An improved practical approach for estimating catchment-scale response functions through wavelet analysis

R. Dwivedi[1], C. Eastoe[2], J. F. Knowles[3,4], L. Hamann[5], T. Meixner[1], P. A. "Ty" Ferre[1], C. Castro[1], W. E. Wright[6], G.-Y. Niu[1], R. Minor[4,7], G. A. Barron-Gafford[4,7], N. Abramson[1], B. Mitra[8,9], S. A. Papuga[10], M. Stanley[11], and J. Chorover[12]

[1]Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, AZ, USA, [2]Department of Geosciences (retired), The University of Arizona, Tucson, AZ, USA, [3]Southwest Watershed Research Center, USDA Agricultural Research Service, Tucson, AZ, USA, [4]School of Geography, Development & Environment, The University of Arizona, Tucson, AZ, USA, [5]United States Geological Survey, Tucson, AZ, USA, [6]Laboratory of Tree-Ring Research, The University of Arizona, Tucson, AZ, USA, [7]Biosphere 2, The University of Arizona, Tucson, AZ, USA, [8]School of Natural Resources and the Environment, The University of Arizona, Tucson, AZ, USA, [9]School of Ecosystem Science and Management, Texas A&M University, College Station, TX, USA, [10]Department of Geology and Environmental Science Program, Wayne State University, Detroit, MI, USA, [11]Mt. Lemmon Water District, Tucson, AZ, USA, [12]Department of Environmental Science, The University of Arizona, Tucson, AZ, USA

Corresponding author: Ravindra Dwivedi (ravindradwivedi@email.arizona.edu)

Abstract

Catchment-scale response functions, such as transit time distribution (TTD) and evapotranspiration time distribution (ETTD), are considered fundamental descriptors of a catchment's hydrologic and ecohydrologic responses to spatially and temporally varying precipitation inputs. Yet, estimating these functions is challenging, especially in headwater catchments where data collection is complicated by rugged terrain, or in semi-arid or sub-humid areas where precipitation is infrequent. Hence, we developed practical approaches for estimating both TTD and ETTD from commonly available tracer flux data in hydrologic inflows and outflows without requiring continuous observations. Using the weighted wavelet spectral analysis method of *Kirchner and Neal* [2013] for $\delta^{18}O$ in precipitation and stream water, we specifically calculated TTDs that contribute to streamflow via spatially and temporally variable flow paths in a sub-humid mountain headwater catchment in Arizona, USA. Our results indicate that composite TTDs most accurately represented this system for periods up to approximately one month and that a Gamma TTD was most appropriate thereafter. The TTD results also suggested that some contribution of subsurface water was beyond the applicable tracer range. For ETTD and using $\delta^{18}O$ as a tracer in precipitation and xylem waters, a Gamma ETTD type best matched the observations, and stable water isotopes were capable tracers for the majority of vegetation source waters. This study contributes to a better understanding of a fundamental question in mountain catchment hydrology; namely, how tracer input fluxes are modulated by spatially and temporally varying subsurface flow paths that support evapotranspiration and streamflow at multiple time scales.

Key words: Stable water isotopes, transit time distribution, ET time distribution, spectral analysis, soil water, subsurface storage, mountain, headwater

1. Introduction

Mountain systems are regarded as the "water towers for humanity" [*Viviroli et al.*, 2007]. They provide critical ecosystem services including groundwater recharge to adjacent alluvial basins, particularly in arid and semi-arid regions [*Milly and Dunne*, 2020] that cover ~41% of earth's surface [*Reynolds et al.*, 2007] and are expected to grow by ~34% due to climate change [*Zeng and Yoon*, 2009]. Despite the importance of mountain systems for water supply, knowledge of their ecohydrological and biogeochemical functioning remains limited [*Milly and Dunne*, 2020]. Fundamental hydrologic questions remain, such as how spatially and temporally varying subsurface flow that supports evapotranspiration and streamflow modulates solute fluxes at multiple time scales. To address this question, the current study examines relationships between input (precipitation) and output (streamflow and evapotranspiration) tracer fluxes by estimating catchment-scale transit time (TTD) and evapotranspiration time (ETTD) distributions. TTDs and ETTDs are is succinct descriptors, in terms of length and velocity, of the numerous flow paths that connect streams and evapotranspiration to multiple source regions [*Godsey et al.*, 2010; *Kirchner*, 2016a]. Both TTD and ETTD are helpful metrics for understanding the short- and long-term memories of subsurface storage systems.

Many studies have estimated stream water TTDs using time-domain convolution, assuming that a whole watershed can be treated as a single elementary volume [*McGuire and McDonnell*, 2006]. In this way, a dominant period-band is assumed for input and output fluxes and storages, such that the TTD type does not change and the TTD parameters are the only variable quantities [*Birkel et al.*, 2012; *Heidbüchel et al.*, 2012]. Similarly, estimating TTD by wavelet analysis of tracer time series entails *a priori* assumption of distribution type [*Onderka and Chudoba*, 2018]. A principal limitation of these approaches is the requirement of continuous observations at a uniform sampling interval for both hydrologic fluxes and stable water isotopes [*Kirchner*, 2016a], with some approaches also requiring an unknowable estimate of total catchment storage [*Benettin et al.*, 2017; *Harman*, 2015]. These requirements are difficult to fulfill for many field studies [*Berghuijs and Kirchner*, 2017; *Kirchner*, 2016b], especially in rugged mountainous areas where access is difficult due to rugged terrain and fractured bedrock aquifers are poorly characterized. If data gaps are filled by interpolation, the effects on the estimated TTD or ETTD are rarely reported, which is problematic because interpolated values may have lower variance than observed values [*Birkel et al.*, 2012; *Feng et al.*, 2004; *Hirsch et al.*, 2010; *Zhang et al.*, 2018]. As a result, the current research seeks to develop catchment-scale TTDs and ETTDs that are resilient to data gaps in both tracer and hydrometric time series.

Major obstacles to estimating ETTDs include a lack of continuous stable isotope data for xylem water and poor understanding of sources of local evapotranspiration and streamflow, which can lead to significant uncertainty [*Brooks et al.*, 2009]. Because source water identification is difficult, most existing approaches use either numerical modeling [*Botter et al.*, 2010; *van der Velde et al.*, 2012] or base ETTD on transport models calibrated to stream water tracer concentrations [*Benettin et al.*, 2017; *Harman*, 2015]. *Benettin et al.* [2017] specifically reported difficulty identifying ETTD parameters in a catchment-scale study based on hydrologic fluxes and stable water isotopes in precipitation and streamflow. Accordingly, a practical approach is needed for estimation of catchment-scale ETTD in the context of discontinuous or incomplete datasets.

The current study proposes novel practical approaches to estimate catchment-scale TTD and ETTD that are moderately resistant to data gaps in tracer time series and hydrometric observations. The proposed approaches do not assume specific TTD or ETTD types, in contrast to existing approaches. Instead, multiple TTD and ETTD types are objectively evaluated for suitability to the observed long-term data. The proposed approaches are practical in the sense that they do not require input from time-consuming detailed flow and transport models. It is therefore expected that the proposed approaches will have wide applicability, especially in remote areas where many study catchments are located and where data gaps are unavoidable. At sites where environmental water samples have been analyzed for several years at irregular intervals, the proposed approaches can also be employed to estimate catchment-scale response functions. The overarching goal of this work was to leverage TTD and ETTD to improve understanding of how the subsurface environment controls hydrologic and transport behavior in a sub-humid mountain headwater catchment.

2. Materials and Methods

2.1. Study site

The proposed approaches for estimating TTD and ETTD with discontinuous time-series data were tested at Marshall Gulch Catchment (MGC; Figure 1), a 1.55 km^2 headwater catchment within the Santa Catalina Mountains Critical Zone Observatory near Tucson, Arizona, USA. The estimated average annual precipitation at MGC for 1981-2010 was ~920 mm [*PRISM Climate Group*, 2018], but the observed average between water years (WY) 2008 and 2017 was only 654 mm. Water year *n* is defined as the period from July 1 of year *n-1* through June 30 of year *n*. Surface elevations range from 2285 to 2632 m above sea level (asl) with a mean of 2428 m asl.

