² | Fixed R ² | Log likelihood |
|--|--|-----------------|----------|----------------------|----------------------|----------------|
| $d_{diff} \sim Elev + I$ | <i>Elev</i> : 0.001 (0.00)* | -1.93 (1.01) | 4.33 | 20.9% | 4.4% | -2254 |
| $d_{diff} \sim \frac{ET}{P} + I$ | $\frac{ET}{P}$: -7.85 (0.77)* | 4.86 (1.08)* | 4.37 | 30.2% | 13.9% | -2209 |
| $d_{diff} \sim F_{irr} + Res + I$ | <i>F_{irr}</i> : -3.49 (0.49)* <i>Res</i> : -1.76 (0.45)* | 0.92 (0.59) | 1.464 | 14.3% | 7.9% | -2225 |
| $d_{diff} \sim \frac{ET}{P} + F_{irr} + I$ | $\frac{ET}{P}$: -6.53 (0.83)* <i>F_{irr}</i> : -1.61 (0.54)* | 4.40 (0.35)* | 1.942 | 22.7% | 14.7% | -2203 |
| $d_{diff} \sim \frac{ET}{P} + F_{irr} + Res + I$ | $\frac{ET}{P}$: -6.11 (0.88)* <i>F_{irr}</i> : -1.68 (0.54)* <i>Res</i> : -1.22 (0.44)* | 4.33 (0.82)* | 1.861 | 23.0% | 15.2% | -2199 |

Table 3. Results of linear mixed effects models with 764 observations and 5 groups. The minimum and maximum group sizes were 48 and 387, respectively. The models do not include any samples from reaches characterized as an intermittent stream or canal or where NWM indicates that the maximum streamflow is $0 \text{ m}^3 \text{ s}^{-1}$. Random effects apply only to the intercepts. An asterisk indicates that a regression coefficient is statistically significant at $p < 0.01$. Conditional R², which gives the total model variance explained, are reported alongside the fixed R², which gives the variance explained by fixed effects (i.e., explanatory variables) and the log-likelihood, which can be used to evaluate the relative performance of different models.

3.5.3 Literature and datasets support isotopic inference of irrigation return flows contributing to streamflow

535 Numerous prior studies have investigated the influence of irrigation on streamflow. Estimates suggest that, depending on the irrigation type, as much as 50% of applied water may recharge groundwater or arrive at surface waters through shallow groundwater infiltration and subsequent discharge to streams (Grafton et al., 2018). Likewise, irrigation has been demonstrated to increase streamflows during low flow periods (Fillo et al., 2021; Essaid and Caldwell, 2017), if the applied water comes from surface water diversions.

540 Local contributions of groundwater to streams from irrigation-based recharge are supported by the d values of groundwater in agricultural regions. Groundwater from regions influenced by agricultural irrigation exhibited lower mean d relative to deserts, including dried terminal lakes and playas, developed areas which may include turf grass irrigation, forested regions, wetlands or open waters and surface waters (Figure 9). Some contribution of this irrigation-recharged groundwater to streams via return flows would decrease the difference between the modeled and observed isotope ratios in our dataset, and is supported
 545 by conclusions from prior isotopic inference of water sources in the Snake River plain (Windler et al., 2021).

The isotopic inference that irrigation return flows are an important missing process in the NWM is supported by an independent statistical comparison of the NWM groundwater discharge with the Jasechko et al. (2021) well water level comparison to stream level dataset and the agricultural water use data. We hypothesize that if NWM accurately represents groundwater discharge to streams, the Jasechko et al. (2021) well water level comparison to stream water level dataset should be able to predict the summer mean NWM groundwater discharge flux with a large proportion of variance explained. However, the Jasechko
 550



et al. (2021) data (expressed as the fraction of well water levels that lie below the proximal river water level) weakly, though significantly ($R^2 = 0.028$, $p < 0.001$) predict the NWM groundwater discharge rates in a simple linear regression. The regression relationship between the variables is negative, as expected, where river reaches with a greater proportion of their well water levels above proximal river water levels correspond to reaches with greater groundwater discharge fluxes (Figure S11). Though
555 the regression is significant, it has almost no predictive capacity, contrary to what we expect if the well water level comparison to stream level dataset was a good predictor of the NWM groundwater discharge to streams.

The weakness of the statistical relationship between the well water level comparison to river water level and the NWM groundwater discharge flux may be related to shallow aquifers, which are not considered by NWM, and/or agricultural irrigation, as well as the water source (surface or ground water) used for that irrigation (Figure S12). We did not assess the potential
560 for NWM groundwater discharge to reflect the presence of shallow aquifers. However, we do observe that the influence of irrigation on groundwater levels is non-stationary, depending on both the groundwater discharge level as well as the source of irrigation water. For this reason the relationship is difficult to assess statistically. In river reaches where NWM indicates little groundwater discharge (0^{th} to 20^{th} percentile qBucket), irrigation sourced from surface water is associated with a larger fraction of well water levels above river water level in a catchment than those without irrigation. Conversely, in river reaches with
565 substantial groundwater discharge (80^{th} to 100^{th} percentile qBucket), agricultural irrigation with water from either surface or groundwater tends to be associated with a smaller fraction of well water levels above river level in the catchment compared to reaches without any agricultural irrigation. Based on these patterns we suggest that in dry areas, irrigation from surface water appears to contribute to groundwater recharge, whereas in wet areas, irrigation appears to contribute to decreased water table elevations. Surface water irrigation tends to contribute to higher water tables, whereas irrigation from groundwater tends to
570 contribute to lower water tables.

