

Water Resources Research®



RESEARCH ARTICLE

10.1029/2022WR033033

Key Points:

- Process-based tracer-aided ecohydrological modeling is upscaled to >100 km² catchments using stable water isotopes
- Isotopes benefit large-scale modeling in substantially improving model robustness and reliability of water age estimates
- Larger-scale water partitioning and drought responses are controlled by heterogeneity in catchment organization

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

X. Yang,
xiaoqiang.yang@ufz.de

Citation:

Yang, X., Tetzlaff, D., Müller, C., Knöller, K., Borchardt, D., & Soulsby, C. (2023). Upscaling tracer-aided ecohydrological modeling to larger catchments: Implications for process representation and heterogeneity in landscape organization. *Water Resources Research*, 59, e2022WR033033. <https://doi.org/10.1029/2022WR033033>

Received 13 JUN 2022
Accepted 14 FEB 2023

Author Contributions:

Conceptualization: Xiaoqiang Yang, Doerthe Tetzlaff, Dietrich Borchardt, Chris Soulsby

Data curation: Christin Müller, Kay Knöller, Dietrich Borchardt

Formal analysis: Xiaoqiang Yang

Funding acquisition: Doerthe Tetzlaff, Dietrich Borchardt

Investigation: Xiaoqiang Yang, Doerthe Tetzlaff, Dietrich Borchardt, Chris Soulsby

Methodology: Xiaoqiang Yang, Chris Soulsby

© 2023. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Upscaling Tracer-Aided Ecohydrological Modeling to Larger Catchments: Implications for Process Representation and Heterogeneity in Landscape Organization

Xiaoqiang Yang^{1,2} , Doerthe Tetzlaff^{2,3,4}, Christin Müller⁵ , Kay Knöller⁵, Dietrich Borchardt¹ , and Chris Soulsby^{2,4}

¹Department of Aquatic Ecosystem Analysis and Management, Helmholtz Centre for Environmental Research – UFZ, Magdeburg, Germany, ²Department of Ecohydrology and Biogeochemistry, Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany, ³Department of Geography, Humboldt University of Berlin, Berlin, Germany, ⁴Northern Rivers Institute, School of Geosciences, University of Aberdeen, Aberdeen, UK, ⁵Department of Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale), Germany

Abstract Stable isotopes of water are ideal tracers to integrate into process-based models, advancing ecohydrological understanding. Current tracer-aided ecohydrological modeling is mostly conducted in relatively small-scale catchments, due to limited tracer data availability and often highly damped stream isotope signals in larger catchments (>100 km²). Recent model developments have prioritized better spatial representation, offering new potential for advancing upscaling in tracer-aided modeling. Here, we adapted the fully distributed EcH₂O-iso model to the Selke catchment (456 km², Germany), incorporating monthly sampled isotopes from seven sites between 2012 and 2017. Parameter sensitivity analysis indicated that the information content of isotope data was generally complementary to discharge and more sensitive to runoff partitioning, soil water and energy dynamics. Multi-criteria calibrations revealed that inclusion of isotopes could significantly improve discharge performance during validations and isotope simulations, resulting in more reasonable estimates of the seasonality of stream water ages. However, capturing isotopic signals of highly non-linear near-surface processes remained challenging for the upscaled model, but still allowed for plausible simulation of water ages reflecting non-stationarity in transport and mixing. The detailed modeling also helped unravel spatio-temporally varying patterns of water storage-flux-age interactions and their interplay under severe drought conditions. Embracing the upscaling challenges, this study demonstrated that even coarsely sampled isotope data can be of value in aiding ecohydrological modeling and consequent process representation in larger catchments. The derived innovative insights into ecohydrological functioning at scales commensurate with management decision making, are of particular importance for guiding science-based measures for tackling environmental changes.

1. Introduction

Ecohydrological partitioning strongly influences the provision of water resources in terms of quantity, quality, location, and timing; thus, fundamentally affecting water-related ecosystem services (Brauman, 2015). In larger scale catchments (>100 km²), ecohydrological functioning often exhibits strong spatio-temporal variability due to small-scale process heterogeneity and non-linearity, both of which result in often marked catchment-wide contrasts in landscape organization (Hunt et al., 2021; Skjøien et al., 2003; Tetzlaff et al., 2011). Process-based models have been widely adopted for capturing such catchment functional heterogeneity, as well as its responses to environmental changes (Thirel et al., 2015). Efforts of the catchment modeling community have been progressively improving the physical realism of process conceptualization (Beven, 2002; Clark et al., 2017; Tetzlaff et al., 2008). As part of this trend, stable isotopes of water (i.e., $\delta^2\text{H}$ and $\delta^{18}\text{O}$), as ideal tracers, have received more attention and been increasingly integrated into process-based modeling (Birkel & Soulsby, 2015). As complementary to the use of more traditional hydrometric data (Fenicia et al., 2008), isotopic tracer information has provided more direct evidence of the velocity of water particle transport and allowed inference of catchment-scale water age distributions (McDonnell & Beven, 2014; Soulsby, Birkel, Geris, Dick, et al., 2015).

In ecohydrological modeling, isotope information has been applied in explicit tracking of pathways of water partitioning within the soil-plant-atmosphere system (Fatichi et al., 2016; Tetzlaff et al., 2021). A full-integration of hydrometric and tracer data in ecohydrological model development and testing, as argued by McDonnell and Beven (2014), potentially paves the way to fundamentally improving our understanding of ecohydrological

Project Administration: Doerthe Tetzlaff, Dietrich Borchardt
Resources: Doerthe Tetzlaff, Dietrich Borchardt
Software: Xiaoqiang Yang
Validation: Xiaoqiang Yang, Doerthe Tetzlaff, Christin Müller, Kay Knöller, Dietrich Borchardt, Chris Soulsby
Visualization: Xiaoqiang Yang, Chris Soulsby
Writing – original draft: Xiaoqiang Yang
Writing – review & editing: Xiaoqiang Yang, Doerthe Tetzlaff, Dietrich Borchardt, Chris Soulsby

processes, in terms of blue (i.e., runoff generation and groundwater recharge) –green (i.e., evaporation and transpiration) water partitioning, storage-flux interactions and associated timescales. Recent strategies of integrating tracers to track detailed flux dynamics in process-based modeling can simultaneously capture velocity (of water particles) and celerity (speed of response) timescales, and more particularly for fully distributed modeling, allowing spatially explicit descriptions of storage-flux-age interactions (Smith, Tetzlaff, Maneta, et al., 2022). The fully distributed ecohydrological EcH₂O model (Maneta & Silverman, 2013) implements advanced integrations of energy balance, water transfer and vegetation dynamics. The EcH₂O-iso model (Kuppel et al., 2018) is an adaptation that further incorporates general tracer concentrations, that is, water isotopes and age tracking, into each of the ecohydrological compartments and enables explicit spatio-temporal representation of water movement across a catchment. The EcH₂O-iso model has been intensively used to explore spatio-temporal interactions between hydrological cycling and plant functioning (Kuppel et al., 2020; Smith et al., 2020), as well as their responses to environmental changes (Neill et al., 2021; Yang et al., 2021).

Of course, utilizing isotope tracers in modeling studies is strongly dependent on the availability of adequate monitoring data sets (Birkel & Soulsby, 2015; McDonnell & Beven, 2014). However, intensive monitoring has mostly been deployed at smaller scales, such as representative experimental plots (Beyer et al., 2020; Landgraf et al., 2022) or small headwater catchments with a drainage area <10 km² (Fenicia et al., 2008; Kuppel et al., 2018). At these smaller scales, isotopic signals in different compartments of the ecohydrological systems can be well captured by logistically feasible, intensive monitoring schemes. Meanwhile, small-scale catchments also allow a finer-resolution of spatio-temporal modeling scales, which are often needed to match the scale of localized hydrological processes (McDonnell et al., 2021) and related isotopic transformations through mixing and fractionation (Iorgulescu et al., 2007; Smith et al., 2021). As catchment size increases, water sources and mixing processes become more complex and heterogeneous; as a result, stream water isotopic signals often become highly damped (compared to highly variable precipitation input signals) and less directly reflecting small-scale hydrological processes (IAEA, 2012). These factors have hindered the upscaling of advanced tracer-aided modeling from small-scale, intensively monitored catchments to larger catchments (i.e., >100–1,000s km²), even though current model developments (like EcH₂O-iso) have prioritized improved spatial representations of catchment characteristics and process heterogeneity regardless of scales (Smith et al., 2021).

At larger scales, isotope monitoring is increasingly becoming available, although inevitably with coarser sampling resolutions (e.g., monthly grab sampled at locations with averaged drainage area >100 km² (Ala-aho et al., 2018; Gibson et al., 2021; IAEA, 2012; Soulsby, Birkel, Geris, & Tetzlaff, 2015; Wollschläger et al., 2016)). Such data sets have been successfully used in aiding larger scale hydrological modeling (e.g., the isoWATFLOOD modeling by Stadnyk et al. (2013)), which largely emphasizes the value of isotope information in improving model calibration/parameterization, thereby inferring improved internal processes and catchment-scale representation of storage-flux dynamics (Holmes et al., 2020; Nan et al., 2021). Given the emerging potential of upscaling of tracer-aided models like EcH₂O-iso to mesoscale catchments (e.g., Smith et al. (2021) in a 66 km², intensively monitored lowland catchment, northern Germany), it remains to be seen whether further upscaling can be supported by spatio-temporally coarser resolution isotope data in larger (>100 km²), highly heterogeneous catchments. Furthermore, it is unclear what additional insights into detailed ecohydrological process representation and heterogeneous catchment functioning can be derived from upscaled tracer-aided modeling. Such evidence-based insights into larger-scale catchment functioning are of particular relevance to water management and decision making, especially for urgent needs of more efficient mitigation measures for water-related problems under the changing conditions (e.g., the recent European-wide droughts).

Here, we used a multi-site isotope monitoring data set from the Bode catchment (ca. 3,200 km²), central Germany. The data set consists of 27 precipitation sites and 24 stream water sites, sampled on a monthly basis (Lutz et al., 2018). We adapted and upscaled the tracer-aided EcH₂O-iso model to apply to the representative Selke sub-catchment (456 km²), where stream water isotopes were sampled over 2012–2017 from seven sites encompassing the main stem and major tributaries of the Selke River. We particularly addressed the following research questions:

1. By integrating a cross-scale isotope data set with the EcH₂O-iso model, what is the potential, and what are the key challenges, for process representation in the upscaling of advanced tracer-aided ecohydrological modeling in larger catchments?
2. Through comprehensive sensitivity analysis and multi-criteria calibration, what is the added value of using cross-scale stream isotopes in providing insight into modeled behaviors?