The mean topographic gradient is approximately 22°. Existing instrumentation provides spatially-distributed measurements of hydrologic and chemical fluxes (Figure 1). MGC is covered by Rocky Mountain aspen forest and woodlands (~32%), madrean upper montane conifer-oak forest and woodland (~28%) at upper elevations, and madrean pine-oak forest and woodlands (~40%) at lower elevations (GIS data obtained from *NatureServe* [2004]). The dominant tree species are Douglas-fir (*Pseudotsuga menziesii*), white fir (*Abies concolor*), ponderosa pine (*Pinus ponderosa*), aspen (*Populus tremuloides*) and bigtooth maple (*Acer gradidentatum*). Bedrock is mostly granite at upper elevations and micaceous schist at lower elevations [*Dickinson et al.*, 2002]. The prevailing soil type is sandy loam [*Holleran*, 2013] from 0 m to 1.5 m deep [*Pelletier and Rasmussen*, 2009]. Soils overlying schist are generally deeper and richer in clay than soils overlying granite [*Heidbüchel et al.*, 2013; *Holleran*, 2013].

2.2. Data sources

2.2.1. Hydrologic fluxes

We used MGC-scale daily precipitation (P), streamflow (Q) and actual evapotranspiration (AET) data between WY 2008 and 2017, (Figure 2A and B). Precipitation records from eight measurement sites (Fig. 1A) were combined into a single time-series using Thiessen polygon-derived weights [*Dwivedi et al.*, 2019b]. Streamflow was measured at 30-minute intervals at the MG-Weir site (Figure 1) using a pressure transducer (U20-001-01; Onset) with maximum error of 0.62 kPa, accuracy of 0.02 kPa, and a known stage-discharge relationship [*Heidbüchel et al.*, 2012]. AET was based on evapotranspiration (ET) measured at the nearby Mt. Bigelow US-MtB eddy covariance tower (see *Knowles et al.* [2020] for details). To obtain the catchment-scale daily AET time series, the observed ET time series was corrected using precipitation, streamflow, and soil-water storage data and a catchment-scale water balance method (Figure 2B) [*Dwivedi et al.*, 2019b; *Troch et al.*, 2017].

2.2.2. Stable water isotopes

2.2.2.1. Precipitation

Bulk precipitation samples were collected at the Schist, Fern Valley, and Granite bulk samplers, the Mt. Lemmon and MG-Weir ISCO autosamplers, and the Mt. Bigelow station (Figure 1A; for site details, see also *Dwivedi et al.* [2019b]). Samples from the MGC and from the Mt. Lemmon sites were analyzed using a DLT-100 laser spectrometer, Los Gatos Research, Inc., model # 908-171 0008 [*Lyon et al.*, 2009], with an analytical precision (1σ) of ±0.37‰ and ±0.12‰ for δ^2H and $\delta^{18}O$, respectively. Samples from Mt. Bigelow were analyzed on a L2120-I cavity ring-down spectrometer, Picarro, Inc., with analytical precision of ±0.20‰ for $\delta^{18}O$ and ±0.7‰ for δ^2H [*Johnson et al.*, 2017]. Measurements were standardized relative to international standards VSMOW and SLAP [*Coplen*, 1993].

Data were averaged arithmetically [*Heidbüchel et al.*, 2012] to yield a catchment-scale time series (Figure 2C; Section S1.1). Dataset intervals between WY 2008 and WY 2012 (reflecting many rainless days) were: minimum 1 d, mean 4 d, median 1 d, and maximum 78 d. For WYs 2015 and 2016, the dataset intervals were minimum 1 d, mean 15 d, median of 4 d, and maximum 280 d. For WY 2015 alone, the dataset had a minimum interval size of 1 d, a mean interval size of 8 d, a median interval size of 4 d, and a maximum interval size of 79 d.

2.2.2.2. Stream water

Samples were collected at the MG-Weir using an autosampler taking daily samples (including sub-daily samples during large runoff events) prior to 2012 and thereafter by grab sampling. Sub-daily samples were volume-weighted to daily resolution (Figure 2C). The stream is ephemeral; hence, the isotope time series has irregular intervals: minimum 1 d, mean 2 d, median 1 d, and maximum 50 d between WY 2008 and WY 2012.

2.2.2.3. Xylem water

Xylem water stable isotope data were collected, mostly bi-weekly, between July 16, 2014 and October 1, 2016 from Douglas-fir trees at Mt. Bigelow, and were analyzed using the L2130-I cavity ring-down spectrometer [*Hamann*, 2018]. For WY 2015, the dataset intervals were: minimum 13 d, mean 15 d, median 14 d, and maximum 35 d (Figure 2C).

2.3. Estimation of catchment-scale TTD and ETTD functions

Whereas spectral analysis is the preferred method for differentiating between TTD types [*McGuire and McDonnell*, 2006; *Onderka and Chudoba*, 2018], it can only be applied to stationary time series data [*Cottis et al.*, 2016]. Therefore, we employed the wavelet analysis method that is suited to both stationary and non-stationary time series data [*Farge*, 1992; *Onderka and Chudoba*, 2018; *Torrence and Compo*, 1998].

Following *McGuire and McDonnell* [2006], for inflow and outflow tracer fluxes that are related by a time-varying TTD, $h(t,\tau)$:

$$Q(t)c_Q(t) = \int_0^{\infty} h(t,\tau)w(t-\tau)c_I(t-\tau)d\tau \qquad (1)$$

where $Q(t)$ is the discharge flux at any time t, $c_I(t)$ is the input tracer parameter at time t, $c_Q(t)$ is the discharge tracer concentration at time t, and $w(t\text{-}\tau)$ is the "weighting term" [*McGuire and McDonnell*, 2006], considered as the amount of precipitation that contributes to outflow Q. For estimating the "weighting-term" i.e., the fraction of precipitation that contributes to streamflow or AET, two time-varying partitioning functions $FP_1(t)$ and $FP_2(t)$ are used. For streamflow, Equation (1) becomes:

$$Q(t)c_Q(t) = \int_0^{\infty} h(t,\tau)(FP_1(t-\tau)\,P(t-\tau))c_P(t-\tau)d\tau \qquad (2)$$

where P$(t\text{-}\tau)$ is the precipitation flux at time t-τ and $FP_1(t-\tau)$ is the partitioning function for precipitation that contributes to streamflow at time t-τ. Representing the tracer flux (the product of hydrologic flux and corresponding tracer concentration), as m, Equation (2) can be expressed as:

$$m_Q(t) = \int_0^\infty h(t,\tau) m_P(t-\tau) d\tau \tag{3}$$

Power spectra ($\hat{\varphi}$) of tracer flux in streamflow (or AET) and precipitation are estimated as functions of time (t) and period (λ). Letting $\hat{\varphi}_P(t,\lambda)$ and $\hat{\varphi}_Q(t,\lambda)$ be the time- and period-dependent power spectra of $m_P(t)$ and $m_Q(t)$, respectively, then the time-variant TTD can be estimated using Equation (4) by fitting the analytical solutions for various types of TTDs (or ETTDs), seeking the best match to the ratio of power spectra between inflow and outflow [*Godsey et al.*, 2010; *Kirchner et al.*, 2001]:

$$|\hat{H}_{TTD}(t,\lambda)|^2 = \frac{\hat{\varphi}_Q(t,\lambda)}{\hat{\varphi}_P(t,\lambda)} \tag{4}$$

Similarly, for ETTD:

$$|\hat{H}_{ETTD}(t,\lambda)|^2 = \frac{\hat{\varphi}_{AET}(t,\lambda)}{\hat{\varphi}_P(t,\lambda)} \tag{5}$$

For estimating the time-averaged TTD (or ETTD), all available power spectra are combined into global power spectra for inflow and outflow [*Torrence and Compo*, 1998] by taking their weighted mean (weighting by degrees of freedom) along the time axis, following the approach of *Kirchner and Neal* [2013]. Letting $\varphi_P(\lambda)$ and $\varphi_Q(\lambda)$ be the time-averaged power spectra of tracer flux in inflow and outflow respectively, then the final time-averaged $H_{TTD}(\lambda)$ or $H_{ETTD}(\lambda)$ are estimated using Equations (6) and (7):

$$|H_{TTD}(\lambda)|^2 = \frac{\varphi_Q(\lambda)}{\varphi_P(\lambda)} \tag{6}$$

$$|H_{ETTD}(\lambda)|^2 = \frac{\varphi_{AET}(\lambda)}{\varphi_P(\lambda)} \tag{7}$$

In this study, we tested two approaches for estimating time-averaged TTD. The first [*Godsey et al.*, 2010] uses power spectra of tracer concentration time series in inflow and outflow, while the second is based on tracer flux rather than concentration.