Some part of this signal is regional. Reaches from more arid basins compose a greater proportion of the lower percentile qBucket reaches, and reaches from humid or seasonally snowy basins compose a greater proportion of the higher percentile qBucket reaches. However, when evaluated by basin, the relationships are similar. The finding is consistent with modeling studies, which showed lower stream discharge when irrigation water came from groundwater, and greater stream discharge
575 when irrigation water came from surface water (Essaid and Caldwell, 2017). Our analysis suggests that agricultural irrigation is likely to influence groundwater levels and groundwater discharge on a landscape scale and produces gaining streams and contributes to streamflow in otherwise arid, losing reaches of rivers.

3.6 Implications of including irrigation return flows into NWM calculations

Our evaluation of the NWM-driven isotope mass balance calculations suggest that the NWM accuracy would be improved
580 by including agricultural return flows in the water sources sustaining streamflow in the NWM. In effect, agricultural return flows are simply groundwater fluxes to streams that occur at lower elevations than the majority of the groundwater discharge sustaining streams. Based on magnitudes of d_{diff} , these lower elevation groundwater fluxes can sometimes be large. Because the NWM is calibrated to actual streamflows which contain these return flows, these fluxes are currently being misallocated in the model. Inaccuracies in any model terms or fluxes influence the model's capacity to project accurate streamflows, particularly



585 under non-stationary hydrologic conditions. Accurate model water source inclusion, particularly at low elevations where water
use and availability is most critical, thus has implications for the model's utility to stakeholders, including water managers and
users.

Under current conditions, agricultural return flows may be critical for sustaining streamflow late in the growing season (Au-
gust or September) or during drought periods. Sustained streamflow in certain reaches is critical for 1) water access for surface
590 water diversions and 2) water availability for species' use. For example, protected fish species survival requires that waterways
meet thresholds of water quality, temperature, and depth for survival (Dibble et al., 2020). Water managers make decisions
about water allocations and reservoir releases in part to meet these habitat needs (Bruckerhoff et al., 2022). Agricultural return
flows have the capacity to help sustain streamflow (Fillo et al., 2021), but with potentially negative effects on water quality,
through agriculture-associated salinization (Miller et al., 2017; Thorslund et al., 2021), increased concentrations of nitrate (Lin
595 et al., 2021), and other nutrients (Stets et al., 2020), contributions of pesticide and fertilizers, or alterations to water temperature
profiles. These contributions of agricultural waters contribute to sustaining flow but threaten water availability. Thus, inclusion
of groundwater return flows from irrigation to rivers in the Western US supports improved assessments of water availability
both through improved modeling of streamflows and enhanced ability to model water quality.

Explicit inclusion of irrigation return flows will assist the NWM in better projecting streamflows during periods of hydrologic
600 non-stationarity, as are likely to characterize the hydroclimatic elements of climate change. Non-stationary processes include
hydrologic changes arising from the ongoing mega-drought of the southwestern US (Williams et al., 2022), intense precipi-
tation events like monsoons or major storm events that are observed to be increasing in intensity with climate change (Pfahl
et al., 2017; Demaria et al., 2019), and projected changes to future snowpack depth and melt timing (Siirila-Woodburn et al.,
2021; Hammond et al., 2023). The ongoing aridification of the southwestern US is characterized by increased evapotranspira-
605 tion (Milly and Dunne, 2020), and changes to groundwater recharge and discharge associated with decreases in snowpack and
changes to snowpack melt patterns (Hammond et al., 2023). Understanding the groundwater flux contributions of areas with
shallow water tables to streamflow during major precipitation events will help better characterize areas at risk for flooding and
inform appropriate water management strategies.

4 Conclusions

610 The isotope mass balance evaluation of the NWM revealed similarities between the isotope mass balance estimated isotope
ratios (modeled) and observed isotope ratios. The mass balance approach captured as much as 75% of the variance in the
observations, depending on the water isotopologue evaluated. This suggests that, on mean, during the summer, the NWM
correctly represents the relative proportions of groundwater and runoff fluxes sustaining streamflow, and the gridded isotope
datasets are appropriate for the analysis.

615 The observation-model differences exhibited spatial and seasonal structure, suggesting that the NWM is missing important
additional water sources that contribute to streamflow. Specifically, the observation-model differences that plot above the
equilibrium line (Figure 2) suggest the importance of direct contributions of snowmelt to streamflow in humid areas. Those that



plot below the equilibrium line suggest the importance of groundwater sources characterized by evapoconcentration in arid areas. We tested the hypothesis that agricultural irrigation return flows are the missing, isotopically evapoconcentrated water source in arid regions, and found them to be a significant predictor of observation-model differences. This finding is supported by multiple lines of evidence including land use pattern influence on isotopic evapoconcentration of groundwaters, a comparison of NWM groundwater discharge and an independent assessment of the potential for groundwater discharge and other isotopic inferences and modeling studies.