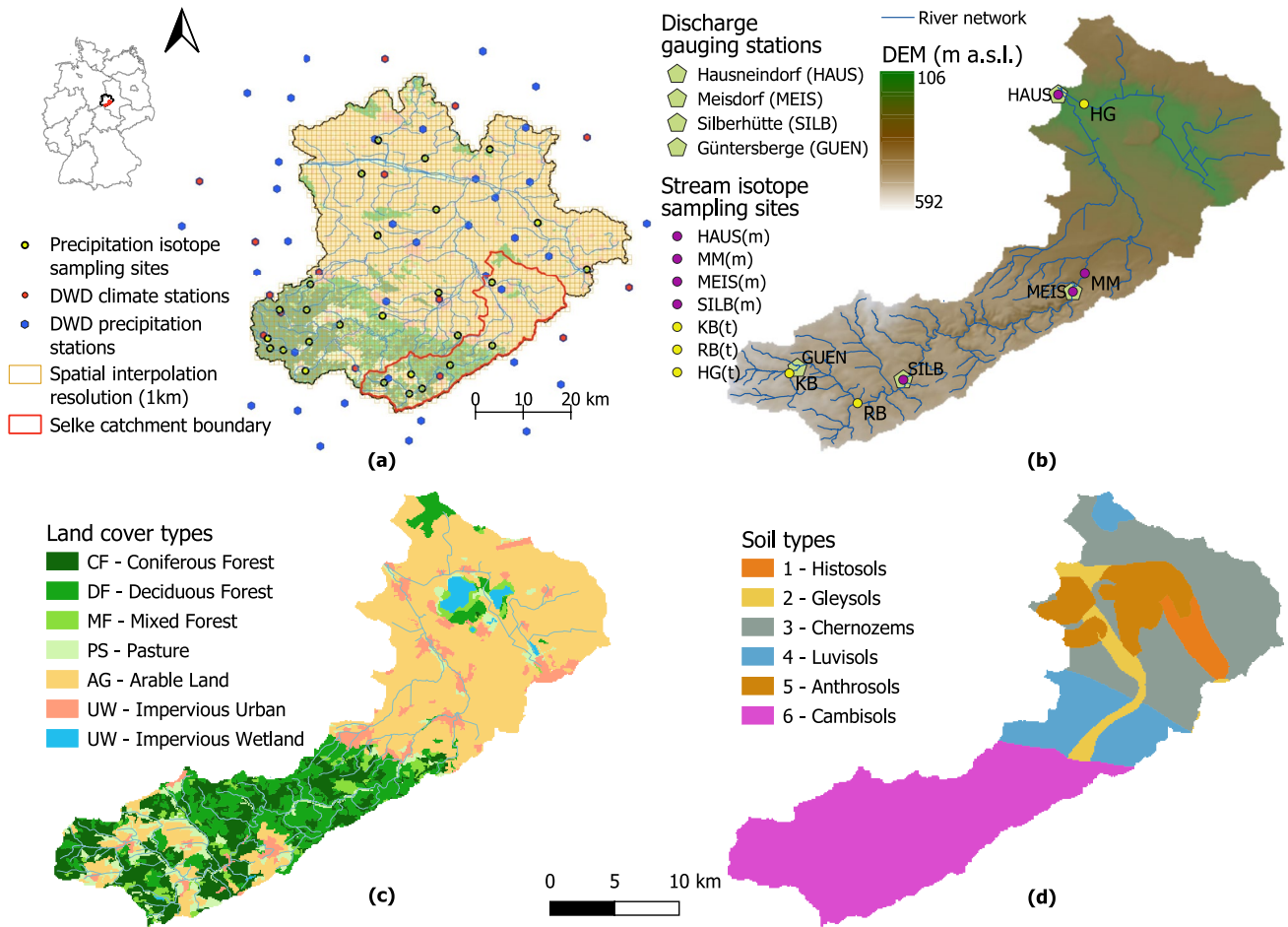


Figure 1. The monitoring networks of the Bode catchment ((a), sampling sites for precipitation isotopes and the German Weather Service—DWD stations for precipitation and climate data) and the Selke sub-catchment ((b), four discharge gauging stations and seven stream water isotope sampling sites). The geographic characteristics of the Selke catchment for the EcH_2O -iso model setup ((b)—Digital Elevation Model, (c)—land cover types, and (d)—soil types). The stream water isotope sampling covers four nested sites in the main stem of the Selke River (i.e., the discharge gauges HAUS, MEIS and SILB, and the MM site near station MEIS) and three tributaries (i.e., KB-Katzsohlbach, RB-Rodelbach, and HG-Hauptseegraben). The land cover (c) and soil type (d) abbreviations were also used for categorized model parameters.

3. Informed by large-scale, process-based modeling, how do the heterogeneous catchment characteristics control ecohydrological functioning and its responses to prolonged droughts?

With this comprehensive modeling and catchment process analysis, we anticipated that the detailed model could inform implications for larger-scale monitoring and catchment water management that are increasingly critical given the implications of climate change.

2. Study Catchment and Data

2.1. The Bode-Selke Catchment

The Bode catchment (ca. 3,200 km²) is a part of the TERrestrial ENvironmental Observatories—TERENO network in Central Germany (Wollschläger et al., 2016) and contributes to the Modular Observation Solutions for Earth Systems—MOSES initiative, which employs specifically designed monitoring modules to investigate the interactions of short-term events and long-term trends across earth compartments with special emphasis on hydrological extremes (Weber et al., 2022). It is located in the transitional region between the Harz Mountain uplands (up to 1,142 m above sea level) and the northern lowland plain (<100 m a.s.l., Figure 1a). It exhibits high gradients of hydro-climatic and physiographic characteristics. Annual precipitation decreases

from >1,700 mm in the Harz Mountains to less than 500 mm in the lowlands, and mean annual temperature increases from 5.0°C to 9.5°C (Wollschläger et al., 2016). The upper Harz Mountain areas are dominated by spruce and beech forests, with significant agricultural areas in the southern lower mountains; while the lowland areas are characterized by intensive arable lands with major crops of winter wheat, winter rapeseed, winter barley and sugar beet (because of highly fertile chernozem soils developed on loess deposits (Altermann et al., 2005)) and urban areas. Climate data were collected from DWD (German Weather Service) climate stations within and surrounding the Bode catchment, including 51 stations for daily precipitation and 15 stations for daily air temperature (daily mean, maximum, and minimum), wind speed and humidity (Figure 1a).

Our modeling focused on the Selke catchment (Figure 1b), a 456 km² sub-catchment located in the south part of the Bode catchment. The Selke catchment exhibits heterogeneous conditions, draining the lower Harz Mountain range to the lowlands, and so is broadly representative of the whole Bode catchment. Along with the topographic gradients, annual precipitation decreases from circa 790 to 450 mm, and mean temperature increases from ca. 4.7°C to 8.5°C (from DWD). The catchment mean annual potential evaporation is ca. 786 mm, and annual runoff depth is 111 mm at the outlet (1993–2020, the State water authority—LHW, Saxony-Anhalt). Contrasting landscape characteristics are well captured by three nested hydrological gauging stations (Figures 1b and 1c): the uppermost mountainous areas (above station Silberhütte—SILB, 99 km²) are dominated by 60% forests (coniferous and deciduous trees) and 36% agricultural lands (arable and pastures); the middle mountainous areas (between stations SILB and Meisdorf—MEIS, 85 km²) are predominated by forests (ca. 88%); and the lowland areas (between station MEIS and the outlet Hausneindorf—HAUS, 272 km²) comprise of 77% arable lands, 13% low density forests and 8% impervious urban/lake areas. There is a mining lake located close to the lowland tributary Hauptsee Graben—HG (Figure 1c). The soil types (Figure 1d, adapted from BGR (2018)) change from mountainous cambisols, to mountain foot luvisols, and to loess lowland chernozems (with patches of histosols, gleysols, and highly sandy anthrosols). Here, we defined the drainage areas above the MEIS gauging station as the upland areas (184 km², dominated by cambisols overlaying the shallow schist/claystone bedrock) and the rest as the lowland areas (272 km², dominated by chernozems developed on the loess deposits). In line with recent European-wide changing climate, the catchment has experienced severe droughts in 2018–2019 (Yang, Rode, et al., 2022; Yang et al., 2021), exhibiting lower precipitation and higher air temperature than averages (Figure S1 in Supporting Information S1).

2.2. Multi-Site Isotope Sampling Data Sets

We used precipitation isotope data sampled from 27 sites throughout the Bode catchment (Figure 1a) between 2012 and 2017. Monthly composite precipitation samples were collected from evaporation-protected precipitation collectors. Stream water samples were grab samples collected monthly from 24 sites in the Bode catchment. For details of the monitoring scheme, please refer to Lutz et al. (2018). We exclusively used four of these sites located within the Selke catchment (Figure 1b, upland tributary sites Katzsohlbach—KB and Rodelbach—RB, lowland tributary site Hauptsee Graben—HG, and main stem site—MM close to the MEIS discharge station). In addition, monthly grab samples were also collected from the three main stem discharge gauging stations between 2013 and 2015 (Figure 1b. Data source: Michael Rode, UFZ; personal communication). Stable water isotope values are reported in per mil (‰, relative to the Vienna Standard Mean Ocean Water standard isotope ratio). Laboratory analysis used a laser-based analyzer (L2130-i, Picarro Inc.; analytical precision of 1.0‰ and 0.3‰ for δ²H and δ¹⁸O, respectively) at the Department of Catchment Hydrology, Helmholtz Centre for Environmental Research—UFZ.

Following earlier work (Yang et al., 2021), we defined four hydrological seasons as Wet (January–March), Drying (April–June), Dry (July–September), and Wetting (October–December). Precipitation isotopes showed clear seasonal patterns, that is, more depleted and more enriched in heavier isotopes during wet-drying and dry-wetting seasons, respectively (Figure 2a); The signal of the stream water isotopes was more seasonally damped at each sampling site, and the isotopes were mostly located around the Local Meteoric Water Line (δ²H = 7.3 × δ¹⁸O + 3.9), except the samples at lowland tributary site HG and the dry-wetting season samples at RB and HAUS (Figure 2b). A more detailed overview of the sampled data set is provided in Text S1 in Supporting Information S1.

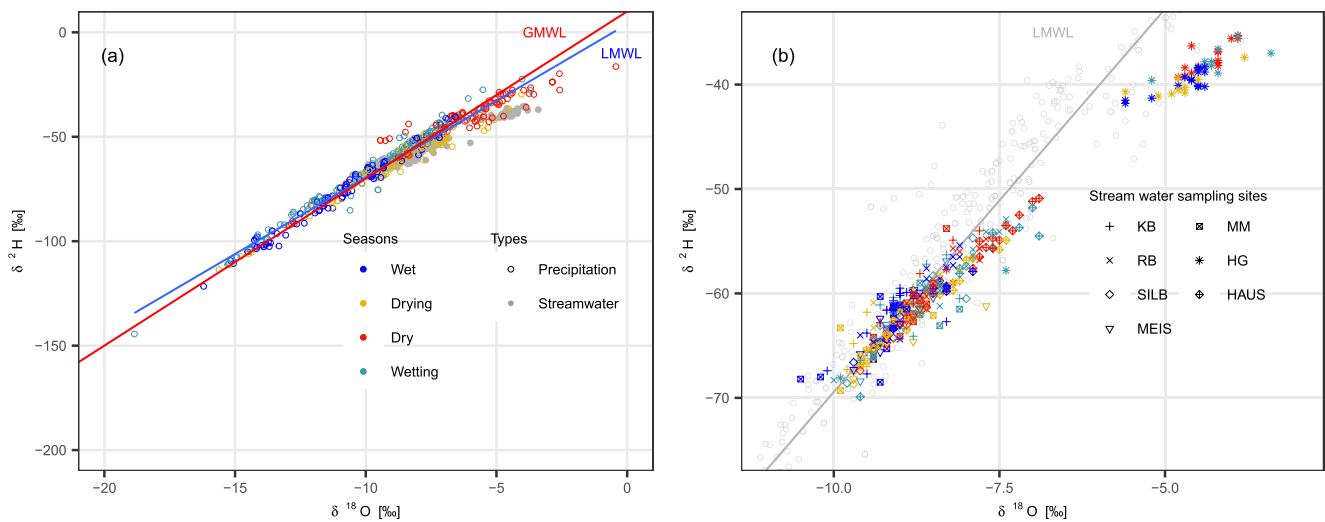


Figure 2. Dual-isotope plots of the sampled (a) precipitation and (b) stream water isotopes in the Selke catchment. The gray circles in panel (b) refer to the precipitation isotopes. LMWL and GMWL represent the Local Meteoric Water Line and the Global Meteoric Water Line, respectively.

3. Methods

3.1. The Ech_2O -Iso Model

The tracer-aided model, Ech_2O -iso (Kuppel et al., 2018), is developed based on the fully distributed process-based ecohydrological model Ech_2O (Maneta & Silverman, 2013). The model integrates modules of energy balance, water balance, vegetation dynamics, and flux tracking (evolving isotope compositions and water ages as tracers through model compartments). Here, we provide only a brief introduction of water balance and tracer related model structures (see the schematic diagram in Figure S2 in Supporting Information S1). For complete model details, please refer to Maneta and Silverman (2013) and Kuppel et al. (2018).

Driven by climatic forcing inputs (i.e., radiation, air temperature, humidity, and wind speed), the energy balance is solved at the canopy and the land surface levels, deriving latent heat fluxes that determine evapotranspiration and snowmelt. Canopy and soil evaporation are constrained by the canopy water storage and first-layer soil water storage, respectively. Transpiration is constrained by the canopy level environmental factors that are implemented via a Jarvis-type stomatal conductance model, with vegetation-dependent parameters representing stomatal properties (Table 1). The transpiration flux is further partitioned to the three soil layers based on soil water storages and root fraction distributions. Please refer to Text S2 in Supporting Information S1 for detailed descriptions of the evaporation and transpiration process conceptualization. The water balance uses linear bucket-type approaches for multiple water storages, including interception, snowpack, surface ponding, and soil water. Throughfall after exceeding canopy interception capacity and direct precipitation enter any surface snowpack or ponding storage, where snowmelt and infiltration processes are also considered. Surface infiltration into the first soil layer is calculated using the Green-Ampt equation (Mein & Larson, 1973), and further infiltrating into deeper layers once the upper layers get saturated. The vertical leakage from the third soil layer enters into the deeper groundwater storage implemented by Yang et al. (2021), which also considers hydrologically active storage and additional passive storage for solute-mixing.