2.3.1. Method 1 (existing method), tracer concentrations

Previous research suggested that the TTD relating stream flow and precipitation can be estimated by: (i) determining the power spectra as a function of wavelength or period (λ) for the concentration time series of a conservative tracer, (ii) taking the ratio of the power spectra, and then (iii) selecting the TTD function whose power spectrum $|H(\lambda)|^2$ or the frequency content of a signal [*Ljung*, 2007], best fits the observed spectrum ratio (Equation 8) [*Godsey et al.*, 2010; *Kirchner*, 2016a; *Kirchner et al.*, 2001]:

$$|H_{TTD}(\lambda)|^2 = \frac{\varphi_{C_Q}(\lambda)}{\varphi_{C_P}(\lambda)} \tag{8}$$

where $\varphi_{C_Q}(\lambda)$ and $\varphi_{C_P}(\lambda)$ are the global power spectra of a conservative tracer in streamflow and precipitation and $|H_{TTD}(\lambda)|^2$ is the power spectrum of the TTD. Note that Method 1 does not yield an ETTD.

2.3.2. Method 2 (proposed method), tracer fluxes

An alternative approach allows for variable streamflow and subsurface storages (see Figure 2B and Figure 3 for the catchment-scale soil-water storage). Following previous approaches that use the convolution integral method in the time domain for TTD estimation [*Botter et al.*, 2010; 2011; *Hrachowitz et al.*, 2010; *Hrachowitz et al.*, 2011; *Hrachowitz et al.*, 2009], calculations are based on tracer fluxes rather than tracer concentrations. In equations (6) and (7), precipitation that contributes to Q or AET is estimated by the time-varying flow-partitioning functions $FP_1(t)$ and $FP_2(t)$:

$$FP_1(t) = \frac{Q_{long-term} + Q(t)}{P_{long-term} + P(t)} \quad (9)$$

$$FP_2(t) = \frac{AET_{long-term} + AET(t)}{P_{long-term} + P(t)} \quad (10)$$

In Equations (9) and (10), $P_{long-term}$, $Q_{long-term}$, and $AET_{long-term}$ are the long-term annual averages of precipitation, streamflow, and AET, whereas P(t), Q(t) and AET(t) are precipitation, streamflow and AET for any daily time step *t*. The time series of $FP_1(t)$ and $FP_2(t)$ for the whole period of record at MGC are shown in Figure 3.

The partitioning functions $FP_1(t)$ and $FP_2(t)$ in Method 2 are similar to the partitioning coefficients of *Botter et al.* [2010; 2011] and the "partial partition function" of *Harman* [2015] with a few differences. The partitioning coefficient in the *Botter et al.* [2010; 2011] approach is defined in terms of an infinite integral. Thus, for a particular flux, the partitioning coefficient depends on the entire hydrologic history of a water parcel from the moment it enters the catchment until discharge. The partitioning coefficient for any outflux is obtained using parameters from a numerical hydrologic model. The "partial partition function" of *Harman* [2015] represents a conditional probability for a water parcel to exit the catchment either as streamflow or as AET after entering the catchment as precipitation. The function is derived by model calibration to a stream water chloride time series.

In this work, $FP_1(t)$ and $FP_2(t)$ are estimated using observations, the time-varying runoff coefficient (Q/P) and the evaporative index (AET/P). To account for the hydrologic history before the study period, the long-term annual averages of inflow and outflow are used to estimate the time-varying $FP_1(t)$ and $FP_2(t)$. Such partitioning functions remain ≤ 1, consistent with the axioms of probability [*Devore*, 2008]. Thus, the probability that a water parcel will exit the catchment can be estimated without performing data-intensive numerical modeling. Moreover, as the sum of the annual AET and Q is not always P (see Figure 3B), any interactions between P, AET, Q, and soil and bedrock storages are also implicitly considered.

The present work estimates time-averaged TTDs or ETTDs, referred to as "marginal distributions" in *Benettin et al.* [2017] and "master TTDs" in *Heidbüchel et al.* [2012] and *Onderka and Chudoba* [2018]. Method 2 uses global power spectra for inflow and outflow rather than data collected on a precipitation event basis, as is customary when using the convolution integral method [*Harman*, 2015; *Heidbüchel et al.*, 2012]. The global power spectrum of a time series serves as an unbiased estimate of its true power spectrum [*Percival*, 1995; *Torrence and Compo*, 1998]. Using the amplitude ratio of sinusoidal tracer-concentration cycles can lead to aggregation biases in mean transit times (mTT) [*Kirchner*, 2016b], and large differences between estimated and true mTTs for a spatially heterogeneous landscape. Aggregation biases are reduced for hypothetical two-box non-stationary models (Figure 5d in *Kirchner* [2016b]). Since both Method 1 and Method 2 fit sinusoidal cycles, albeit with various periods, to the observed tracer fluxes in precipitation and streamflow, it is expected that the mTT from the global power spectrum ratio, which approaches the average behavior of a non-stationary system [*Kirchner*, 2016b], will be more representative of true catchment behavior than seasonal mTTs. Therefore, we propose the use of time-averaged response functions as a basis for practical examination of catchment hydrologic and transport behavior.

2.3.3. Spectral time series analysis

For Method 1 and new Method 2, Foster's [*Foster*, 1996] Weighted Wavelet Transform (WWT) method, adapted by *Kirchner and Neal* [2013], was used for estimating time- and period-dependent power spectra of $\delta^{18}O$ time series data as fitted sinusoidal cycles. *Zhang et al.* [2018] recommended the WWT method for analysis of time series data with irregular time steps.

To estimate the TTD using Method 1, power spectra were estimated in input and output $\delta^{18}O$ time series data. Given the irregularity of data intervals, we estimated Nyquist frequencies using the median time intervals [*Godsey et al.*, 2010], which was 1 d for both precipitation and stream water data; however, the fundamental frequency was based on the full analysis period, i.e., WY 2008 through WY 2012. For ETTD estimation, precipitation and xylem water time series data for WY 2015 only were utilized. The Nyquist frequency was based on the median time interval, 14 d, for $\delta^{18}O$ in xylem waters and the fundamental frequency was based on the full analysis period, WY 2015. For both TTD and ETTD, the limiting frequency was set to twice the fundamental frequency [*Godsey et al.*, 2010].

For Method 2, daily precipitation fluxes were either used as observed or combined into 7-d brackets when precipitation was irregularly distributed. Data from Mt. Bigelow were available as 14-d brackets. Daily streamflow and AET values were used to estimate tracer fluxes in these outflows. The power spectra of tracer concentrations or fluxes were computed at even multiples of the fundamental frequency up to the Nyquist frequency for both methods. Although $\delta^{18}O$ values are strictly not tracer concentrations, the mathematics for TTD estimation are the same for $\delta^{18}O$ values as for conservative tracer concentrations [*Kirchner*, 2016a]. As a result, we use the terms “tracer concentration” (Method 1) and “tracer flux” (Method 2) to describe the use of $\delta^{18}O$ as a tracer for TTD or ETTD estimation.

2.3.4. Data suitability assessment

To assess the applicability of the WWT method to irregular time series data, we followed *Zhang et al.* [2018] in generating self-similar, normally-distributed white, pink, and red noises with no data gaps. These data were generated at a daily time scale between WY 2008 and WY 2012 and for WY 2015 (solid curves in Figure 4). The expected spectral slopes (β_e) between power spectra and period in log-log space for white, pink, and red noises are 0, 1, and 2, respectively [*Witt and Malamud*, 2013]. These are referred to as regular time series data. Synthetic irregular time series data were also generated by deleting data from the regular time series in intervals for which data were not available for precipitation, stream water, or xylem water (points in Figure 4). The WWT method was applied to each regular and irregular time series, and the results were used for correlation analysis and prediction of spectral slopes (β_p) and standard errors. Following *Kleinbaum and Kupper* [1978], a t-statistic was used for statistical comparison (at the 95% confidence interval) of the predicted spectral slopes of regular and irregular time series data. As found by *Zhang et al.* [2018], the spectral slopes of synthetic noises were far from their expected values when the *Kirchner* [2005] filter was applied during benchmarking tests, but less so when the filter was not used (Table S1). Therefore, we did not apply the *Kirchner* [2005] filter in this study.