Our findings suggest that the NWM accuracy would be improved by including agricultural return flows into the NWM water sources. Agricultural return flows function as lower elevation baseflow fluxes, and are likely to be critical for sustaining streamflow during drought periods or late in the growing season. Inclusion of this specific source into the more general groundwater fluxes would thus improve the ability to meet water manager and water user NWM data needs. Specifically, water managers use predictions of reach-specific flows at lower elevations during summer precipitation events and monsoons to assess flood risk, or to inform dam releases (if dam releases are incorporated into the NWM) to assess the volume of water required to achieve specific management goals like fish species preservation or dam water level maintenance for hydropower production. Likewise, explicit inclusion of irrigation return flows in NWM calculations will assist in accurately predicting and projecting streamflows in heavily managed sections of river in the event of changing irrigation practices, increased evapotranspiration, or water supply reductions and fallowing of agricultural fields, which would change or halt irrigation groundwater fluxes. Finally, our findings have implications for areas at risk for diminished water availability due to issues of quality, arising from the entrainment of fertilizer and pesticides and as well as dissolution and delivery of salts.

Data availability. Data are publicly available (Reddy et al., 2023).

Author contributions. Conceptualization and funding acquisition: ALP and OLM; Data curation: PCL, MCM, MK, JR, and JRB; Methodology, investigation and formal analysis: ALP and PCL; Project administration and supervision: ALP; Visualization: ALP; writing (original draft): ALP; Writing - review: OLM, PCL, MCM, JR, MK and JRB

Competing interests. The authors declare no competing interests

Acknowledgements. The authors acknowledge financial support from the US Geological Survey (USGS) Water Mission area Water Quality Processes Division. We appreciate comments from two USGS peer reviewers, whose comments improved the quality of the manuscript. The views expressed in this article are those of the author(s) and do not necessarily represent the views or the policies of the US Environmental Protection Agency. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.



References

- Aemisegger, F. and Sjolte, J.: A Climatology of Strong Large-Scale Ocean Evaporation Events. Part II: Relevance for the Deuterium Excess Signature of the Evaporation Flux, *Journal of Climate*, 31, 7313 – 7336, <https://doi.org/https://doi.org/10.1175/JCLI-D-17-0592.1>, 2018.
- Ala-aho, P., Tetzlaff, D., McNamara, J. P., Laudon, H., Kormos, P., and Soulsby, C.: Modeling the isotopic evolution of snowpack and snowmelt: Testing a spatially distributed parsimonious approach, *Water Resources Research*, 53, 5813–5830, <https://doi.org/https://doi.org/10.1002/2017WR020650>, 2017.
- Anderson, L., Berkelhammer, M., and Mast, M. A.: Isotopes in North American Rocky Mountain Snowpack 1993–2014, *Quaternary Science Reviews*, 131, 262–273, <https://doi.org/https://doi.org/10.1016/j.quascirev.2015.03.023>, 2016.
- Barnhart, T. B., Molotch, N. P., Livneh, B., Harpold, A. A., Knowles, J. F., and Schneider, D.: Snowmelt rate dictates streamflow, *Geophysical Research Letters*, 43, 8006–8016, <https://doi.org/https://doi.org/10.1002/2016GL069690>, 2016.
- Beria, H., Larsen, J. R., Ceperley, N. C., Michelon, A., Vennemann, T., and Schaeffli, B.: Understanding snow hydrological processes through the lens of stable water isotopes, *WIREs Water*, 5, e1311, <https://doi.org/https://doi.org/10.1002/wat2.1311>, 2018.
- Bowen, G.: The Online Isotopes in Precipitation Calculator, version 2.0, https://wateriso.utah.edu/waterisotopes/pages/information/oipc_info.html, 2006.
- 660 Bowen, G.: The Online Isotopes in Precipitation Calculator, version 3.1, <http://www.waterisotopes.org>, 2022a.
- Bowen, G. J.: Spatial analysis of the intra-annual variation of precipitation isotope ratios and its climatological corollaries, *Journal of Geophysical Research: Atmospheres*, 113, <https://doi.org/https://doi.org/10.1029/2007JD009295>, 2008.
- Bowen, G. J.: Gridded maps of the isotopic composition of meteoric waters, <http://www.waterisotopes.org>, uSA grids downloaded Nov 2021, 2022b.
- 665 Bowen, G. J., Wassenaar, L. I., and Hobson, K. A.: Global application of stable hydrogen and oxygen isotopes to wildlife forensics, *Oecologia*, 143, 337–348, <https://doi.org/10.1007/s00442-004-1813-y>, 2005.
- Bowen, G. J., Putman, A. L., Brooks, J. R., Bowling, D. R., Oerter, E. J., and Good, S. P.: Inferring the source of evaporated waters using stable H and O isotopes, *Oecologia*, 187, 1025–1039, <https://doi.org/10.1007/s00442-018-4192-5>, 2018.
- Bowen, G. J., Cai, Z., Fiorella, R. P., and Putman, A. L.: Isotopes in the water cycle: Regional-to global-scale patterns and applications, *Annual Review of Earth and Planetary Sciences*, 47, <https://doi.org/10.1146/annurev-earth-053018-060220>, 2019.
- 670 Bowen, G. J., Guo, J. S., and Allen, S. T.: A 3-D groundwater isoscape of the contiguous USA for forensic and water resource science, *PLOS ONE*, pp. 1–22, <https://doi.org/10.1371/journal.pone.0261651>, 2022.
- Brooks, J. R., Gibson, J. J., Birks, S. J., Weber, M. H., Rodecap, K. D., and Stoddard, J. L.: Stable isotope estimates of evaporation : inflow and water residence time for lakes across the United States as a tool for national lake water quality assessments, *Limnology and Oceanography*, 59, 2150–2165, <https://doi.org/https://doi.org/10.4319/lo.2014.59.6.2150>, 2014.
- 675 Brooks, P. D., Gelderloos, A., Wolf, M. A., Jamison, L. R., Strong, C., Solomon, D. K., Bowen, G. J., Burian, S., Tai, X., Arens, S., Briefer, L., Kirkham, T., and Stewart, J.: Groundwater-Mediated Memory of Past Climate Controls Water Yield in Snowmelt-Dominated Catchments, *Water Resources Research*, 57, e2021WR030605, <https://doi.org/https://doi.org/10.1029/2021WR030605>, 2021.
- Bruckerhoff, L. A., Wheeler, K., Dibble, K. L., Mihalevich, B. A., Neilson, B. T., Wang, J., Yackulic, C. B., and Schmidt, J. C.: Water Storage Decisions and Consumptive Use May Constrain Ecosystem Management under Severe Sustained Drought, *JAWRA Journal of the American Water Resources Association*, 58, <https://doi.org/https://doi.org/10.1111/1752-1688.13020>, 2022.