Lateral exchanges (and runoff generation where grid cells link to the channel) are enabled at three levels, that is, the surface ponding storage (surface overland flow), the gravitational storage of the third soil layer (groundwater flow), and the deeper groundwater storage (deeper GW flow). The two subsurface flows are calculated using linear kinematic methods and transported to corresponding storage of the downstream grid cell, while ponded surface water is transported as the surface overland flow at the end of each time step. For grid cells where a channel is present, groundwater and deeper GW contributions to the channel are calculated using an exponential decay function of active storage with channel-seepage parameters. Similar for channel-connected cells, any surface ponding becomes surface overland flow to the channel. Stream water is routed along the river network using a kinematic-wave approach considering the channel roughness. Key model parameters are

Table 1
Brief Descriptions of the Key Parameters of the EcH_2O -Iso Model

Parameter	Description	Initial range
Ks_S	Effective horizontal soil hydraulic conductivity (m s^{-1})	[1e-05, 0.01]
$Aniso_S$	Ratio of vertical to horizontal hydraulic conductivity	[0.01, 0.5]
$B\lambda_S$	Soil pore size distribution parameter of the Brooks-Corey method	[3, 12]
$PsiAE_S$	Soil air entry pressure (m)	[0.05, 0.9]
$snowmeltCoe_S$	Snowmelt coefficient	[1e-09, 2e-07]
$leakance_S$	Vertical leakage to the deeper groundwater storage	[1e-06, 5e-03]
$Sheatcap_S$	Heat capacity of soil solid particles ($\text{J m}^3 \text{K}^{-1}$)	[1e + 04, 1e + 08]
$Sroughns_S$	Local surface roughness (m)	[0.1, 0.4]
$HyFrac_d_g_w_S$	Hydrologically active Fraction of the deeper groundwater storage	[1e-05, 0.1]
$CWSmax_L$	Maximum canopy water storage capacity (m)	[1e-04, 0.1]
$gsmax_L$	Maximum stomatal conductance (m s^{-1})	[1e-03, 0.05]
$gs_vpd_coeff_L$	Stomatal sensitivity to vapor pressure deficit	[1e-05, 0.01]
$gs_light_coeff_L$	Stomatal sensitivity to light	[1, 500]
$gs_psi_c_L$	Stomatal sensitivity to soil moisture content	[0.5, 10]
$gs_psi_d_L$	Soil moisture suction potential at which stomatal function is reduced by 50%	[0.01, 8.5]
$Kroot_L$	Exponential parameter controlling the root distributions in different soil layers	[0.1, 5]
$KBeers_L$	Light extinction coefficient as per Beer's law	[0.1, 1]
$ChanGWSeep_S$	Coefficient of groundwater seepage to channel	[0.5, 20]
$ChanDGWSeep_S$	Coefficient of deeper groundwater seepage to channel	[0.01, 10]
$ManningRiv_S$	River routing parameter related to channel roughness (Manning's n)	[1, 20]

Note. The labels “_L” and “_S” indicate land cover type and soil type dependent parameters, respectively, and should be replaced for each type as abbreviated in Figures 1c and 1d. Note that the channel seepage and routing parameters were assigned as soil-type dependent in this study.

listed in Table 1 with initial ranges set based on reference values from EcH_2O documentation (<https://ech2o.readthedocs.io>, last accessed 13 September 2022) and EcH_2O -iso applications in a similar northern German catchment (Smith et al., 2021). Please also refer to the associated open data (Yang et al., 2022b) for the full parameter list.

The tracer metrics (i.e., isotopic composition and water ages) are fully integrated into the ecohydrological cycling based on the full-mixing assumption in each storage compartment (Kuppel et al., 2018). Water age is treated as a conservative tracer concentration and the age of each storage increases by one after each time-step. All tracer concentrations are updated at each time step according to inflows and previous-step storage, weighted by the corresponding water amount; the concentrations of outflows are equal to that of the storage. Throughfall water is assumed to have the same isotopic composition and age (i.e., zero) as that of incoming precipitation. Isotopic fractionation by evaporation is considered in the first soil layer using the Craig-Gordon fractionation model (Craig & Gordon, 1965) with modifications for saturation in a porous medium, detailed in Kuppel et al. (2018).

Transpiration flux is contributed from the three soil layers based on vegetation-dependent root distributions (parameter $Kroot$, Table 1) and availability of soil water content (Text S2 in Supporting Information S1). The transpiration water ages are calculated as the flux-weighted means over the soil layers. Here, we further evaluated the transpiration Exit Age Ratio (EAR; adapted from Soulsby et al. (2016)) calculated as the ratio of the transpiration water age and the age of soil-stored water (i.e., the volume-weighted mean age over the soil layers). The EAR ($\in [0, 1]$) indicates the transpiration water age extraction characteristics, with lower values highlighting heterogeneity of source water ages (root water uptake typically from the upper, younger soil water) and higher values highlighting the preference of tapping into well-mixed, less-variable stores (Kuppel et al., 2020).

3.2. Ech_2O -Iso Adaptations for Larger Scale Modeling

The Ech_2O -iso ecohydrological model has mostly been applied to intensively monitored plots or smaller experimental catchments (up to 66 km²). Hence, some of the technical aspects of implementation needed to be adjusted for larger-scale catchments and the coarser spatio-temporal resolution modeling at these larger scales:

1. An input map of areal proportion of each modeling cell was added, given that the boundary cells might not be fully covered by the catchment domain.
2. Topographic characteristics (i.e., slope, drainage network, and stream morphology) could not be derived directly from a resampled digital elevation model (DEM) due to the coarse resolution at the modeling grid scale (e.g., 1 km²). Therefore, this study retrieved the information from Yang et al. (2018) in the modeled Selke catchment using multi-scale routing model implementations (Samaniego et al., 2010): first, the main drainage line of each modeling cell was derived according to the flow direction (D8 algorithm) and flow accumulation at the finer resolution DEM (e.g., 100 m); Second, the drainage direction at the 1 km² model resolution was assigned as the flow direction of the outlet of the main drainage line, and the slope was calculated as the elevation difference between the inlet and the outlet, divided by the main drainage length.
3. For a better representation of spatial variability, multiple soil types and land cover types were allowed for each modeling cell (rather than one soil or land cover type for each cell). This was supported by using input maps of type-specific areal proportions (Smith et al., 2021). Hydrological parameters were assigned as soil-type dependent, and the values in each grid cell were aggregated using the area-weighted geometric means if multiple soil types were present.
4. For grid cells with channels, the surface ponded water was partitioned into surface overland flow draining to the down-slope cell and surface overland flow to the channel, according to their relative lengths. Importantly, re-infiltration of any surface run-on into the first soil layer and the subsequent vertical tracer mixing (for both water isotopes and age) throughout the soil profile was activated, especially as the modeling resolution was coarse and the river network was relatively dense.

3.3. Sensitivity Analysis and Multi-Criteria Calibration

Sensitivity analysis was used for parameter ranking and diagnosing the information content of multiple data sources in revealing modeling behavior. The Morris method (Morris, 1991) was used: the elementary effect (EE—ratio between model output variation and parameter variation) of each parameter was calculated by changing one parameter at a time (i.e., one trajectory); then, multiple trajectories were repeated with randomly generated starting points, and the average of absolute EEs for each parameter was taken as the global sensitivity index. With low computation requirement, the Morris method is particularly suitable for screening and ranking purposes (Pianosi et al., 2016; Yang et al., 2019). The parameter sets were sampled using the radial-based Latin-Hypercube sampling strategy, due to its efficiency and coverage within the sampling space (Campolongo et al., 2011).

The most sensitive parameters were used in multi-criteria model calibrations using the Monte Carlo based method of Ala-aho et al. (2017) (see also flowchart in Figure S3 in Supporting Information S1). A large number of model runs (400,000 for this study) was first conducted with randomly sampled parameter sets. For each evaluation data source, model performance metrics were calculated for each model run and ranked among all runs (i.e., the data-specific empirical cumulative distribution function—CDF). A threshold quantile was iteratively determined among the CDFs of all data types and corresponding evaluation metrics, above which the common best 100 runs were detected. This method avoids potential uncertainty of objective aggregation between the multi-criteria and is flexible for analyzing the value of specific data in constraining the model performance (Yang et al., 2021). Moreover, the large number of model runs can be easily parallelized, which significantly reduces the computing time.

The evaluation metrics for sensitivity analysis were Kling-Gupta Efficiency—KGE (Gupta et al., 2009) for discharge data and Mean Absolute Error (MAE) for stream water $\delta^2\text{H}$. The MAE was used for isotopes because the coarse (~monthly) sampling had limited information on the short-term variability of isotope dynamics. The global sensitivity indices were calculated from 100 trajectories, as the maximum number suggested by Pianosi et al. (2016). For the multi-criteria calibration, we used log-transformed Nash-Sutcliffe Efficiency (NSE) and percentage bias of the water balance (PBIAS) for discharge, to ensure comparable weights for both high- and low-flow periods (Krause et al., 2005) and the overall water balance performance, respectively, and MAE for stream isotopes ($\delta^2\text{H}$). In total, 400,000 model runs were computed, and a two-stage calibration strategy similar

to Yang et al. (2021) was adopted (Figure S3 in Supporting Information S1): first, the best 40,000 runs according only to discharge metrics were selected to ensure the overall water balance modeling; the second stage selected the best 100 runs out of these 40,000 runs, using discharge data alone (the “discharge-only” calibration) and both discharge and isotope data (the “discharge plus isotope” calibration). All computations were performed at the High-Performance Computing Cluster EVE, a joint effort of both the Helmholtz Centre for Environmental Research—UFZ and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig.

3.4. Model Set-Up and Data Processing

The daily EcH_2O -iso modeling (spatial resolution of $1 \text{ km} \times 1 \text{ km}$) was setup in the Selke catchment (Table S1 in Supporting Information S1) during 2012–2019. Topographic characteristics (i.e., slope and drainage network at the $1 \text{ km} \times 1 \text{ km}$ resolution) were derived from Yang et al. (2018) based on the original 100 m DEM. The Selke river network was retrieved from the State water authority—LHW (<https://lhw.sachsen-anhalt.de/>, last accessed on 13 September 2022). Stream channel length was calculated directly from the LHW river network, and channel width was estimated using the empirical equation (i.e., $5.4 \cdot Q_m^{0.5}$ (Rode et al., 2016), where Q_m denotes the annual mean discharge obtained from Yang et al. (2018) for each stream reach).

To make most use of the whole monitoring data sets, model climate inputs (spatial resolution of $1 \text{ km} \times 1 \text{ km}$) were prepared at the Bode catchment level. Gridded precipitation input was spatially interpolated using the kriging method with elevation as external drift, while for other climate inputs, the Inverse Distance Weighted (IDW) method was used due to the relatively low number of stations (the “*hydroTSM*” package in *R* software (Zambrano-Bigiarini, 2020)). Incoming short- and long-wave radiation measurements are rarely available for DWD stations, therefore the daily radiation input data were retrieved from the ERA5 hourly reanalysis data with $0.25^\circ \times 0.25^\circ$ spatial resolution (Hersbach et al., 2018), downloaded from the Copernicus Climate Change Service Climate Data Store.

The daily gridded precipitation isotope inputs were prepared based on the isotope monitoring data set from the Bode catchment. First, the multi-site monthly composite data were spatially interpolated for each 1 km^2 grid cell. For months having measurements from more than 13 sites, the kriging method drifted by elevation was used; for the remaining months, the IDW method was used instead. Second, the temporal interpolation in each grid cell was conducted in two steps: (a) given the seasonality of isotope signals, mean and standard deviation were calculated separately for January–December and weighted by recorded precipitation amount; and (b) the daily values were then randomly generated following the normal distributions. The interpolated daily and sampled monthly $\delta^2\text{H}$ data fit well with each other, and comparisons at the Selke sampling sites were provided in Figure S4 in Supporting Information S1.