2.3.5. Optimization of model parameters using the Downhill Simplex method

The Downhill Simplex method was used for estimating the optimal model parameters for various TTD and ETTD types (for more details, see section S2). The Downhill Simplex method is a derivative-free local search method for estimating optimum model parameters using the model outputs closest to observations following pre-selected model performance criteria [*Gupta*, 2016; *Nelder and Mead*, 1965] within an allowed parameter space. The modified Kling Gupta efficiency or KGE’ [*Gupta et al.*, 2009; *Kling et al.*, 2012] is used as the criterion of best TTD fit [*Heidbüchel et al.*, 2012]:

$$KGE' = \sqrt{(\frac{Cov_{m,o}}{\sigma_m\sigma_o} - 1)^2 + (\frac{\sigma_m/\mu_m}{\sigma_o/\mu_o} - 1)^2 + (\frac{\mu_m}{\mu_o} - 1)^2} \quad (11)$$

where $Cov_{m,o}$ is the covariance between the modeled and observed time series, σ_m and σ_o represent one standard deviation, and μ_m and μ_o are the means of the modeled and observed time series, respectively. A KGE' value of zero indicates a perfect fit, and a value of ∞ indicates no fit. KGE' criteria were estimated in a log-transformed space [*Godsey et al.*, 2010]. A criterion response surface, which is the locus of all values of KGE' for a set of model parameters in the plausible parameter space was constructed for each TTD or ETTD [*Gupta*, 2016].

The types of TTDs or ETTDs that we considered included piston flow, exponential, gamma, one-dimensional fixed path advection-dispersion (ADE-1x), and multiple-path advection-dispersion (ADE-nx) models. Mathematical expressions for the models and their power spectra are given in Section S2. For the piston flow and exponential TTD (or ETTD) models, the mTT (mTT; τ_o) was the only fitted parameter. For the gamma model, the fitted parameters were shape (α) and mTT (τ_o; note that the scale parameter β for a gamma distribution is τ_o/α). For the ADE-1x and ADE-nx models, the average Peclet number (*Pe*), i.e., the ratio of advective to dispersive transport rates, and mTT were the fitted parameters. In this work, we considered the following model parameter ranges: (i) 0.1 to 5 years (5 years is the maximum applicable range when using stable isotopes as tracers) for mTT [*Godsey et al.*, 2010; *McGuire and McDonnell*, 2006; *Stewart et al.*, 2010], (ii) 0.1 to 10 for the gamma shape parameter (α), and 0.1 to 100 for the Peclet number (*Pe*), following evaluated in ranges considered applicable for solutes at catchment scale [*Kirchner and Neal*, 2013; *Kirchner et al.*, 2001].

3. Results

3.1. Evaluation of the irregular time series

3.1.1. Long time series; Nyquist frequency based on median daily sampling interval

The calculated spectral slopes (β_c) of the regular time series data were very close to their expected (β_e) values for all noise types when the Nyquist frequency was based on the median daily time step ("regular" rows in Table 1 A and B). However, for irregular data that mimicked actual data gaps (Figure 2C), the correlation coefficient between spectral power and period was reduced relative to the regular time series data. Further, the β_c for each noise type was lower than β_e and the difference between β_e and β_c depended on the noise type ("irregular" rows in Table 1). For example, using a pink noise ($\beta_e = 1$) time series based on precipitation $\delta^{18}O$, the values of R^2 and $|\beta_e - \beta_c|$, i.e., the absolute values of the difference between β_e and β_c, were 0.92 and 0.16 for the irregular time series, and 0.96 and 0.03 for the regular time series, respectively. Using red noise data ($\beta_e = 2$) based on precipitation $\delta^{18}O$, R^2 and $|\beta_e - \beta_c|$ increased to 0.97 and 0.50 for the irregular time series and 0.99 and 0.03 for the regular time series. Moreover, the calculated slope difference between the irregular and regular time series data for the same noise type depended on the degree of irregularity in the time series data and the noise type. For white and red noises, $|\beta_{c,regular} - \beta_{c,irregular}|$ was lower for stream water than for precipitation, but the opposite was true for pink noise. Overall, $|\beta_{c,irregular} - \beta_{c,regular}| \leq 0.5$.

3.1.2. Short time series; Nyquist frequency based on median bi-weekly sampling interval

Agreement between the expected and calculated slopes was generally poorer in these cases (Tables 1C and D), especially for the regular time series data. For aggregated data, e.g., precipitation $\delta^{18}O$, $|\beta_e - \beta_c|$ values based on regular time series data were 0.21, 0.25, and 0.59 for white, pink, and red noises, respectively. For instantaneous samples, e.g., xylem water $\delta^{18}O$, $|\beta_e - \beta_c|$ values based on regular time series data were 0.17, 0.44, and 0.02 for white, pink, and red noises, respectively.

3.1.3. Statistical comparison of calculated spectral slopes

For the daily median sampling intervals, the *t*-values were generally greater than the critical value, while for the bi-weekly median sampling intervals, the *t*-values were generally less than the critical value. In other words, the null hypothesis (that the calculated slopes of regular and irregular time series were statistically similar) was rejected with 95% confidence for all but one case with daily median sampling intervals, but could not be rejected for the cases with 14-d median sampling intervals. The effect of data gaps can therefore be evaluated for most of the datasets with 1-d median intervals, but not for the datasets with 14-d median intervals.

3.2. TTD estimation using $\delta^{18}O$ time series data with gaps; modeled tracer outflow

3.2.1. Estimated TTD and TTD parameters using Method 1 and Method 2

Both Method 1 and Method 2 indicated composite TTDs (Figure 5A1 and A2). For periods above 0.04 years with Method 1 and above 0.1 years with Method 2, the Gamma TTD type more closely matched observations within the permissible parameter space (Table 2; Figure 5A1; Figure 5A2). For Method 2, the ADE-nx TTD performed as well as the Gamma TTD type in terms of the KGE' criterion, but the estimated Pe parameter was at the edge of the allowable range, i.e., 0.1 to 10. The simpler Gamma distribution (Equation S2.5) was therefore preferred for periods above 0.1 years. For the Gamma TTDs, the mTTs were 1.21 years for Method 1, and 0.82 years for Method 2 (Table 2), and both mTTs were within the specified parameter range. The Gamma shape parameters (α) were 0.40 using Method 1 and 0.52 using Method 2.

For periods below 0.04 years with Method 1 and below 0.1 years with Method 2, both methods suggested piston flow power spectra (horizontal solid red lines in Figures 5A1 and A2). Therefore, we combined piston flow power spectra at lower periods with Gamma power spectra at higher periods using Equation S2.18. The dashed red lines in Figures 5A1 and 5A2 correspond to 5% piston flow and 95% Gamma TTD for Method 1 and 10% piston flow and 90% Gamma TTD for Method 2; these combinations provided acceptable matches to the observations. KGE' values were lower for the composite TTD than the Gamma TTD for the whole period range and for both methods (Table 2). Both methods produced smooth response surfaces with the best-fit Gamma TTD type, which indicates robust model optimization (Figures 5B1 and 5B2). When the optimization models were run three times for each TTD type, starting from different initial model parameters, the results for each run were not significantly different (Figures 5B1 and B2; Tables S2-S4). Both $\delta^{18}O$ and $\delta^{2}H$ produced similar results for each TTD type (Table S5).

A sensitivity analysis that considered the influence of spatio-temporal variability in precipitation and temporal variability in stream water on the estimated TTD type and TTD parameters for both methods confirmed that the Gamma TTD type most closely matched the observed power spectrum within the permissible parameter space (see Section S3). Method 2 was less sensitive to spatial and temporal variabilities than Method 1 (Section S3; Tables S6-S8). This was corroborated by the coefficients of variation of the mean age and α parameters, which were 13% and 3%, respectively, for Method 1 (Table S7) versus 4% and 0.2% for Method 2 (Table S8).