- Bureau of Reclamation, Department of the Interior: Interior Department Announces Actions to Protect Colorado River System, Sets 2023 Operating Conditions for Lake Powell and Lake Mead, Bureau of Reclamation News; Multimedia, <https://www.usbr.gov/newsroom/#/news-release/4294>, accessed Aug 18, 2022, 2022.
- 685 Burt, E. I., Coayla Rimachi, D. H., Ccahuana Quispe, A. J., Atwood, A., and West, A. J.: Isotope-derived young water fractions in streamflow across the tropical Andes mountains and Amazon floodplain, *Hydrology and Earth System Sciences*, 27, 2883–2898, <https://doi.org/10.5194/hess-27-2883-2023>, 2023.
- Carroll, R. W. H., Deems, J., Maxwell, R., Sprenger, M., Brown, W., Newman, A., Beutler, C., Bill, M., Hubbard, S. S., and Williams, K. H.: Variability in observed stable water isotopes in snowpack across a mountainous watershed in Colorado, *Hydrological Processes*, 690 36, e14 653, <https://doi.org/https://doi.org/10.1002/hyp.14653>, 2022.
- Condon, L. E. and Maxwell, R. M.: Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion, *Science Advances*, 5, eaav4574, <https://doi.org/https://dx.doi.org/10.1126/sciadv.aav4574>, 2019.
- Corcoran, M. C., Thomas, E. K., and Boutt, D. F.: Event-Based Precipitation Isotopes in the Laurentian Great Lakes Region Reveal Spatiotemporal Patterns in Moisture Recycling, *Journal of Geophysical Research: Atmospheres*, 124, 5463–5478, 695 <https://doi.org/https://doi.org/10.1029/2018JD029545>, 2019.
- Craig, H.: Isotopic Variations in Meteoric Waters, *Science*, 133, 1702–1703, <http://www.jstor.org/stable/1708089>, 1961.
- Craig, H. and Gordon, L. I.: Deuterium and oxygen 18 variations in the ocean and marine atmosphere, in: *Stable Isotopes in Oceanographic Studies and Paleotemperatures*, p. 9, 1965.
- Demaria, E. M. C., Hazenberg, P., Scott, R. L., Meles, M. B., Nichols, M., and Goodrich, D.: Intensification of the North American Monsoon Rainfall as Observed From a Long-Term High-Density Gauge Network, *Geophysical Research Letters*, 46, 6839–6847, 700 <https://doi.org/https://doi.org/10.1029/2019GL082461>, 2019.
- Dewitz, J. and U.S. Geological Survey: National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021), <https://doi.org/https://doi.org/10.5066/P9KZCM54>, 2021.
- Dibble, K., Yackulic, C., and Bestgen, K.: Water temperature models, data and code for the Colorado, Green, San Juan, Yampa, and White 705 rivers in the Colorado River basin, <https://doi.org/https://doi.org/10.5066/P9HFKV7Q>, 2020.
- Dieter, C., Maupin, M., Caldwell, R., Harris, M., Ivahnenko, T., Lovelace, J., Barber, N., and Linsey, K.: Estimated use of water in the United States in 2015: U.S. Geological Survey Circular 1441, Tech. rep., U.S. Geological Survey, <https://doi.org/https://doi.org/10.3133/cir1441>, 2018.
- Essaid, H. I. and Caldwell, R. R.: Evaluating the impact of irrigation on surface water – groundwater interaction and stream temperature in an agricultural watershed, *Science of The Total Environment*, 599-600, 581–596, 710 <https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.04.205>, 2017.
- Feng, X., Faiia, A. M., and Posmentier, E. S.: Seasonality of isotopes in precipitation: A global perspective, *Journal of Geophysical Research: Atmospheres*, 114, <https://doi.org/https://doi.org/10.1029/2008JD011279>, 2009.
- Fergus, C. E., Brooks, J. R., Kaufmann, P. R., Herlihy, A. T., Pollard, A. I., Weber, M. H., and Paulsen, S. G.: Lake Water Levels and 715 Associated Hydrologic Characteristics in the Conterminous U.S., *Journal of the American Water Resources Association*, 56, 450–471, <https://doi.org/https://doi.org/10.1111/1752-1688.12817>, 2020.
- Fergus, C. E., Brooks, J. R., Kaufmann, P. R., Pollard, A. I., Herlihy, A. T., Paulsen, S. G., and Weber, M. H.: National framework for ranking lakes by potential for anthropogenic hydro-alteration, *Ecological Indicators*, 122, 107 241, <https://doi.org/https://doi.org/10.1016/j.ecolind.2020.107241>, 2021.