Daily discharge data at HAUS, MEIS, and SILB in 2012–2014 and 2018 were used for calibration, covering both normal (before 2018) and recent drought (2018 and 2019) years. The discharge data at the above three stations in 2015–2017 and 2019 were used for temporal validation, and the whole period data at GUEN for spatial validation. Stream isotope data from six sites (i.e., HAUS, MM, MEIS, SILB, KB, and RB) in 2012–2014 and 2015–2017 were used for stream isotope calibration and validation, respectively. Additionally, EcH_2O -iso requires initial conditions of water storages, as well as tracer concentrations and ages therein, therefore, the period 2010–2011 was used as the model warming-up period to stabilize these initial inputs.

4. Results

4.1. Model Parameter Sensitivity Variations Based on Discharge and Isotope Data at Multiple Sites

The sensitivity of the EcH_2O -iso parameters was ranked separately for discharge and isotope data at individual stations and then cross compared (Figure 3). For the upland discharge and isotope data (i.e., above the MEIS/MM station), the parameter rankings were generally similar, with the common sensitive parameters related to canopy interception (*CWSmax*), topsoil hydraulic properties (*Ks*), snowmelt (*snowmeltCoe*) and leakage to the deeper groundwater (*leakance*); interestingly, parameters related to vegetation stomatal properties (*gsmax*, *gs_light_coeff*, and *gs_vpd_coeff*) and soil field capacity (*BCLambda*) were more sensitive based on discharge data, while parameters related to soil heat fluxes (*Sheatcap*), subsurface lateral fluxes (*HyFrac_dgw*, *ChanDGWSeep*, and *ChanGWSeep*) and vertical soil hydraulic conductivity (*Aniso*) were more sensitive based on isotope data. The rankings also exhibited slight variations at different upland

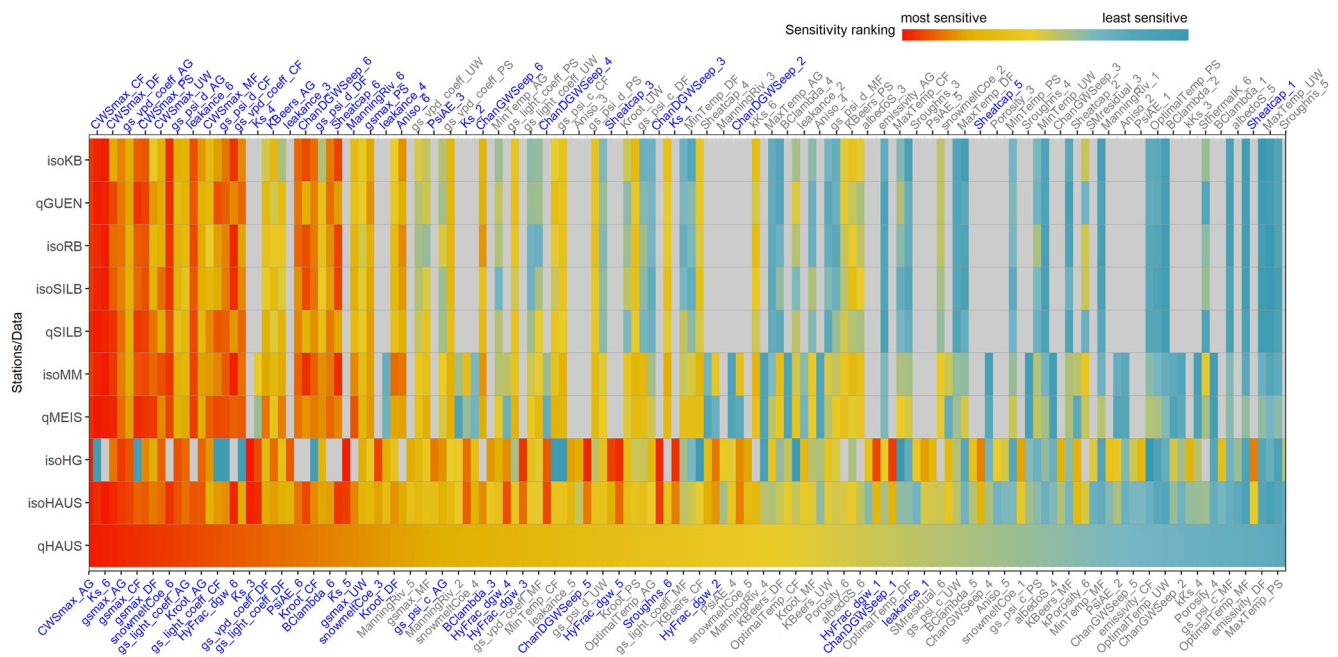


Figure 3. Parameter sensitivity rankings based on discharge (“q”) and isotope (“iso”) data at different stations and sampling sites. The top 20% sensitive parameters for each criterion were combined for model calibrations (parameter names labeled in blue font color). Labels at the end of vegetation- and soil-type dependent parameters corresponded to abbreviations in Figures 1c and 1d, respectively. The full sensitivity ranking results were provided in the associated repository (Yang et al., 2022).

stations. For instance, at MEIS/MM, the channel routing parameter (*ManningRiv*) and vegetation parameters of deciduous forest (labeled with “_DF”) become more sensitive primarily due to the increased importance of channel routing and deciduous forest area, respectively. In the lowlands, parameters related to hydraulic conductivity (*Ks*), deeper groundwater dynamics (*ChanDGWSeep* and *HyFrac_dgw*), soil heat fluxes (*Sheatcap*) and crop canopy-interception and transpiration were similarly sensitive to isotope simulations (isoHG, Figure 3; parameters labeled with 1–5 and “AG” represented different soil types and extensive arable lands in the lowlands, respectively).

The sensitivity rankings for discharge and isotopes diverged more significantly at the outlet HAUS (qHAUS and isoHAUS, Figure 3), compared to that of the upland gauging stations. The ranking for discharge generally followed that of the uplands, though with agricultural vegetation parameters ranked higher. In contrast, the ranking for isotopes showed more an equal combination of the most sensitive parameters from the uplands and the lowlands (isoMM and isoHG, respectively; Figure 3).

4.2. Discharge and Tracer Performance of the Upscaled Ech_2O -Iso Modeling

The top 20% sensitive parameters for each sensitivity ranking shown in Figure 3 were then combined and included in model calibrations. For both “discharge-only” and “discharge plus isotope” calibration schemes, the Ech_2O -iso discharge simulations captured the complex temporal variations at all four gauging stations across all the spatial scales reasonably well (Figure 4). At the two downstream stations HAUS and MEIS (Figures 4b and 4c), the two calibration schemes performed similarly well (KGEs mostly >0.55 and PBIAS mostly within $\pm 15\%$), with a slight degradation for the “discharge plus isotope” scheme. At the upland stations SILB and GUEN (Figures 4d and 4e), the simulations performed slightly better in terms of overall water balance (PBIAS mostly within $\pm 10\%$), but with lower KGE values; especially for the “discharge-only” scheme, the simulations largely overestimated peak flows (e.g., at SILB, median simulations overestimated by 26.8% during periods with discharge >90th quantile, Figure 4d), followed by very quick recessions (KGEs decreased to around 0.40, with the lowest mean value of 0.19 at SILB in the validation period). Importantly, the inclusion of isotope information significantly improved the performance during the high-flow periods of each hydrological year, thereby significantly mitigating the degradation of discharge performance in the uplands (e.g., KGEs mostly >0.45, Figures 4d and 4e). Although both calibration schemes over-estimated the extreme low-flows during the drought 2018–2019, the prolonged

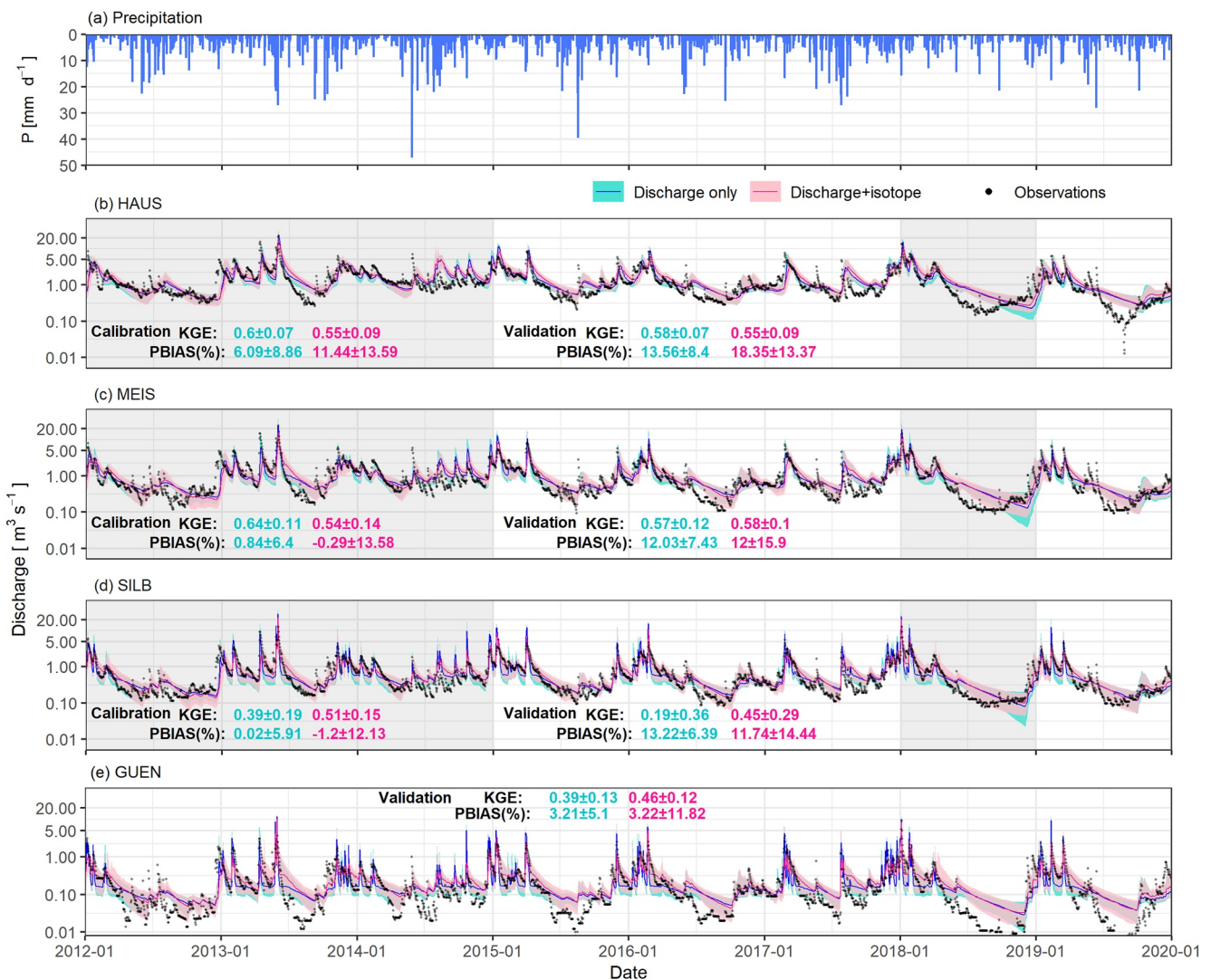


Figure 4. Catchment mean daily precipitation (a) and discharge performance at the four nested gauging stations ((b)—the outlet HAUS, (c)—MEIS, (d)—SILB, and (e)—GUEN, with drainage areas of 456, 184, 99, and 28 km², respectively) in the Selke catchment. The lines and uncertainty ranges represent the median values and 90% of simulation ranges, respectively, of the best 100 runs from the two calibration schemes. The performance metrics were shown as mean ± standard deviation of the best 100 runs. The calibration periods are shaded (the whole GUEN data were used for a spatial validation). Note that the y-axes are log-transformed for a better presentation.

discharge decreases and the subsequent recovering patterns were captured, and the simulation uncertainty during the driest periods was better constrained with the isotope information.