3.2.2. Modeled tracer outflow

Comparison of modeled tracer concentrations using the composite TTDs demonstrates that Method 1 did not accurately reproduce the observed concentrations. In contrast, tracer fluxes modeled with Method 2 were within their observed ranges (Figure 5C1 vs. C2). Recall that Method 1 uses tracer concentrations (shown as mean $\pm\sigma$ in the figure) whereas Method 2 uses tracer fluxes. We attribute the discrepancy between observed and modeled $\delta^{18}O$ in stream water (Fig. 5C1) to the lack of resolution as a result of aggregation for $\delta^{18}O$ in precipitation, in particular those precipitation $\delta^{18}O$ values that are associated with quick-flow responses, which may have had a significant impact on Method 1 modeling (see Section S4). With Method 2 and using the composite TTDs, the modeled tracer fluxes remained within the range of observed values, even beyond WY2012, except for large runoff events that were underestimated. The attenuation of large events may result from the averaging process inherent in the calculation of power spectra. The underestimated annual flux for WY 2014 may represent an artifact of low data density during that time. Nonetheless, in a general sense, the estimated composite TTDs produced outflow tracer fluxes that matched observations both within and beyond the TTD estimation period.

3.3. ETTD estimation using a $\delta^{18}O$ time series with data gaps using Method 2

3.3.1. Catchment-scale ETTD

Gamma and ADE-nx ETTD types performed equally well, although the slope estimation statistics were poorer for the limited ETTD dataset. Between the Gamma and ADE-nx ETTDs, the KGE' values for the ADE-nx and Gamma ETTDs were approximately similar, but the second parameter, Pe, for the ADE-nx ETTD (Pe = 100) was at the edge of the permissible parameter space, i.e., between 0.1 and 100 (Figure 6A). A composite ETTD type was investigated because the power spectrum ratios appeared to be horizontal up to a period of 0.1 years. However, a Gamma ETTD for all periods performed better (KGE' = 0.45) than composite ETTDs, e.g., piston flow plus Gamma for periods $\leq$ 0.1 years and Gamma for periods > 0.1 years (KGE' = 1.17).

The smooth response surface of the Gamma ETTD (Figure 6B; Figure S10) indicated that the model parameters were robust. Three model runs using different initial model parameters resulted in identical optimal model parameters (Tables S9-S11). When spatial and temporal variability in precipitation and temporal variability in xylem water were considered, a Gamma ETTD performed better than the ADE-nx ETTD types for some scenarios, whereas an ADE-nx ETTD performed better in others (section S5 and Table S13). Note that the stable water isotope data for precipitation were much sparser during the ETTD estimation period than between WY 2008 and WY 2012. To err on the conservative side, the current analysis used the largest standard deviation of all the temporal standard deviations for precipitation $\delta^{18}O$ between WY 2008 and WY 2012. For all scenarios, the estimated second parameter for the ADE-nx ETTD was at the edge of the permissible parameter space (Table S15). Therefore, a Gamma ETTD type was selected as a more appropriate model. The estimated mTT and α parameters for the Gamma ETTD were 0.65 years (σ = 0.05 years; coefficient of variation = 7%; Table S14) and 1.32 (unitless) (σ = 0.03 (unitless); coefficient of variation = 2 %; Table S15), respectively.

The estimated ETTD and ETTD parameters suggested minimal contribution of subsurface waters with residence time beyond the range of applicability of stable water isotope tracers. For example, the contribution of the subsurface waters with residence time beyond the 5 years range of applicability of stable water isotope tracers, was 0.020 ± 0.007% (Figure 6D).

3.3.2. Modeled tracer fluxes for the ETTD

Comparison of the observed and modeled AET tracer fluxes shows that the modeled fluxes were within the range of observations (Figure 6C). The sparseness as well as a larger temporal aggregation of the precipitation isotope data for the estimation period had a significant impact on estimated tracer fluxes in the obtained AET. Beyond the estimation period (WY 2015), the data were even sparser, leading to artifacts in the modeled tracer fluxes. However, the Gamma ETTD generally modeled outflow tracer fluxes that matched observations both within and beyond the ETTD estimation period.

4. Discussion

4.1. A practical approach for assessment of acceptable data gaps

As shown in Section 3.1, gaps in irregular time series data interacted to affect spectral analysis of various synthetic self-similar noise types. Thus, the proposed Method 2 can be applied to estimate TTD and ETTD if the power-spectrum slopes of the regular and irregular time series compare satisfactorily. Following [*Zhang et al.*, 2018], the comparison is considered satisfactory if the slopes can be distinguished statistically for a selected confidence interval, and if they differ by ≤ 0.5. When these criteria are satisfied, Method 2 can provide an adequate estimate of the spectral slope of real-world time series data with gaps. We compared slopes of irregular and regular time series (rather than with expected slopes) because the WWT method produces slopes that differ somewhat from the expected slopes, even for regular (ideal) time series (Table S1).

4.2. The estimated TTD at MGC

The composite TTD at MGC contributes to an improved understanding of the combined effects of diverse subsurface flow paths and storages that support streamflow. Previously, *Heidbüchel et al.* [2012] used stable water isotopes in precipitation and streamflow between 2007 and 2010, along with the Downhill Simplex method and the KGE' criterion, to represent subsurface processes at MGC with an exponential TTD, and determined a mTT of ~1.4 years. However, the exponential TTD is limited to representing the subsurface as a well-mixed linear reservoir [*Godsey et al.*, 2010; *McGuire et al.*, 2005; *van der Velde et al.*, 2014]. Using the same tracers, the current study determined that a combination of piston flow and Gamma TTDs yielded a mTT of 0.82 years at MGC (Figure 7A and B). Whereas the Piston flow TTD represents quick responses from short-term storages such as overland and near surface flow, the Gamma TTD represents a combination of quick responses from short-term storages and long-term memory of the subsurface reservoir [*Kirchner et al.*, 2001]. This framework is supported by recent work at MGC that used multiple tracers and end-member mixing analysis to show that near-surface flows and soil water storages are the dominant contributors to streamflow [*Dwivedi et al.*, 2019a], and thus that an incompletely mixed reservoir provides a more realistic model of the subsurface water storage than a well-mixed reservoir [*Sprenger and Allen*, 2020; *Sprenger et al.*, 2016; *Sprenger et al.*, 2018; *Zhao et al.*, 2013].

The estimated TTD at MGC suggests some contribution (about 3%) of subsurface waters with residence times greater than 5 years, beyond the applicability range of stable water isotope tracers (Figure 7B). Longer-term tracers would be required to model slower flow paths. In a study addressing a similar situation, *Stewart et al.* [2010] suggested including ^3H along with stable water isotope tracers for revealing "hidden streamflow" [*Stewart et al.*, 2012]. Using ^3H as a tracer, *Dwivedi et al.* [2019a] reported that some streamflow at MGC is supported by deep groundwater with mean residence times up to 16 years. Hence, the current study confirms that frequency-domain rather than time-domain observations may work better for differentiating TTD types. We attribute differences between the current and previous results at MGC to differences in the period of record and/or to our use of weighted wavelet analysis instead of the time-domain approach.

4.3. New insight into vegetation source waters

The ETTD emerging from this study was less statistically robust than the TTD, as indicated by scattered results for the power-spectrum slopes (section 3.2.2 and 3.2.3). However, the calculated ETTD confirmed that the water source for vegetation includes both short-term flow and storage, as well as longer-term storage in deep soil and fractured bedrock, especially by vegetation with tap roots [*Dwivedi et al.*, 2019b; *Oliver and Ryker*, 1990; *Wright*, 2001]. A more detailed understanding of the relationships between vegetation, water and streamflow could potentially be obtained by using a more robust ETTD, for which a longer, more complete dataset of xylem isotopes would be required.

Although the TTD and the ETTD derived here were both characterized by gamma distributions, the shape parameters differed (Figures 7B, 7C). The TTD shape parameter was less than one, indicating a long tail [*Godsey et al.*, 2010], whereas the ETTD shape parameter was greater than one, indicating a shorter ETTD tail. This suggests that vegetation at MGC has less access than stream water to groundwater of long residence time, which exemplifies a "good common sense" [*Klemes*, 1986] understanding of plant-accessible subsurface water.