- 720 Fergus, C. E., Brooks, J. R., Kaufmann, P. R., Pollard, A. I., Mitchell, R., Geldhof, G. J., Hill, R. A., Paulsen, S. G., Ringold, P., and Weber, M.: Natural and anthropogenic controls on lake water-level decline and evaporation-to-inflow ratio in the conterminous United States, *Limnology and Oceanography*, 67, 1484–1501, <https://doi.org/https://doi.org/10.1002/lno.12097>, 2022.
- Fillo, N. K., Bhaskar, A. S., and Jefferson, A. J.: Lawn Irrigation Contributions to Semi-Arid Urban Baseflow Based on Water-Stable Isotopes, *Water Resources Research*, 57, e2020WR028 777, <https://doi.org/https://doi.org/10.1029/2020WR028777>, 2021.
- 725 Friedrich, K., Grossman, R. L., Huntington, J., Blanken, P. D., Lenters, J., Holman, K. D., Gochis, D., Livneh, B., Prairie, J., Skeie, E., Healey, N. C., Dahm, K., Pearson, C., Finnessey, T., Hook, S. J., and Kowalski, T.: Reservoir Evaporation in the Western United States: Current Science, Challenges, and Future Needs, *Bulletin of the American Meteorological Society*, 99, 167 – 187, <https://doi.org/https://doi.org/10.1175/BAMS-D-15-00224.1>, 2018.
- Gentile, A., Canone, D., Ceperley, N., Gisolo, D., Previati, M., Zuecco, G., Schaeffli, B., and Ferraris, S.: Towards a conceptualization of the hydrological processes behind changes of young water fraction with elevation: a focus on mountainous alpine catchments, *Hydrology and Earth System Sciences*, 27, 2301–2323, <https://doi.org/10.5194/hess-27-2301-2023>, 2023.
- 730 Gochis, D., Barlage, M., Dugger, A., Karsten, L., McAllister, M., McCreight, J., Mills, J., RafieeiNasab, A., Read, L., Sampson, K., Yates, D., and Yu, W.: The WRF-Hydro Modeling System Technical Description (Version 5.0), Tech. rep., National Center for Atmospheric Research, <https://doi.org/https://doi.org/10.5065/D6J38RBJ>, nCAR Technical Note, 2018.
- 735 Gochis, D., Barlage, M., Cabell, R., Dugger, A., Fanfarillo, A., FitzGerald, K., McAllister, M., McCreight, J., RafieeiNasab, A., Read, L., Frazier, N., Johnson, D., Mattern, J. D., Karsten, L., Mills, T. J., and Fersch, B.: WRF-Hydro v5.1.1, <https://doi.org/https://doi.org/10.5281/zenodo.3625238>, dataset, 2020a.
- Gochis, D. J., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., McAllister, M., McCreight, J., RafieeiNasab, A., Read, L., Sampson, K., Yates, D., and Zhang, Y.: The WRF-Hydro modeling system technical description, (Version 5.1.1), Tech. rep., UCAR, <https://ral.ucar.edu/sites/default/files/public/projects/wrf-hydro/technical-description-user-guide/wrf-hydrov5.2technicaldescription.pdf>, nCAR
- 740 Technical Note, 2020b.
- Grafton, R. Q., Williams, J., Perry, C. J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., and Allen, R. G.: The paradox of irrigation efficiency, *Science*, 361, 748–750, <https://doi.org/10.1126/science.aat9314>, 2018.
- Hammond, J. C. and Kampf, S. K.: Subannual Streamflow Responses to Rainfall and Snowmelt Inputs in Snow-Dominated Watersheds of the Western United States, *Water Resources Research*, 56, e2019WR026 132, <https://doi.org/https://doi.org/10.1029/2019WR026132>, 2020.
- 745 Hammond, J. C., Sexstone, G. A., Putman, A. L., Barnhart, T. B., Rey, D. M., Driscoll, J. M., Liston, G. E., Rasmussen, K. L., McGrath, D., Fassnacht, S. R., and Kampf, S. K.: High Resolution SnowModel Simulations Reveal Future Elevation-Dependent Snow Loss and Earlier, Flashier Surface Water Input for the Upper Colorado River Basin, *Earth’s Future*, 11, e2022EF003 092, <https://doi.org/https://doi.org/10.1029/2022EF003092>, 2023.
- 750 Hansen, C., Shafiei Shiva, J., McDonald, S., and Nabors, A.: Assessing Retrospective National Water Model Streamflow with Respect to Droughts and Low Flows in the Colorado River Basin, *Journal of the American Water Resources Association*, 55, 964–975, <https://doi.org/https://doi.org/10.1111/1752-1688.12784>, 2019.
- Harris, I., Osborn, T., Jones, P., and D., L.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, *Sci Data*, 7, <https://doi.org/https://doi.org/10.1038/s41597-020-0453-3>, 2020.
- 755 Hicke, J., Lucatello, S., Mortsch, L., Dawson, J., Aguilar, M. D., Enquist, C., Gilmore, E., Gutzler, D., Harper, S., Holsman, K., Jewett, E., Kohler, T., and Miller, K.: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth



- Assessment Report of the Intergovernmental Panel on Climate Change, chap. 14: North America, p. 1929–2042, Cambridge University Press, Cambridge, UK and New York, NY, USA, <https://doi.org/10.1017/9781009325844.016>, 2022.
- 760 Jasechko, S., Birks, S. J., Gleeson, T., Wada, Y., Fawcett, P. J., Sharp, Z. D., McDonnell, J. J., and Welker, J. M.: The pronounced seasonality of global groundwater recharge, *Water Resources Research*, 50, 8845–8867, <https://doi.org/10.1002/2014WR015809>, 2014.
- Jasechko, S., Seybold, H., Perrone, D., Ying, F., and Kirchner, J.: Widespread potential loss of streamflow into underlying aquifers across the USA, *Nature*, 591, 391–395, <https://doi.org/10.1038/s41586-021-03311-x>, 2021.
- Kornfield, M.: Rio Grande runs dry in Albuquerque for the first time in 40 years, <https://www.washingtonpost.com/climate-environment/2022/07/22/rio-grande-drought/>, [Online; posted July 22, 2022, accessed Aug 18, 2022], 2022.
- 765 Li, D., Wrzesien, M. L., Durand, M., Adam, J., and Lettenmaier, D. P.: How much runoff originates as snow in the western United States, and how will that change in the future?, *Geophysical Research Letters*, 44, 6163–6172, <https://doi.org/10.1002/2017GL073551>, 2017.
- Lin, J., Compton, J. E., Hill, R. A., Herlihy, A. T., Sabo, R. D., Brooks, J. R., Weber, M., Pickard, B., Paulsen, S. G., and Stoddard, J. L.: Context is Everything: Interacting Inputs and Landscape Characteristics Control Stream Nitrogen, *Environmental Science & Technology*, 770 55, 7890–7899, <https://doi.org/10.1021/acs.est.0c07102>, 2021.
- McGill, L. M., Brooks, J. R., and Steel, E. A.: Spatiotemporal dynamics of water sources in a mountain river basin inferred through $\delta^2\text{H}$ and $\delta^{18}\text{O}$ of water, *Hydrological Processes*, 35, e14063, <https://doi.org/10.1002/hyp.14063>, 2021.
- Miller, M. P., Buto, S. G., Susong, D. D., and Rumsey, C. A.: The importance of base flow in sustaining surface water flow in the Upper Colorado River Basin, *Water Resources Research*, 52, 3547–3562, <https://doi.org/10.1002/2015WR017963>, 2016.
- 775 Miller, M. P., Buto, S. G., Lambert, P. M., and Rumsey, C. A.: Enhanced and updated spatially referenced statistical assessment of dissolved-solids load sources and transport in streams of the Upper Colorado River Basin, Tech. rep., US Geological Survey, <https://doi.org/10.3133/sir20175009>, 2017.
- Miller, O. L., Miller, M. P., Longley, P. C., Alder, J. R., Bearup, L. A., Pruitt, T., Jones, D. K., Putman, A. L., Rumsey, C. A., and McKinney, T.: How will baseflow respond to climate change in the Upper Colorado River Basin?, *Geophysical Research Letters*, n/a, e2021GL095085, 780 <https://doi.org/10.1029/2021GL095085>, 2021a.
- Miller, O. L., Putman, A. L., Alder, J., Miller, M., Jones, D. K., and Wise, D. R.: Changing climate drives future streamflow declines and challenges in meeting water demand across the southwestern United States, *Journal of Hydrology X*, 11, 100074, <https://doi.org/10.1016/j.hydroa.2021.100074>, 2021b.
- Milly, P. and Dunne, K. A.: Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation, *Science*, 367, 785 1252–1255, <https://doi.org/10.1126/science.aay9187>, 2020.
- Mote, P., Li, S., Lettenmaier, D., Xiao, M., and Engle, R.: Dramatic declines in snowpack in the western U.S., *npj Clim Atmos Sci*, 1, <https://doi.org/10.1038/s41612-018-0012-1>, 2021.
- National Oceanographic and Atmospheric Administration: National Water Model CONUS Retrospective Dataset, <https://registry.opendata.aws/nwm-archive>, data accessed in 2022, 2022.
- 790 Nickolas, L. B., Segura, C., and Brooks, J. R.: The influence of lithology on surface water sources, *Hydrological Processes*, 31, 1913–1925, <https://doi.org/10.1002/hyp.11156>, 2017.
- Nusbaumer, J., Wong, T. E., Bardeen, C., and Noone, D.: Evaluating hydrological processes in the Community Atmosphere Model Version 5 (CAM5) using stable isotope ratios of water, *Journal of Advances in Modeling Earth Systems*, 9, 949–977, <https://doi.org/10.1002/2016MS000839>, 2017.