The inclusion of isotope data in the calibration improved the isotope simulations, both in terms of goodness-of-fit metrics (i.e., smaller MAEs) and better constrained uncertainty (Figures 5a–5f), although the overall errors remained quite high at some sites. The median values of the “discharge-only” calibration were consistently biased to the upper boundary of the uncertainty ranges; moreover, the simulated median isotopes and ages failed to reflect the prolonged drought disturbances in last 2 years (solid blue and dashed gray lines, Figures 5a–5f). In contrast, the isotope simulations of the “discharge plus isotope” scheme reproduced the highly damped stream water signals relatively well (MAEs of $\delta^2\text{H}$ mostly <7‰), especially during the validation periods with lower MAEs and the simulated median values closely following the observations (Figures 5a–5f). In particular, the performance was consistent across the spatial scales, ranging from smaller scales at upland sites KB and RB to larger scales at the downstream main stem sites MEIS/MM and HAUS (with drainage areas <15 and >200 km², respectively). The model simulations also captured the significant fluctuations during the wet high-flow periods, though we were not able to validate this due to the monthly observations. The simulated uncertainties in dryer

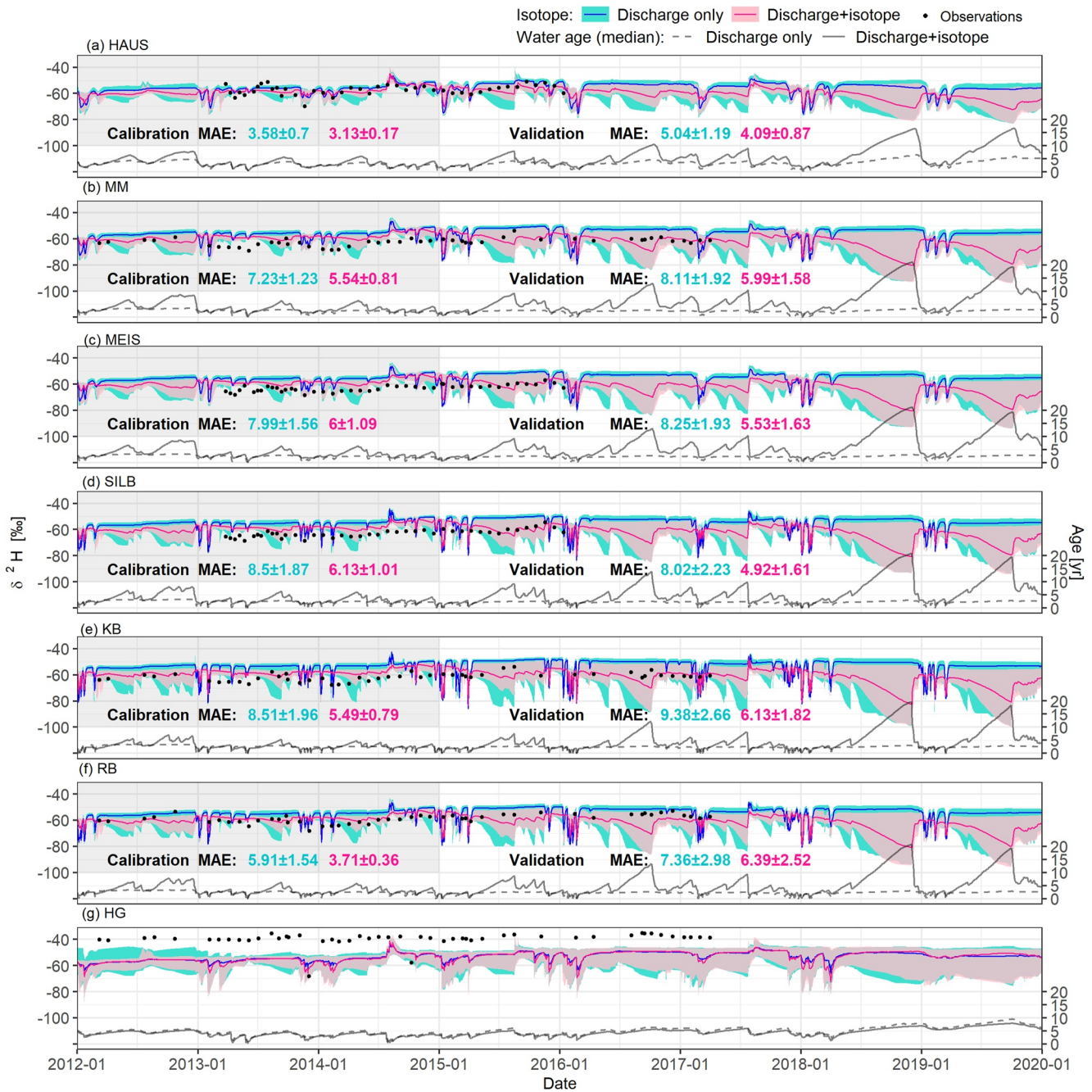


Figure 5. Tracer simulations (isotopes and water ages) at the sampling sites (a-HAUS, b-MM, c-MEIS, d-SILB, e-KB, f-RB, g-HG). The lines and uncertainty ranges represent the median values and 90% of the simulation ranges, respectively, of the best 100 runs from the two calibration schemes. The isotope performance metrics (Mean Absolute Error, MAE) are shown as mean \pm standard deviation of the best 100 runs. The calibration periods are shaded, and the data at site HG (g) were not included in calibrations.

periods and in the drought 2018–2019 were also higher and could be probably better constrained if more and higher resolution data were available (e.g., Yang et al., 2021).

In the calibration period 2012–2014, the isotope simulations generally overestimated the observations in 2013–2014 (except for sites HAUS and RB, Figures 5a–5f). Surprisingly in the validation period 2015–2017, these systematic overestimates were mostly eliminated, whereas the simulations at sites HAUS and RB showed larger underestimates. The exclusive lowland tributary HG exhibited heavily enriched and strongly fractionated isotope signals (Figure 2b), and model performance degraded largely showing a continuous underestimation, although

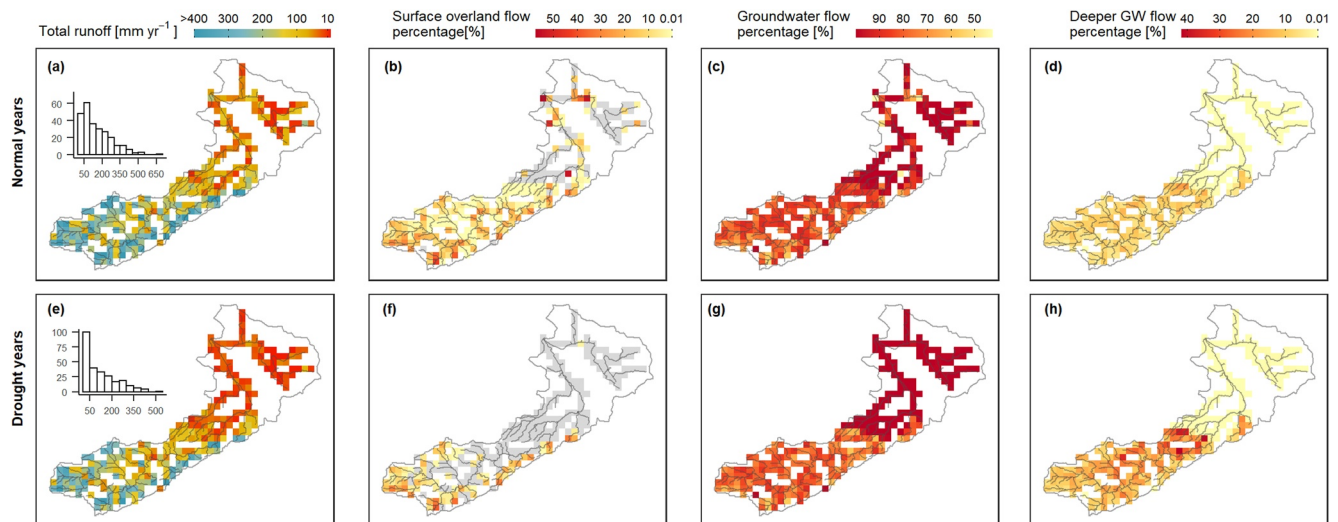


Figure 6. Catchment-scale annual runoff generation (a and e) and runoff component partitioning (b and f: surface overland flow; c and g: groundwater flow; d and h: deeper GW flow) in the normal years (2012–2017, upper panels) and the drought years (2018–2019, lower panels). The inserted histogram plots in panels (a and e) represent the distribution of annual total runoff across the catchment. Channel cells with overland flow percentage <0.01% were gray colored in panels (b and f). All values were calculated from the median simulations of the best 100 runs based on the discharge plus isotope calibration scheme.

the simulated isotopes were significantly more enriched compared to the upland simulations (e.g., the median $\delta^2\text{H}$ simulations at HG and MM were $-53.7\text{‰} \pm 4.0\text{‰}$ and $-59.4\text{‰} \pm 3.0\text{‰}$, respectively; ANOVA test, $p < 0.001$; Figures 5b and 5g). Interestingly, exceptions were that two observation points in December 2013 and October 2014 exhibited abnormal depletions, which fitted with the simulated levels.

The temporal variability of isotopic tracer behavior has also been correspondingly reflected in the stream water age simulations—a useful metric to characterize hydrological functioning (the median values of the best 100 runs shown as gray solid lines in Figure 5). Stream water ages were much younger during high-flow periods (as low as <0.5 years) and continuously increased (up to ca. 10 years in the normal years 2012–2017) with decreasing discharge, in accordance to the strongly fluctuating and gradually depleting isotope simulations, respectively. Such seasonal patterns of water ages were generally consistent across the spatial scales, with the dry period ages slightly increased from the upland tributary sites (KB and RB, Figures 5e and 5f) to the main stem sites (SILB and MEIS/MM, Figures 5b–5d). Notably, the ages during April–September were significantly younger in 2013–2014 than the other years (e.g., mean values <3.4 years vs. >5.1 year at the main stem sites, Figures 5a–5d), indicating relatively wetter conditions. This is supported by the higher surface overland flow contributions during these periods (Figure S5 in Supporting Information S1). In the drought years 2018–2019, water ages in the wet season (January–March) were similar to that of the normal years, while followed by dramatic increases (up to ca. 20 years) crossing the upland tributary sites to the main stem sites (Figures 5a–5f). At the exclusive lowland tributary HG, the ages were 3.9 ± 1.0 years, showing a more stationary state compared to other sites (e.g., the age standard deviation at MM was 2.3 years); interestingly, the ages showed a much more moderate increase in the drought 2018–2019 (5.7 ± 1.5 years), but the seasonal recovery behaviors were not observed during the intervening wet season of 2019. The ages at the outlet HAUS (Figure 5a) varied seasonally, following the general pattern of the Selke upland sites, while with reduced dynamic amplitudes (e.g., the ages increased to ca. 15 years during the droughts) due to the flow-weighted mixing of the main Selke and the lowland tributary.