Stable water isotopes are useful for identifying vegetation source water at MGC, consistent with studies elsewhere [*Bowling et al.*, 2017; *Dawson and Ehleringer*, 1991; *Dawson and Simonin*, 2011; *Dwivedi et al.*, 2019b; *Hartsough et al.*, 2008; *Jackson et al.*, 1999; *Li et al.*, 2007; *Meinzer et al.*, 1999; *Meißner et al.*, 2012; *Moreira et al.*, 2000; *Sprenger et al.*, 2016; *Sprenger et al.*, 2018; *Stratton and Meinzer*, 2000]. However, we acknowledge potential limitations of the isotope approach including a lack of distinctive isotopic signatures for all source waters, incomplete understanding of fractionation processes during plant water uptake, the spatial and temporal heterogeneity of natural systems, inconsistent methodologies for collecting and analyzing sap water and subsurface water storages, and differing subsurface water use (ecohydrological niches) among various vegetation types, across broad climatic gradients, and within single catchments [*Allen et al.*, 2018; *Lin and Sternberg*, 1993; *Penna et al.*, 2018; *Silvertown et al.*, 2015; *Vargas et al.*, 2017]. Consequently, clear research objectives, consideration of the scale of analysis, and careful evaluation of local vegetation characteristics are advised when using Method 2.

5. Conclusions

The modified weighted wavelet transform method presented by the current study was useful for estimating catchment-scale TTDs and ETTDs from realistic tracer and hydrologic flux time series datasets with gaps. Further, it is practical because it requires no intensive computation of flow and transport models, exploits commonly available data in hydrologic inflows and outflows, and can be tested for resistance to data gaps. We specifically recommend application of the proposed method in situations when the power-spectrum slopes of regular (no gaps) and irregular (gaps matching those of the dataset) synthetic time series can be statistically distinguished for a selected confidence interval, and when they differ by ≤ 0.5.

The relationship between streamflow and soil water storage at the Marshall Gulch, Arizona, USA catchment (MGC) was best described by a combination of Piston flow and Gamma TTDs at short periods (below 0.1 years) and a Gamma TTD at longer periods. Overall, the mTT was 0.8 years and the scale parameter was 0.5 (unitless). The TTD indicated small contributions of subsurface water with a residence time greater than 5 years. Stable water isotopes in precipitation and xylem water supported a Gamma ETTD, which is useful for tracing principal vegetation source waters at MGC. The proposed method sheds light on how tracer input fluxes that vary in space and time are modulated by subsurface flow paths of different length and transit time, and how such fluxes support evapotranspiration and streamflow at multiple time scales.

Acknowledgements, Samples, and Data

We thank Drs. James Kirchner, Milan Onderka, and Qian Zhang for methodological guidance and feedback on an earlier manuscript version (Kirchner), and Colin Clark and Joe Gibney for help with the analytical solution of the ADE-nx equation. We also thank Dr. J. McIntosh for helpful comments on the preliminary version of this paper. We appreciated equipment support from Alissa Coes and isotope data and support from Nicole Weber, Adam Weber, Matej Durcik, Tim Corley, and Ingo Heidbüchel. We acknowledge National Science Foundation Grants EAR-1331408 and EAR-0724958, and NSF career award EAR-1255013/1755722 to Dr. Shirley (Kurc) Papuga. R.D. acknowledges additional funding from a Geological Society of America Graduate Student Research Grant, a Horton Research Grant, and Graduate and Professional Student Council Research and Travel grants from the University of Arizona. T.M., T.F., and R.D. were supported by a Water Resources Research Center 104(b) grant. The corresponding author thanks his wife, Jessie Dwivedi, and their children, Darcy and Noah Dwivedi, for their patience and support. All data used in this work can be accessed from: http://www.czo.arizona.edu/data-catalog/, https://www.geo.arizona.edu/node/154, or from the supporting information of this work. The authors declare no financial conflicts of interest.

References

Allen, S. T., J. W. Kirchner, S. Braun, R. T. W. Siegwolf, and G. R. Goldsmith (2018), Seasonal origins of soil water used by trees, *Hydrology and Earth System Sciences Discussions*, 1-23.

Benettin, P., C. Soulsby, C. Birkel, D. Tetzlaff, G. Botter, and A. Rinaldo (2017), Using SAS functions and high resolution isotope data to unravel travel time distributions in headwater catchments, *Water Resources Research*, *53*, 1864-1878.

Berghuijs, W. R., and J. W. Kirchner (2017), The relationship between contrasting ages of groundwater and streamflow, *Geophysical Research Letters*, *44*, 8925-8935.

Birkel, C., C. Soulsby, D. Tetzlaff, S. Dunn, and L. Spezia (2012), High-frequency storm event isotope sampling reveals time-variant transit time distributions and influence of diurnal cycles, *Hydrological Processes*, *26*(2), 308-316.

Botter, G., E. Bertuzzo, and A. Rinaldo (2010), Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, *Water Resources Research*, *46*(W03514), 1-18.

Botter, G., E. Bertuzzo, and A. Rinaldo (2011), Catchment residence and travel time distributions: The master equation, *Geophysical Research Letters*, *38*(11), n/a-n/a.

Bowling, D. R., E. S. Schulze, and S. J. Hall (2017), Revisiting streamside trees that do not use stream water: can the two water worlds hypothesis and snowpack isotopic effects explain a missing water source?, *Ecohydrology*, *10*(1), e1771.

Brooks, J. R., H. R. Barnard, R. Coulombe, and J. J. McDonnell (2009), Ecohydrologic separation of water between trees and streams in a Mediterranean climate, *Nature Geoscience*, *3*(2), 100-104.

Coplen, T. B. (1993), Reporting of stable carbon, hydrogen, and oxygen isotopic abundance, paper presented at Reference and intercomparison materials for stable isotopes of light elements (1-3 December), International Atomic Energy Agency, Vienna, Austria.

Cottis, R. A., A. M. Homborg, and J. M. C. Mol (2016), The relationship between spectral and wavelet techniques for noise analysis, *Electrochimica Acta, 202*, 277-287.

Dawson, T. E., and J. R. Ehleringer (1991), Streamside trees that do not use streamwater, *Nature, 350*(6316), 335-337.

Dawson, T. E., and K. A. Simonin (2011), The Roles of Stable Isotopes in Forest Hydrology and Biogeochemistry, in *Forest hydrology and biogeochemistry: Synthesis of Past Research and Future Directions*, edited by D. F. Levia, D. E. Carlyle-Moses and T. Tanaka, pp. 137-161, Springer.

Dickinson, W. R., D. M. Hirschberg, G. Pitts, G. Stephen, and S. K. Bolm (2002), Spatial digital database of the geologic map of Catalina Core Complex and San Pedro Trough, Pima, Pinal, Gila, Graham, and Cochise Counties, Arizona, 1-25 pp, U.S. Geological Survey, Menlo Park, CA.

Dwivedi, R., T. Meixner, J. McIntosh, P. A. T. Ferré, C. J. Eastoe, G.-Y. Niu, R. L. Minor, G. Barron-Gafford, and J. Chorover (2019a), Hydrologic functioning of the deep Critical Zone and contributions to streamflow in a high elevation catchment: testing of multiple conceptual models, *Hydrological Processes, 33*, 476-494, doi: 10.1002/hyp.13363.

Dwivedi, R., C. Eastoe, J. F. Knowles, W. E. Wright, L. Hamann, R. Minor, B. Mitra, T. Meixner, J. McIntosh, P. A. T. Ferre, C. Castro, G.-Y. Niu, G. A. Barron-Gafford, N. Abramson, S. A. Papuga, M. Stanley, J. Hu, and J. Chorover (2019b), Vegetation source water identification using isotopic and hydrometric observations from a subhumid mountain catchment, *Ecohydrology*, doi: 10.1002/eco.2167.

Farge, M. (1992), Wavelet transforms and their applications to Turbulence, *Annual Review of Fluid Mechanics 24*, 395-458

Feng, X., J. W. Kirchner, and C. Neal (2004), Spectral Analysis of Chemical Time Series from Long-Term Catchment Monitoring Studies: Hydrochemical Insights and Data Requirements, in *Biogeochemical Investigations of Terrestrial, Freshwater, and Wetland Ecosystems across the Globe*, edited by R. K. Wieder, M. Novák and M. A. Vile, pp. 221-235, Springer-Science + Business Media, LLC, the Netherlands.

Foster, G. (1996), Wavelets for period analysis of unevenly sampled time series, *The Astronomical Journal*, *112*(4), 1709-1729.