- 795 Oerter, E., Malone, M., Putman, A., Drits-Esser, D., Stark, L., and Bowen, G.: Every apple has a voice: using stable isotopes to teach about food sourcing and the water cycle, *Hydrology and Earth System Sciences*, 21, 3799–3810, <https://doi.org/10.5194/hess-21-3799-2017>, 2017.
- Pfahl, S., O’Gorman, P., and Fischer, E.: Understanding the regional pattern of projected future changes in extreme precipitation, *Nature Climate Change*, 7, 423–427, <https://doi.org/https://doi.org/10.1038/nclimate3287>, 2017.
- 800 Putman, A. L. and Bowen, G. J.: A global database of the stable isotopic ratios of meteoric and terrestrial waters, *Hydrology and Earth System Sciences*, 23, 4389–4396, <https://doi.org/https://doi.org/10.5194/hess-23-4389-2019>, 2019.
- Putman, A. L., Fiorella, R. P., Bowen, G. J., and Cai, Z.: A global perspective on local meteoric water lines: Meta-analytic insight into fundamental controls and practical constraints, *Water Resources Research*, 55, 6896–6910, <https://doi.org/https://doi.org/10.1029/2019WR025181>, 2019.
- 805 Ramirez, R.: Great Salt Lake is ‘in trouble’ as level falls to lowest on record for second year in a row, <https://www.cnn.com/2022/07/06/us/great-salt-lake-record-low-climate/index.html>, [Online; posted July 6, 2022, accessed Aug 18, 2022], 2022.
- Reddy, J., Longley, P. C., McDonnell, M. C., Katoski, M. P., Miller, O. L., and Putman, A.: Hydrogen and oxygen stable isotope mass balance evaluation of the National Water Model (v2.1) streamflow, runoff and groundwater flows, <https://doi.org/https://doi.org/10.5066/P9NOD5ES>, U.S. Geological Survey data release, 2023.
- 810 Rubel, F. and Kottek, M.: Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification, *Meteorologische Zeitschrift*, 19, 135–141, <https://doi.org/10.1127/0941-2948/2010/0430>, 2010.
- Seabold, S. and Perktold, J.: Statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, <https://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf>, 2010.
- Seo, B.-C., Krajewski, W. F., and Quintero, F.: Multi-Scale Hydrologic Evaluation of the National Water Model Streamflow Data Assimilation, *JAWRA Journal of the American Water Resources Association*, 57, 875–884, <https://doi.org/https://doi.org/10.1111/1752-1688.12955>, 2021.
- Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., Collins, W. D., et al.: A low-to-no snow future and its impacts on water resources in the western United States, *Nature Reviews Earth & Environment*, 2, 800–819, <https://doi.org/https://doi.org/10.1038/s43017-021-00219-y>, 2021.
- 820 Solder, J. and Beisner, K.: Critical evaluation of stable isotope mixing end-members for estimating groundwater recharge sources: case study from the South Rim of the Grand Canyon, Arizona, USA, *Hydrogeol J*, 28, 1575–1591, <https://doi.org/https://doi.org/10.1007/s10040-020-02194-y>, 2020.
- Stets, E. G., Sprague, L. A., Oelsner, G. P., Johnson, H. M., Murphy, J. C., Ryberg, K., Vecchia, A. V., Zuellig, R. E., Falcone, J. A., and Riskin, M. L.: Landscape Drivers of Dynamic Change in Water Quality of U.S. Rivers, *Environmental Science & Technology*, 54, 4336–4343, <https://doi.org/10.1021/acs.est.9b05344>, 2020.
- 825 The pandas development team: pandas-dev/pandas: Pandas, <https://doi.org/10.5281/zenodo.3509134>, 2020.
- Theobald, D. M., Stevens, D. L., White, D., Urquhart, N. S., Olsen, A. R., and Norman, J. B.: Using GIS to Generate Spatially Balanced Random Survey Designs for Natural Resource Applications, *Environmental Management*, 40, 134–146, <https://doi.org/10.1007/s00267-005-0199-x>, 2007.
- 830 Thorslund, J., Bierkens, M., and Oude Essink, G. e. a.: Common irrigation drivers of freshwater salinisation in river basins worldwide, *Nature Communications*, 12, 4232, <https://doi.org/https://doi.org/10.1038/s41467-021-24281-8>, 2021.