4.3. Spatio-Temporal Variability of Water Flux and Age Dynamics and Their Responses to Drought

The best-performing model simulations provided further insights into the highly heterogeneous catchment functioning and its responses to the severe droughts 2018–2019, as well as the effects of spatial scale on process representation. Due to the high gradients of climatic and physiographic conditions, the runoff generation and component partitioning strongly contrasted between the uplands and the lowlands (Figure 6). For the normal years, the high-runoff producing upland areas (i.e., >75th quantile annual runoff in upland grid cells, with

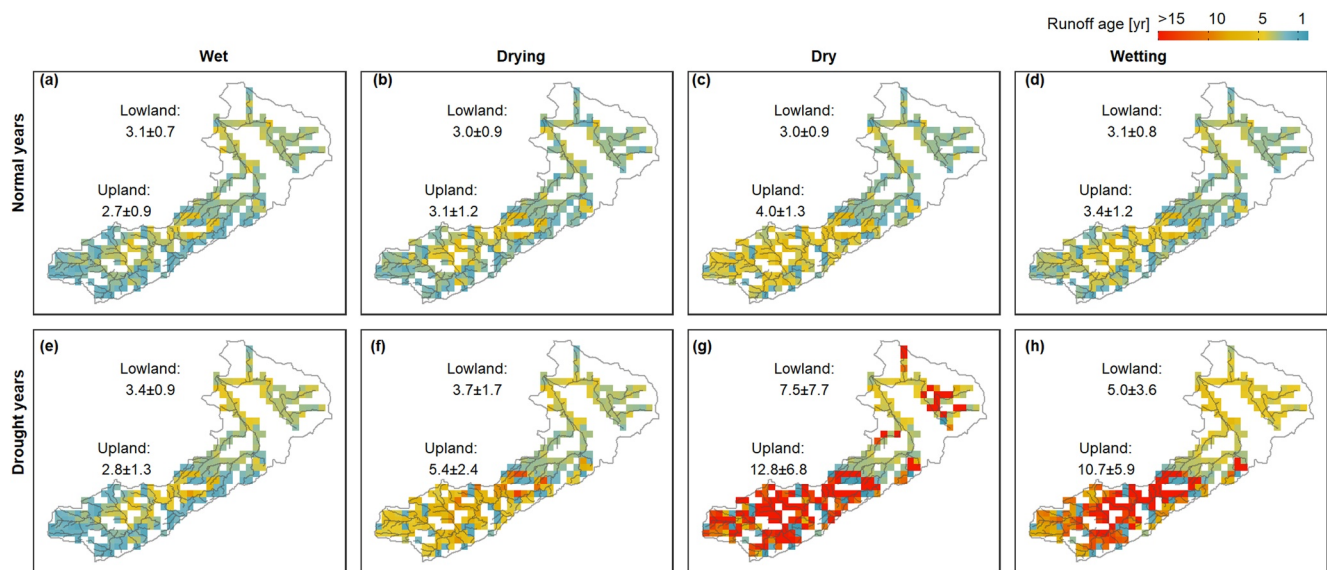


Figure 7. Spatial distributions of seasonal mean runoff ages weight by daily runoff amount in the normal (a–d) and the drought years (e–h). All values were calculated from the median simulations of the best 100 runs based on the discharge plus isotope calibration scheme.

corresponding annual precipitation as high as 828.0 ± 60.6 mm, Figure S1a in Supporting Information S1) generated 390.9 ± 79.0 mm annual runoff, with overland flow contributions accounting for 19.6% of total runoff in these cells; while the lowland areas exhibited an average annual runoff of 70.7 mm, with a large number of cells having almost no overland flow (Figures 6a and 6b). The groundwater flow from gravitational drainage of soil water storage dominated the total runoff, accounting for 81.3% and 93.9% of total runoff in the uplands and the lowlands, respectively, with the deeper GW flow contributing an additional 8.0% and 1.2%, respectively (Figures 6c and 6d). Such heterogeneous runoff generation patterns also responded differently to drought stress (Figures 6e–6h). The upland areas exhibited 18.4% decreases of annual runoff compared to the normal years (191.4 ± 102.0 vs. 234.7 ± 116.4 mm; ANOVA test, $p = 0.00125$), while its runoff partitioning changed in terms of significantly increased deeper GW flow contributions (an increase of up to $15.0 \pm 7.0\%$, ANOVA test, $p < 0.001$; Figure 6h vs. Figure 6d) and large areas with nearly no surface overland flow (except for the high-runoff cells; Figure 6f). In contrast, for the lowlands, the annual total runoff homogeneously decreased as low as 36.8 ± 32.2 mm, with nearly no surface overland flow ($<0.01\%$, Figure 6f) and only a marginal contribution from the deeper groundwater (mean of 2.5%, Figure 6h).

The simulated runoff ages provided further insights into the spatio-temporal heterogeneity of runoff mechanisms and hydrological functioning (Figure 7). Runoff ages of each grid cell exhibited strong seasonal variations in the uplands and a steadier pattern in the lowlands. The spatial variability of upland runoff ages also varied seasonally: the ages were generally 1-year younger in the high-runoff headwaters than in the central cells during the wet and drying seasons (e.g., means of 2.5 vs. 3.9 years during the drying period in Figure 7b); then, the uplands become more uniformly older during the dry season (spatial mean of 4.0 years with the lowest coefficient of variation as 0.325, Figure 7c), followed by spatially varying age recoveries during the re-wetting season (similar age ranges and spatial patterns to the drying season, Figures 7b and 7d). The lowland areas exhibited more homogeneous spatial patterns across seasons (mean ages of ~ 3.0 years with standard deviations <1.0 years, Figures 7a–7d). The runoff partitioning changes under drought stress induced marked impacts on the spatio-temporal variability of runoff ages for the uplands (Figures 7e–7h): the runoff ages and spatial patterns were similar to those of the normal years during the wet season, but ages dramatically increased and homogenized during the drying season (similar spatial pattern to the normal year dry season, with even older ages; Figure 7f vs. Figure 7c); the ages increased from 5 years to >10 years as the drought further developed during the dry and wetting seasons (Figures 7g and 7h). The lowland runoff ages showed relatively moderate and spatially homogeneous increases in all seasons of the drought years (e.g., mean ages of 3.7 and 5.0 years during the drying and wetting seasons, respectively, Figures 7f and 7h), which is in accordance with the reduced soil water replenishment due to reduced precipitation (annual means of 547.1 ± 58.0 vs. 400.5 ± 51.2 mm in the normal and drought years, respectively, Figure S1 in Supporting Information S1) and marginal changes in runoff source waters (Figure 6).

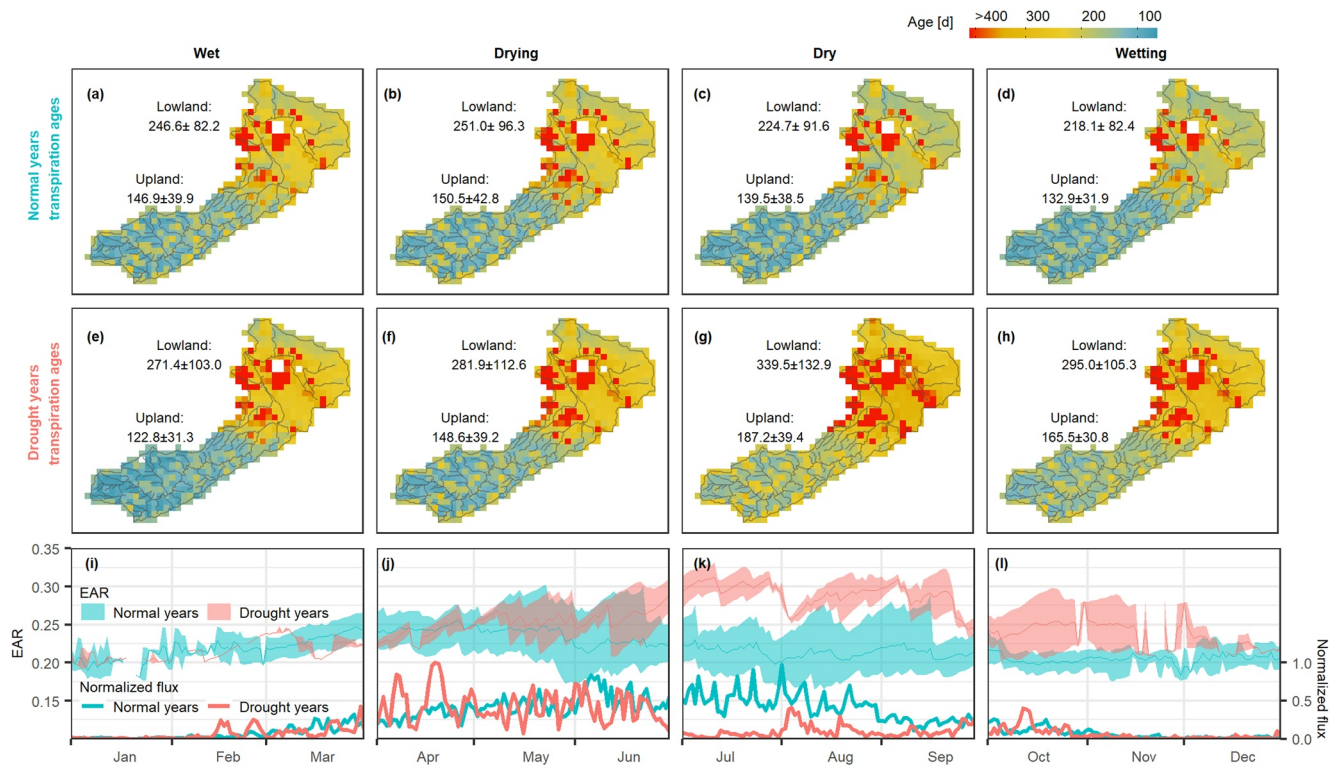


Figure 8. Spatial distributions of transpired water ages in the normal (a–d) and the drought years (e–h), and the catchment-averaged seasonal patterns of the transpiration exit age ratio—EAR and the min-max normalized transpiration fluxes (i–l). The presented water ages in panels (a–h) are seasonal mean ages weighted by daily transpiration fluxes, and the EARs are averaged across the whole catchment. All values are based on the median simulations of the calibrated best 100 runs. Note that the lowland mining lake areas are excluded.

The simulated ecohydrological fluxes also showed high spatial variations between the uplands and the lowlands (annual evapotranspiration means of 436.5 vs. 369.4 mm, respectively in the normal years, and 361.0 vs. 311.2 mm, respectively in the drought years; detailed seasonal patterns of the flux partitioning into transpiration, soil evaporation and interception evaporation shown in Figure S6 in Supporting Information S1). Transpiration, as the major component of catchment ecohydrological fluxes involving all three-layer soil water dynamics, showed spatio-temporally varying usage of soil water both quantitatively and qualitatively (Figures 8 and 9), although this was not specifically constrained by xylem isotopes. In the normal years, the transpiration fluxes were similarly high during April–September as the major vegetation period (mean of $0.6 \pm 0.24 \text{ mm d}^{-1}$, Figures 8j–8k, Figure S6 in Supporting Information S1); however, the transpiration water ages were somewhat older during the drying season than the dry season across the catchment (means of 150 vs. 140 days and 251 vs. 225 days in the uplands and lowlands, respectively; Figures 8b and 8c). Meanwhile the transpiration EARs were also higher during the drying season, indicating further higher degrees of vertical mixing of soil waters (Figures 8j, 8k, 9b and 9c; ANOVA test, $p < 0.001$). The uplands exhibited significantly younger transpiration water ages, but higher EARs, in the channel-connected grid cells compared to the hillslope cells not connected to the channel across all seasons (e.g., during the drying season, mean ages and EARs were 123.1 vs. 196.7 days and 0.242 vs. 0.236, respectively, Figures 8 and 9b; ANOVA test, $p < 0.001$). The lowland transpiration water ages were mostly >200 days, with some scattered areas exhibiting exceptionally old ages (i.e., >400 days).

The modeling implied that the severe drought in 2018–2019 had dramatically altered the seasonal and spatial patterns of transpiration. The transpiration fluxes were largely maintained during the drying season and dramatically decreased as the drought further developed; importantly, the increases of EARs started earlier since the drying season and further increased during the dry season (Figures 8j and 8k). Moreover, these drought-induced changes of water ages showed significant spatial variability (Figures 8e–8h and 9e–9h). The upland transpiration water ages were significantly younger than that in the normal years during the wet season (means of 123 vs. 147 days, Figures 8a and 9a; ANOVA test, $p < 0.001$) and continuously increased in the following seasons (i.e.,

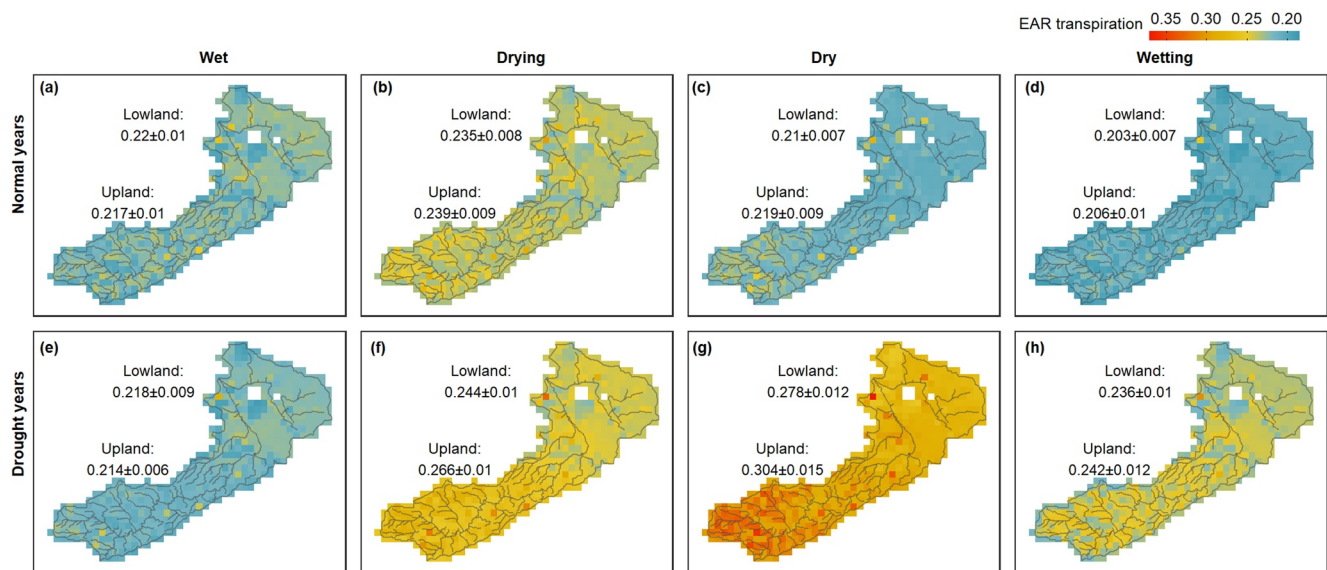


Figure 9. Spatial and seasonal variations of the transpiration exit age ratio—EAR in the normal (a–d) and the drought years (e–h). All values are based on the median simulations of the calibrated best 100 runs. Note that the lowland mining lake areas are excluded.

means of 187 and 165 days during the drying and wetting seasons); while the lowland transpiration ages were significantly older than that of the normal years in all seasons (e.g., means of 282 vs. 251 days during the drying seasons), with an increasing expansion of the high-age (>400 days) areas (Figures 8a–8h). Interestingly, the upland channel-connected cells still exhibited younger transpiration water ages than non-channel connected cells (Figures 8e–8h). Such landscape organizational differences did not always apply for transpiration age extractions (indicated by EARs) under drought stress, that is, EARs in the uplands were spatially more homogeneous during the wet and drying seasons (Figures 9e and 9f) but spatially varying during the dry and wetting seasons (i.e., with higher EARs in channel-connected cells than non-channel cells, Figures 9g and 9h).