Godsey, S. E., W. Aas, T. A. Clair, H. A. de Wit, I. J. Fernandez, J. S. Kahl, I. A. Malcolm, C. Neal, M. Neal, S. J. Nelson, S. A. Norton, M. C. Palucis, B. L. Skjelkvåle, C. Soulsby, D. Tetzlaff, and J. W. Kirchner (2010), Generality of fractal 1/f scaling in catchment tracer time series, and its implications for catchment travel time distributions, *Hydrological Processes*, *24*(12), 1660-1671.

Gupta, H. V. (2016), Fundamentals – Systems Approach to Hydrological Modeling (HWRS 528), edited, The University of Arizona, Tucson, AZ.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, *377*(1-2), 80-91.

Hamann, L. (2018), Temporal Dynamics of Tree Source Water in Sky Island Ecosystems with Ephemeral Snow Pack: A Case Study using Pseudotsuga Meziesii (Douglas Fir), 131 pp, The University of Arizona, Tucson.

Harman, C. J. (2015), Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, *Water Resources Research*, *51*, 1-30.

Hartsough, P., S. R. Poulson, F. Biondi, and I. G. Estrada (2008), Stable Isotope Characterization of the Ecohydrological Cycle at a Tropical Treeline Site, *Arctic, Antarctic, and Alpine Research*, *40*(2), 343-354.

Heidbüchel, I., P. A. Troch, and S. W. Lyon (2013), Separating physical and meteorological controls of variable transit times in zero-order catchments, *Water Resources Research*, *49*(11), 7644-7657.

Heidbüchel, I., P. A. Troch, S. W. Lyon, and M. Weiler (2012), The master transit time distribution of variable flow systems, *Water Resources Research*, *48*(6), 1-19.

Hirsch, R. M., D. L. Moyer, and S. A. Archfield (2010), Weighted regression on time, discharge and season (WRTDS) with an application to Chesapeake bay rivers inputs, *Journal of the American Water Resources Association (JAWRA)*, *46*(5), 857-880.

Holleran, M. E. (2013), Quantifying catchment scale soil variability in Marshall Gulch, Santa Catalina Mountains Critical Zone Observatory, 102 pp, The University of Arizona, Tucson.

Hrachowitz, M., C. Soulsby, D. Tetzlaff, and M. Speed (2010), Catchment transit times and landscape controls—does scale matter?, *Hydrological Processes*, *24*, 117-125.

Hrachowitz, M., C. Soulsby, D. Tetzlaff, and I. A. Malcolm (2011), Sensitivity of mean transit time estimates to model conditioning and data availability, *Hydrological Processes*, *25*(6), 980-990.
Hrachowitz, M., C. Soulsby, D. Tetzlaff, J. J. C. Dawson, and I. A. Malcolm (2009), Regionalization of transit time estimates in montane catchments by integrating landscape controls, *Water Resources Research*, *45*(5), doi: 10.1029/2008wr007496.
Jackson, P. C., F. C. Meinzer, M. Bustamante, G. Goldstein, A. Franco, P. W. Rundel, L. Caldas, E. Igler, and F. Causin (1999), Partitioning of soil water among tree species in a Brazilian Cerrado ecosystem, *Tree Physiology*, *19*, 717-724.
Johnson, J. E., L. Hamann, D. L. Dettman, D. Kim-Hak, S. W. Leavitt, R. K. Monson, and S. A. Papuga (2017), Performance of induction module cavity ring-down spectroscopy (IM-CRDS) for measuring delta(18) O and delta(2) H values of soil, stem, and leaf waters, *Rapid communications in mass spectrometry : RCM*, *31*(6), 547-560.
Kirchner, J. W. (2016a), Aggregation in environmental systems- Part 1: Seasonal tracer cycles quantify young water fractions, but not mean transit times, in spatially heterogeneous catchments, *Hydrology and Earth System Sciences*, *20*(1), 279-297.
Kirchner, J. W. (2016b), Aggregation in environmental systems -Part 2: Catchment mean transit times and young water fractions under hydrologic nonstationarity, *Hydrology and Earth System Sciences*, *20*(1), 299-328.
Kirchner, J. W., and C. Neal (2013), Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection, *Proceedings of the National Academy of Sciences of the United States of America*, *110*(30), 12213-12218.
Kirchner, J. W., X. Feng, and C. Neal (2001), Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations, *Journal of Hydrology*, *254*, 82-101.
Kleinbaum, D. G., and L. L. Kupper (1978), *Applied regression analysis and other multivariable methods*, 556 pp., Duxbury Press, North Scituate, Massachusetts.
Klemes, V. (1986), Dilettantism in Hydrology: Transition or Destiny, *Water Resources Research*, *22*(9), 177S-188S.
Kling, H., M. Fuchs, and M. Paulin (2012), Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, *Journal of Hydrology*, *424-425*, 264-277.
Knowles, J. F., R. L.Scott, R. L.Minor, and G. A.Barron-Gafford (2020), Ecosystem carbon and water cycling from a sky island montane forest, *Agricultural and Forest Meteorology*, *281*(107835).
Li, S.-G., H. Romero-Saltos, M. Tsujimura, A. Sugimoto, L. Sasaki, G. Davaa, and D. Oyunbaatar (2007), Plant water sources in the cold semiarid ecosystem of the upper Kherlen River catchment in Mongolia: A stable isotope approach, *Journal of Hydrology*, *333*(1), 109-117.
Lin, G., and L. d. S. L. Sternberg (1993), Hydrogen Isotopic Fractionation by Plant Roots during Water Uptake in Coastal Wetland Plants, in *Stable Isotopes and Plant Carbon-Water Relations*, edited by J. R. Ehleringer, A. E. Hall and G. D. Farquhar, pp. 497-510, Academic Press, Inc.
Ljung, L. (2007), System Identification*Rep. LiTH-ISY-R-2809*, Automatic Control at Linköpings universitet.
Lyon, S. W., S. L. E. Desilets, and P. A. Troch (2009), A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus δ18O, *Hydrological Processes*, *23*(14), 2095-2101.
McGuire, K. J., and J. J. McDonnell (2006), A review and evaluation of catchment transit time modeling, *Journal of Hydrology*, *330*(3-4), 543-563.
McGuire, K. J., J. J. McDonnell, M. Weiler, C. Kendall, B. L. McGlynn, J. M. Welker, and J. Seibert (2005), The role of topography on catchment-scale water residence time, *Water Resources Research*, *41*(W05002), 1-14.
Meinzer, F. C., J. L. Andrade, G. Goldstein, N. M. Holbrook, J. Cavelier, and S. J. Wright (1999), Partitioning of soil water among canopy trees in a seasonally dry tropical forest, *Oecologia*, *121*, 293-301.
Meißner, M., M. Köhler, L. Schwendenmann, and D. Hölscher (2012), Partitioning of soil water among canopy trees during a soil desiccation period in a temperate mixed forest, *Biogeosciences*, *9*(8), 3465-3474.
Milly, P. C. D., and K. A. Dunne (2020), Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation, *Science*, doi: 10.1126/science.aay9187.
Moreira, M. Z., L. d. S. L. Sternberg, and D. C. Nepstad (2000), Vertical patterns of soil water uptake by plants in a primary forest and an abandoned pasture in the eastern Amazon: an isotopic approach, *Plant and Soil*, *222*, 95-107.
NatureServe (2004), Southwest Regional GAP Analysis Project—Land Cover Descriptions. RS/GIS Laboratory, 1-178 pp, College of Natural Resources, Utah State University, Logan, UT.
Nelder, J. A., and R. Mead (1965), A simplex method for function minimization, *The Computer Journal*, *7*(4), 308-313.
Oliver, W. W., and R. A. Ryker (1990), Pinus pmderosa Dougl. ex Laws., in *Silvics of North America*, edited by R. M. Burns and B. H. Honkala, United States Department of Agriculture, Washington, DC.