- Towler, E., Foks, S. S., Dugger, A. L., Dickinson, J. E., Essaid, H. I., Gochis, D., Viger, R. J., and Zhang, Y.: Benchmarking high-resolution hydrologic model performance of long-term retrospective streamflow simulations in the contiguous United States, *Hydrology and Earth System Sciences*, 27, 1809–1825, <https://doi.org/10.5194/hess-27-1809-2023>, 2023.
- 835 Tulley-Cordova, C. L., Putman, A. L., and Bowen, G. J.: Stable Isotopes in Precipitation and Meteoric Water: Sourcing and Tracing the North American Monsoon in Arizona, New Mexico, and Utah, *Water Resources Research*, 57, e2021WR030039, <https://doi.org/10.1029/2021WR030039>, 2021.
- University of East Anglia Climatic Research Unit and Harris, I.C. and Jones, P.D. and Osborn, T. : CRU TS4.05: Climatic Research Unit (CRU) Time-Series (TS) version 4.05 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2020),
840 <https://catalogue.ceda.ac.uk/uuid/c26a65020a5e4b80b20018f148556681>, accessed January, 2022, 2021.
- U.S. Environmental Protection Agency: National lakes assessment: A collaborative survey of the nations’s lakes, Tech. rep., Office of Water and Office of Research and Development, https://www.epa.gov/sites/default/files/2013-11/documents/nla_newlowres_fullrpt.pdf, 2009.
- U.S. Environmental Protection Agency: National Lakes Assessment 2012: A Collaborative Survey of Lakes in the United States, Tech. rep., Office of Water and Office of Research and Development, https://www.epa.gov/sites/default/files/2016-12/documents/nla_report_dec_2016.pdf, 2016a.
845
- U.S. Environmental Protection Agency: National rivers and streams assessment 2008-2009: A collaborative survey, Tech. rep., Office of Water and Office of Research and Development, https://www.epa.gov/sites/default/files/2016-03/documents/nrsa_0809_march_2_final.pdf, 2016b.
- U.S. Environmental Protection Agency: National rivers and streams assessment 2013-2014: A collaborative survey, Tech. rep., Office of Water and Office of Research and Development, https://www.epa.gov/system/files/documents/2021-10/nrsa_13-14_report_508_ci_2021-10-15.pdf, 2020.
850
- U.S. Geological Survey: National Hydrography Dataset (ver. 2.1), <https://www.usgs.gov/national-hydrography/access-national-hydrography-products>, accessed Oct, 2021, 2019.
- U.S. Geological Survey: National Water Information System: U.S. Geological Survey web interface,
855 <https://doi.org/http://dx.doi.org/10.5066/F7P55KJN>, 2022.
- U.S. Geological Survey, National Geospatial Technical Operations Center: Watershed Boundary Dataset (WBD) - USGS National Map Downloadable Data Collection, <https://www.sciencebase.gov/catalog/item/51361e87e4b03b8ec4025c22#:~:text=Citation,Data%20Collection%3A%20U.S.%20Geological%20Survey.,> scienceBase Data Release, 2023.
- Wes McKinney: Data Structures for Statistical Computing in Python, in: *Proceedings of the 9th Python in Science Conference*, edited by
860 Stéfan van der Walt and Jarrod Millman, pp. 56 – 61, <https://doi.org/10.25080/Majora-92bf1922-00a>, 2010.
- Williams, A., Cook, B., and Smerdon, J.: Rapid intensification of the emerging southwestern North American megadrought in 2020–2021, *Nat. Clim. Chang.*, 12, <https://doi.org/10.1038/s41558-022-01290-z>, 2022.
- Windler, G., Brooks, J. R., Johnson, H. M., Comeleo, R. L., Coulombe, R., and Bowen, G. J.: Climate Impacts on Source Contributions and Evaporation to Flow in the Snake River Basin Using Surface Water Isoscapes ($\delta^2\text{H}$ and $\delta^{18}\text{O}$), *Water Resources Research*, 57, e2020WR029157, <https://doi.org/10.1029/2020WR029157>, 2021.
865
- Wolf, M. A., Jamison, L. R., Solomon, D. K., Strong, C., and Brooks, P. D.: Multi-Year Controls on Groundwater Storage in Seasonally Snow-Covered Headwater Catchments, *Water Resources Research*, 59, e2022WR033394, <https://doi.org/10.1029/2022WR033394>, 2023.



870 Xia, C., Liu, Y., Meng, Y., Liu, G., Huang, X., Chen, Y., and Chen, K.: Stable isotopes reveal the surface water-
groundwater interaction and variation in young water fraction in an urbanized river zone, *Urban Climate*, 51, 101641,
<https://doi.org/https://doi.org/10.1016/j.uclim.2023.101641>, 2023.

875 Yang, Z., Qian, Y., Liu, Y., Berg, L. K., Hu, H., Dominguez, F., Yang, B., Feng, Z., Gustafson Jr, W. I., Huang, M., and Tang, Q.: Irrigation
Impact on Water and Energy Cycle During Dry Years Over the United States Using Convection-Permitting WRF and a Dynamical Recy-
cling Model, *Journal of Geophysical Research: Atmospheres*, 124, 11 220–11 241, <https://doi.org/https://doi.org/10.1029/2019JD030524>,
2019.