5. Discussion

5.1. Upscaling Tracer-Aided Ecohydrological Modeling to Larger Catchments: Opportunities and Challenges for Process Representation

Through successful use at the plot-scale and in smaller catchments, the fully distributed EcH_2O -iso model has shown flexible structures for representing intra- and inter-grid heterogeneity and catchment spatial organization (e.g., the hillslope-channel configuration) (Kuppel et al., 2020; Smith, Tetzlaff, Maneta, et al., 2022; Yang et al., 2021). This highlights the model's potential suitability for upscaling and the necessity for testing in larger catchments like the Selke.

First, the model allows spatially distributed climate inputs and catchment physiographic properties (e.g., soil types and vegetation species/types) to be represented independent of modeling resolution. This ensures adequate representation of multiple environmental drivers and their various interactions that encompass the functional heterogeneity across larger catchments. For example, given the similar landscape characteristics and geology, the spatial variability of runoff processes in the Selke uplands (Figure 6) was primarily driven by the spatial pattern of precipitation (Figure S1 in Supporting Information S1). Meanwhile, the lowland variability of water ages in transpiration fluxes and the soil water sources (Figure 8, Figure S7 in Supporting Information S1) was likely explained by the heterogeneity of soils and related hydraulic properties. In addition, the process heterogeneity within a single model cell became more pronounced with coarser modeling resolution at larger scales (1 km × 1 km in this study). The EcH_2O -iso model addresses this issue through separate calculation of vegetation processes for each vegetation type present in a grid cell and then areal-weighting them for the overall grid outputs. Additionally, hydrological parameters related to soil processes are soil-type dependent and aggregated using the area-weighted geometric means if multiple soil types are present in one grid cell. Such designs of model structure and parameterization at the finer input resolution have advantages in terms of increasing representation of spatial

heterogeneity and modeling robustness across scales (Beven, 1995; Clark et al., 2017). Second, beyond heterogeneity and process representation, the EcH₂O-iso model structure accommodates catchment organizational heterogeneity, and thus, allows emergent process representation rather than forcing only by distributed parameterization. Accommodating such catchment organizational principles is essential in representing overall catchment functioning (McDonnell et al., 2007; Sivapalan, 2005).

However, the upscaling application to the Selke catchment required some model adaptations (Section 2.2). Specifically, initial simulations highlighted the importance of re-infiltration of surface overland flow from up-slope cells. This process is not considered in the original model when grid cells include river channels; while for coarser modeling resolutions with dense river networks, switching on this process was needed to reproduce stream isotope simulations and damping variability (detailed trial-and-error results not shown due to the limited space). Such upscaling adaptations need to be carefully evaluated since the importance of specific processes might differ at different scales and new processes might emerge (Fatichi et al., 2016). The redistribution of surface runoff between multiple pathways, for example, re-infiltrating into soil, transport to a down-slope grid cell, or drainage to the stream channel as an overland flow component, is highly non-linear at small scales and extremely challenging to constrain for larger-scale investigation with coarser resolution modeling. For example, the clear systematic mismatch of our isotope simulations at some sites during March-May 2013 and June-August 2014 (Figure 5) was likely due to the abnormally wet weather that invoked significant overland flow contributions (Figure 4a, Figure S5 in Supporting Information S1). This also links to the general limits of coarse-resolution modeling in reproducing localized processes with shorter characteristic length scales (Skøien et al., 2003), and particularly so for the tracer mixing between compartments with contrasting isotopic signatures (Smith et al., 2021). In addition, although the simulated $\delta^2\text{H}$ in the drier, warmer lowlands were substantially more enriched in both stream waters and soil waters than in the uplands (Figure 5, Figure S8 in Supporting Information S1), the simulations could not reach the enrichment level of the HG observations. However, beyond modeling uncertainties in the lowland loess areas, this mismatch is most likely explained by an effect of the interactions between stream water and the nearby mining lake water (impacted by open water evaporation) through complex groundwater dynamics in this region (Z.-Y. Zhang et al., 2021). Note that the well-matched two depleted observations at HG were sampled during the first flow-events after dry summers, likely before stream-lake hydraulic connectivity was re-established. In addition, the surface open water evaporation and the associated isotopic fractionation effects on isotope signatures would need more explicit representation of lake impacts for future modeling of larger-scale catchments where lakes and reservoirs are likely to be present (e.g., our simulations missed the isotope observations at RB, which were also impacted by an upstream small reservoir, Figure 5f).

Although the logarithmic NSE discharge metric was used particularly to better constrain low-flow simulations (Krause et al., 2005), the prolonged low flows under the droughts were still slightly overestimated (means of 0.51 vs. the observed 0.21 m³l⁻¹ at the outlet station, note the logarithmic y-axes in Figure 4). It is generally very difficult to further overcome such subtle mismatches of summer low flows and would need further effort to resolve alternative model structures (Staudinger et al., 2011). We also noted that our ecohydrological model representation of deeper groundwater dynamics and their interactions with surface waters remains overly simplistic. Moreover, at some sites the overall isotope simulations exhibited considerable mismatches against the monthly observations. This primarily reflects the challenges of upscaling tracer-aided modeling as detailed above and the general limitation of coarse resolution isotope data available in large catchments for calibration. Similar to the monthly samples in the Selke/Bode catchment, large-scale catchments are unfortunately often lacking of high-resolution sampling, which is necessary for characterizing short-term dynamics of tracer and solute transport. Therefore, our tracer-aided modeling sought to reproduce the general seasonality and spatial patterns of isotopes, which still increases confidence in the general flux-storage-age representation compared to conventional hydrological models (Birkel et al., 2014). This is also the reason that we chose MAE as the metric of isotope calibration, rather than the typical KGE or NSE for discharge simulations with high-resolution timeseries. We noted that the isotope performance might be further improved through different calibration strategies (e.g., using different evaluation metrics and calibration methods), though the general spatio-temporal patterns of catchment functioning derived in this study are still insightful (see Section 5.3).

There was also increased uncertainty in the isotope simulations and the inferred water ages during the droughts (e.g., periods with water ages >15 years, Figure 5), which was most likely due to the increased contributions of deeper groundwater flows with generally much older ages. We stress, however, that we were not intending to derive exact age values in these periods (which would require other tracers), but rather to use the water

stable isotopes to indicate changes in catchment functioning. In addition, and in line with the challenge of all process-based modeling in larger catchments, we acknowledge that the model estimates of internal processes might differ from the real catchment functioning, and that additional data from different sources could be critical to further verify and/or constrain the modeling. For example, the model simulated reasonable total evapotranspiration (e.g., 396.5 ± 43.7 and 332.6 ± 36.2 mm in the normal and drought years, respectively), similar to other regional estimates (e.g., flux tower measurements in Gillefalk et al. (2021) and modeled estimates in Smith et al. (2021)). However, the model seemed to underestimate the proportion of transpiration, especially in the lowlands (Figure S6 in Supporting Information S1), given that this process is only indirectly constrained by stream isotopes as part of the modeling work flow while not specifically by vegetation isotopes, for example, in xylem waters. To address this would need direct calibration objectives based on, for example, measurements of sap flux (H. Wang et al., 2019), soil water isotopes and/or energy information from an eddy covariance flux tower.

5.2. The Value of Tracer Information on Aiding Ecohydrological Modeling and Implications of Isotope Monitoring at Larger Scales

We demonstrated that the information content of even relatively coarsely sampled isotope tracer data can still be a useful complement to conventional hydrometric data and enhance larger-scale tracer-aided modeling through comprehensive cross-scale parameter sensitivity analyses and model calibrations.

The most sensitive Ech_2O -iso parameters for discharge and isotope data were generally similar, including parameters related to canopy interception (*CWSmax*), snowmelt (*snowmeltCoe*), soil hydraulic properties (e.g., *Ks* and *BCLambda*), vegetation stomatal and transpiration properties (e.g., *gsmax*, *gs_light_coeff*, *gs_vpd_coeff*, and *Kroot*), and runoff generation (e.g., *ChanGWSeep*, *leakance*, and *ChanDGWSeep*) (Figure 3). This indicates the consistency of the information content of hydrometric and tracer data (Lischeid, 2008). Meanwhile, parameters related to the overall water balance ranked higher for discharge simulations, while parameters related to runoff partitioning, soil water dynamics and the energy balance ranked higher for isotopes. This supports the utility of isotope data in constraining internal processes in terms of more realistic representation of storage-flux interactions, water partitioning and the associated evolution of water ages. These findings are in line with Birkel et al. (2014) who included soil water isotope measurements in sensitivity analysis in the well-studied Bruntland Burn catchment, Scotland (3.2 km², with annual precipitation of around 1,000 mm and ET of ~350 mm). The inclusion of isotope data in calibration maintained similarly good discharge simulations, with particularly significant improvements at the upper stations in validation (e.g., mean KGEs improved from 0.19 to 0.45 at SILB; Figures 4d and 4e), though slight degradation in fit at the two downstream stations where the “discharge-only” performances were already good (Figures 4b and 4c). Crucially, the $\delta^2\text{H}$ tracer simulations and stream water age estimates improved significantly in terms of seasonal patterns and responses to nonstationary high-flow perturbations (Figure 5), although the overall performance of isotope simulations was still inhibited by the general challenges of the upscaled tracer-aided modeling as detailed in Section 5.1.

Embracing the upscaling challenges discussed above, the isotope-informed stream water age simulations (Figure 5, Figure S9 in Supporting Information S1) agreed quite well with the independent age estimations from Lutz et al. (2018) who used lumped transit time models (e.g., our simulated mean age was 2.9 years at the main stem site MM during February 2012–May 2015, compared with that of 2.2 years in Lutz et al. (2018)). Moreover, the simulated stream water ages were significantly younger during the abnormally wet periods of March–May 2013 and June–August 2014 (Figure 5). Therefore, our upscaled tracer-aided modeling plausibly provided a first larger-scale approximation that captures storage-flux-age interactions and flow path distributions, as well as their nonstationary dynamics, in larger catchments.