Onderka, M., and V. Chudoba (2018), The Wavelets show it – the transit time of water varies in time, *Journal of Hydrology and Hydromechanics*, *66*(3), 295-302.
Pelletier, J. D., and C. Rasmussen (2009), Geomorphically based predictive mapping of soil thickness in upland watersheds, *Water Resources Research*, *45*(9), 1-15.
Penna, D., L. Hopp, F. Scandellari, S. T. Allen, P. Benettin, M. Beyer, J. Geris, J. Klaus, J. D. Marshall, L. Schwendenmann, T. H. M. Volkmann, J. von Freyberg, A. Amin, N. Ceperley, M. Engel, J. Frentress, Y. Giambastiani, J. McDonnell, G. Zuecco, P. Llorens, R. T. W. Siegwolf, T. E. Dawson, and J. Kirchner (2018), Tracing ecosystem water fluxes using hydrogen and oxygen stable isotopes: challenges and opportunities from an interdisciplinary perspective, *Biogeosciences Discussions*, 1-23, doi: 10.5194/bg-2018-286.
Percival, D. P. (1995), On Estimation of the Wavelet Variance, *Biometrika*, *82*(3), 619-631.
PRISM Climate Group (2018), PRISM Gridded Climate Data for the Location: Lat: 32.4263 Lon: -110.7610 Elev: 7703ft, edited, Oregon State University.
Reynolds, J. F., D. M. S. Smith, E. F. Lambin, B. L. T. II, M. Mortimore, S. P. J. Batterbury, T. E. Downing, H. Dowlatabadi, R. J. Fernández, J. E. Herrick, E. Huber-Sannwald, H. Jiang, R. Leemans, T. Lynam, F. T. Maestre, M. Ayarza, and B. Walker (2007), Global Desertification: Building a Science for Dryland Development, *Science*, *316*(11 May), 847-851.
Silvertown, J., Y. Araya, D. Gowing, and W. Cornwell (2015), Hydrological niches in terrestrial plant communities: a review, *Journal of Ecology*, *103*(1), 93-108.
Sprenger, M., and S. T. Allen (2020), What Ecohydrologic Separation Is and Where We Can Go With It, *Water Resources Research*, *56*(7), doi: 10.1029/2020wr027238.
Sprenger, M., H. Leistert, K. Gimbel, and M. Weiler (2016), Illuminating hydrological processes at the soil-vegetation atmosphere interface with water stable isotopes, *Reviews of Geophysics*, *54*, 674-704.
Sprenger, M., D. Tetzlaff, J. Buttle, H. Laudon, H. Leistert, C. P. J. Mitchell, J. Snelgrove, M. Weiler, and C. Soulsby (2018), Measuring and Modeling Stable Isotopes of Mobile and Bulk Soil Water, *Vadose Zone Journal*, *17*(1), 0, doi: 10.2136/vzj2017.08.0149.
Stewart, M. K., U. Morgenstern, and J. J. McDonnell (2010), Truncation of stream residence time: how the use of stable isotopes has skewed our concept of streamwater age and origin, *Hydrological Processes*, *24*(12), 1646-1659.
Stewart, M. K., U. Morgenstern, J. J. McDonnell, and L. Pfister (2012), The 'hidden streamflow' challenge in catchment hydrology: a call to action for stream water transit time analysis, *Hydrological Processes*, *26*(13), 2061-2066.
Stratton, L. C., and G. G. F. C. Meinzer (2000), Temporal and Spatial Partitioning of Water Resources among Eight Woody Species in a Hawaiian Dry Forest, *Oecologia*, *124*(3), 309-317.
Torrence, C., and G. P. Compo (1998), A practical guide to wavelet analysis, *Bulletin of the American Meteorological Society*, *79*(1), 61-78.
Troch, P., R. Dwivedi, T. Liu, A. Meira, T. Roy, R. Valdés-Pineda, M. Durcik, S. Arciniega, and J. A. Brena-Naranjo (2017), Catchment-scale groundwater recharge and vegetation water use efficiency, Abstract [H51B-1260], in *2017 Fall Meeting (11-15 Dec)*, *AGU*, edited by AGU, New Orleans, LA.
U.S. Geological Survey (2018), USGS NED 1 arc-second n33w111 1 x 1 degree ArcGrid 2018, edited, U.S. Geological Survey.
van der Velde, Y., P. J. J. F. Torfs, S. E. A. T. M. van der Zee, and R. Uijlenhoet (2012), Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, *Water Resources Research*, *48*(6).
van der Velde, Y., I. Heidbüchel, S. W. Lyon, L. Nyberg, A. Rodhe, K. Bishop, and P. A. Troch (2014), Consequences of mixing assumptions for time-variable travel time distributions, *Hydrological Processes*.
Vargas, A. I., B. Schaffer, L. Yuhong, and L. Sternberg (2017), Testing plant use of mobile vs immobile soil water sources using stable isotope experiments, *The New phytologist*, *215*(2), 582-594.
Viviroli, D., H. H. Dürr, B. Messerli, M. Meybeck, and R. Weingartner (2007), Mountains of the world, water towers for humanity: Typology, mapping, and global significance, *Water Resources Research*, *43*(7), 1-13.
Witt, A., and B. D. Malamud (2013), Quantification of Long-Range Persistence in Geophysical Time Series: Conventional and Benchmark-Based Improvement Techniques, *Surveys in Geophysics*, *34*(5), 541-651.
Wright, W. E. (2001), Delta-deuterium and delta-oxygen-18 in mixed conifer systems in the United States southwest: The potential of delta-oxygen-18 in Pinus ponderosa tree rings as a natural environmental recorder, 328 pp, The University of Arizona, Tucson, AZ.
Zeng, N., and J. Yoon (2009), Expansion of the world's deserts due to vegetation-albedo feedback under global warming, *Geophysical Research Letters*, *36*(17), 1-5.
Zhang, Q., C. J. Harman, and J. W. Kirchner (2018), Evaluation of statistical methods for quantifying fractal scaling in water-quality time series with irregular sampling, *Hydrology and Earth System Sciences*, *22*(2), 1175-1192.

Zhao, P., X. Tang, P. Zhao, C. Wang, and J. Tang (2013), Identifying the water source for subsurface flow with deuterium and oxygen-18 isotopes of soil water collected from tension lysimeters and cores, *Journal of Hydrology*, *503*, 1-10.

List of Figures and Tables:

Figure 6. (A) Observations (black points) and fitted power spectra (dimensionless) of various ETTD types: Exponential (Exp), Gamma (Gam), ADE-1x, ADE-nx, and a composite ETTD with 10% Piston Flow and 90% Gamma up to a period ≤ 0.1 years and 100% Gamma for periods > 0.1 years. (B) Response surface plots for the best fitting Gamma ETTD type. (C) Observed and modeled (mean ± 1σ) $M_{AET}=\delta^{18}O x AET$ in AET. (D) The best-fit Gamma ETTD type as a function of residence time of evapo-transpired water. The table in (A) shows 1σ (in parentheses) for the Gamma ETTD type based on the spatial and temporal variability in precipitation and temporal variability in xylem water isotopic composition (see section S4) where parameter 1 is the mean age (in years), and parameter 2 (not applicable for the piston flow TTD type) is the shape parameter (dimensionless) for the Exp and Gamma distributions and the Peclet number for the ADE-1x and ADE-nx ETTD types. The Gamma ETTD parameters in the composite ETTD in (A) are estimated using the observation data for periods >0.1 years. The 1σ error bars in (C) are visible only if they are larger than the point symbol.

Figure 7. A conceptual model for MGC that illustrates the composite TTDs for short (0.1 years) (A) and long (B) periods, and (C) the applicable catchment-scale ETTD. (D) A conceptual model of MGC showing subsurface storages that support streamflow and AET (modified from *Dwivedi et al.* [2019a] and *Dwivedi et al.* [2019b]).

Table 1. Correlation coefficients, p-values, spectral slopes, standard errors of calculated slopes, and t- and t-critical (t_{crit}) values of the estimated slopes between spectra and period in log-log space for various synthetic self-similar normally distributed noise types and considering regular and irregular time series data. The irregular time series data mimic gaps in the observed $\delta^{18}O$ time series (Figure 2).

Table 2. Estimated model parameters (columns 2 and 3) for various stream response function types (column 1) with their final estimated KGE' value (column 4) for periods from 0.04 to 5 years (Method 1) and 0.1 years to 5 years (Method 2). Parameter 1 represents mean age (years) for all tested TTD types (except piston flow TTD). Parameter 2 is the shape parameter (α, dimensionless) for the exponential type (set to 1) and gamma type TTDs, and the Peclet number for the ADE-1x or ADE-nx TTDs; parameter 2 is not applicable for the piston flow TTD. Numbers in brackets for the Gamma TTD type show 1σ based on spatial and temporal variability of $\delta^{18}O$ in precipitation and temporal variability in streamflow (see section S2). A low value of KGE' corresponds to a better model fit to the observations.