Most coupled flow-tracer modeling studies in small-scale experimental catchments are based on high-resolution sampling (e.g., daily) and include distributed measurements in soil and groundwater compartments (Smith et al., 2021; Soulsby, Birkel, Geris, Dick, et al., 2015). However, such intensive isotope data are less feasible at larger scales, and often available stream water isotopes are coarsely sampled (e.g., monthly), exhibiting highly damped temporal patterns compared to rainfall (like the Selke). Notwithstanding that the isotopic amplitude might be masked by the coarse sampling frequency, the strong damping effects are most likely due to the dominance of subsurface flow path interactions at larger scales, which involve strong internal mixing with stored water (e.g., McDonnell & Beven, 2014; McDonnell et al., 2010). This is also consistent with independent perceptual understanding of the well-studied Selke catchment derived from hydrological and nitrate investigations based

on process-based modeling (Yang et al., 2018) and data-driven analysis (X. Zhang et al., 2020). Therefore, a strategic implication for large-scale monitoring is that the availability of the tracer information can be invaluable, even though the data might be inevitably coarse. Of course, higher frequency data collection, particularly during storm events can provide additional insights into catchment function (e.g., Stevenson et al., 2021). Indeed, we observed that our upscaled modeling was weaker in reproducing the depleted isotope signals induced by enhanced surface overland flow dynamics under wetter weather conditions. Optimizing isotope tracer monitoring design (e.g., higher frequency sampling that targets characteristic flow regimes (Stevenson et al., 2021; L. Wang et al., 2019)) could perhaps better aid representation of highly non-linear processes and modeling constraints, meanwhile reducing the logistical and financial burdens that are often already high for larger-scale monitoring.

5.3. Insights Into the Heterogeneous Ecohydrological Functioning Under Varying Hydroclimatic Conditions

The fully distributed EcH_2O -iso modeling uniquely provided spatio-temporally explicit insights that allowed us to disentangle blue- and green-water flux dynamics and the associated output age distributions across scales. Such detailed information is particularly critical in larger-scale catchments that often show contradictory upland-lowland behaviors (e.g., Tetzlaff et al., 2011), like the Selke catchment.

In the Selke uplands, the high-elevation headwaters exhibited much higher precipitation and annual runoff than the leeward-positioned central areas. The thin soils and shallow bedrock have resulted in a relatively flashy flow regime, with rapid replenishment of soil water storage deficits (Dupas et al., 2017; Yang, Rode, et al., 2022). Coupled with the spatially varying precipitation, the simulated total runoff and contributions of surface overland flow exhibited significant spatial variability (Figure 6b). This has further resulted in younger runoff ages in the more runoff producing headwaters, except for the dry season when the uplands were dry and deeper groundwater uniformly sustained the low baseflows (Figures 7a–7d). It is notable that stream water along the main stem showed more homogeneity and maintained young ages moving downstream, due to the runoff-weighted age mixing and the dominance of younger runoff from the wetter headwaters (Figure S9 in Supporting Information S1). This implies that stream water tracer signals are significantly impacted by the hillslope-channel organization (e.g., outlined as a first-order control in mountainous regions by Jencso et al. (2009)) and hydroclimatic variability in larger catchments, and that spatially explicit investigations are important to show the actual transit time distributions across the catchment (e.g., Hrachowitz et al., 2010; Soulsby et al., 2010). In contrast, the dryer, low-gradient lowlands homogeneously exhibited more stationary hydrological regimes, due to the flat topography and limited lateral hydrological connectivity (Figure S10 in Supporting Information S1).

The prolonged droughts in 2018–2019 provided an opportunity to investigate spatio-temporal variations of catchment water balance dynamics under nonstationary conditions. The upland annual runoff was largely maintained in the drought years, but with significant changes in runoff partitioning (i.e., relatively deeper GW runoff contribution almost doubled and there were extensive areas generating no overland flow; Figures 6e–6h). Intensive groundwater monitoring data in the Selke's headwater Schäfertal catchment (Yang et al., 2021) provided direct evidence that the regional deeper groundwater system could sustain significant stream flows under the droughts, primarily because of the upland shallow schist/claystone bedrock with significant fractures (Jiang et al., 2014; Wollschläger et al., 2016). This further resulted in the general increases in runoff ages (Figures 7e–7h), since longer flow paths through the fractured bedrock typically exhibits longer water transit times, although with similar catchment forms and hydrological regimes (Hale & McDonnell, 2016). In contrast for the lowlands, the drought induced significant decreases in annual runoff but only marginal changes in runoff partitioning and more moderate increases in runoff ages. On the one hand, the groundwater fluxes from soil water storage (Figure S10 in Supporting Information S1) plausibly sustained the lower stream flows due to the flat lowland topography with low hydraulic gradients (Tetzlaff et al., 2011). On the other hand, the deeper groundwater storage in the lowland loess aquifer decreased dramatically in recent drought years (groundwater level measurements from the water authority LHW, <https://lhw.sachsen-anhalt.de/>, last accessed on 13 September 2022), and thus, did not evoke significant nonstationary ages responses in lowland stream waters.

Transpiration fluxes exhibited significantly younger ages in the upland channel-connected grid cells with higher EARs (e.g., mean ages of 123 vs. 197 days during the drying season; Figures 8a–8d and 9a–9d), indicating more uniformly younger soil water sources in these cells. This likely resulted from interactions between blue- and green-water dynamics: the lateral groundwater seepage to channels has plausibly

enhanced overland flow re-infiltration and the consequent vertical exchanges between soil layers, resulting in whole-profile younger soil water ages (Figure S7 in Supporting Information S1). In the flat lowlands, due to limited lateral flux exchanges, the spatial patterns of transpiration water ages were consistently similar to those of soil water ages in all layers; the areas with much older ages were characterized as higher outgoing potentials, such as the highly sandy areas and the higher-slope headwater grid-cells (Figure 8, Figures S7 and S10 in Supporting Information S1). In responding to the drought disturbances, transpiration fluxes exhibited a delayed response compared to that of EARs (Figures 8i–8l and 9). Such differences indicated that the transpiration fluxes at the beginning of the droughts were maintained largely by deeper layer soil water sources with less-variable age distributions (i.e., higher EARs). Furthermore, compared to the absolute transpiration ages that exhibit landscape organizational and upland-lowland variabilities, the transpiration EARs were more spatially homogeneous, and therefore, can be a sensitive indicator of environmental disturbance (like the droughts) impacts across larger catchments. Assessing catchment-scale vegetation water use is still fundamentally challenging (Asadollahi et al., 2022; Evaristo et al., 2019; Soulsby et al., 2016), and the results of our upscaled modeling largely depended on the preliminary conceptualization and specific catchment settings in this study. In particular, our transpiration simulations were not directly constrained by isotopes. Therefore, increasingly available information from, for example, xylem water samples, will be useful to further reduce model uncertainty and increase the process inferences (e.g., Smith, Tetzlaff, Landgraf, et al., 2022). Also, data-driven improved conceptualization and specific considerations of the complex transpiration processes and their impacts on water use would be advantageous, for example, taking advantage of recent advances from in situ experiments and controlled lysimeters (Benettin et al., 2021; Evaristo et al., 2019; Smith, Tetzlaff, Landgraf, et al., 2022).

6. Conclusions

We upscaled process-based, tracer-aided ecohydrological modeling using the EcH₂O-iso model in the large, heterogeneous Selke catchment (456 km²). Several adaptations to EcH₂O-iso were made to further improve the model's process and spatial representations. To our knowledge, such tracer-aided, process-based ecohydrological modeling has not been conducted at such large spatial scales. The novel larger-scale modeling used spatially distributed flow and isotope data across the Selke catchment with contrasting landscape properties. Multi-site and multi-data sensitivity analysis revealed that the information content of discharge and stream isotope data was generally “consistent” (similar sensitive parameters) and “orthogonal” (e.g., parameters related to runoff partitioning and dynamics in soils ranked high for isotopes). Further, the inclusion of isotope information in multi-criteria calibration generally preserved the discharge performance with improved transferability to validations, and more importantly, substantially improved the isotope performance even at such large scales (though these remained uncertain due to the general challenges of the upscaled tracer-aided modeling using relatively coarse tracer data for calibration). This still provided the basis for more plausible inferences of stream water ages in terms of general seasonal dynamics and spatial variability.

Following comprehensive process representation and isotope-aided model constraints, the novel fully distributed modeling has unique advantages in providing new insights into spatial heterogeneities in ecohydrological functioning in larger catchments, in terms of tracer mixing and water ages, and their response to environmental disturbances like droughts. The upland areas exhibited marked spatial variability in hydroclimatic conditions and further in simulated ages of terrestrial water sources; the hydrological regimes under the drought disturbances were largely maintained by the increased contributions from deeper groundwater in the fractured bedrock, while associated with dramatically increased stream water ages. In contrast, the flat and drier lowlands exhibited more constant flow regimes and water age distributions, with moderate responses to drought disturbances, primarily due to the flat topography and limited hydrological connectivity. As a major green-water flux, transpiration water ages exhibited strong spatial patterns related to catchment organization in the uplands but were generally more homogeneous in the lowlands. The more spatially generalized transpiration EAR, indicating transpiration water age extraction preference, can be a sensitive indicator of functional changes across larger catchments under environmental disturbance (like the droughts). All these innovative insights from the upscaled tracer-aided ecohydrological modeling are of particular relevance to inform larger-scale management of heterogeneous catchments under environmental change and climate extremes.

Data Availability Statement

The data used in this study are available at Yang et al. (2022b) via <https://doi.org/10.5281/zenodo.6628889>. The model code and modifications of Ech_2O -iso are also publicly available via GitHub repository (https://github.com/XYang-EcoHydroWQ/Ech2O-iso_largescale), and Yang et al. (2022a) via <https://doi.org/10.5281/zenodo.7117112>.

Acknowledgments

This work was supported by the Modular Observation Solutions for Earth Systems—MOSES project and the Terrestrial Environmental Observatories—TERENO project. Funding was also received for Doerthe Tetzlaff through the Einstein Research Unit “Climate and Water under Change” from the Einstein Foundation Berlin and Berlin University Alliance. Contributions by Chris Soulsby were supported by the Leverhulme Trust through the ISO-LAND project (RPG 2018 375). The authors thank Michael Rode and Ralf Merz for permitting the use of the water stable isotope data and for constructive discussions related to the data analysis. The authors would like to thank Marco Maneta and Aaron Smith for their support and discussion on the modeling. The authors thank the Editor Peter Troch, the Associate Editor and three anonymous reviewers for their constructive comments. The authors also thank the German Weather Service, the Federal Institute for Geosciences and Natural Resources and the State Agency for Flood Protection and Water Management Saxony-Anhalt for the model setup data. Open Access funding enabled and organized by Projekt DEAL.

References

- Ala-aho, P., Soulsby, C., Pokrovsky, O. S., Kirpotin, S. N., Karlsson, J., Serikova, S., et al. (2018). Using stable isotopes to assess surface water source dynamics and hydrological connectivity in a high-latitude wetland and permafrost influenced landscape. *Journal of Hydrology*, 556, 279–293. <https://doi.org/10.1016/j.jhydrol.2017.11.024>
- Ala-aho, P., Tetzlaff, D., McNamara, J. P., Laudon, H., & Soulsby, C. (2017). Using isotopes to constrain water flux and age estimates in snow-influenced catchments using the STARR (Spatially distributed Tracer-Aided Rainfall–Runoff) model. *Hydrology and Earth System Sciences*, 21(10), 5089–5110. <https://doi.org/10.5194/hess-21-5089-2017>
- Altermann, M., Rinklebe, J., Merbach, I., Körschens, M., Langer, U., & Hofmann, B. (2005). Chernozem—Soil of the year 2005. *Journal of Plant Nutrition and Soil Science*, 168(6), 725–740. <https://doi.org/10.1002/jpln.200521814>
- Asadollahi, M., Nehemy, M. F., McDonnell, J. J., Rinaldo, A., & Benettin, P. (2022). Toward a closure of catchment mass balance: Insight on the missing link from a vegetated lysimeter. *Water Resources Research*, 58(4), e2021WR030698. <https://doi.org/10.1029/2021wr030698>
- Benettin, P., Nehemy, M. F., Asadollahi, M., Pratt, D., Bensimon, M., McDonnell, J. J., & Rinaldo, A. (2021). Tracing and closing the water balance in a vegetated lysimeter. *Water Resources Research*, 57(4), e2020WR029049. <https://doi.org/10.1029/2020wr029049>
- Beven, K. (1995). Linking parameters across scales: Subgrid parameterizations and scale dependent hydrological models. *Hydrological Processes*, 9(5–6), 507–525. <https://doi.org/10.1002/hyp.3360090504>
- Beven, K. (2002). Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrological Processes*, 16(2), 189–206. <https://doi.org/10.1002/hyp.343>
- Beyer, M., Kühnhammer, K., & Dubbert, M. (2020). In situ measurements of soil and plant water isotopes: A review of approaches, practical considerations and a vision for the future. *Hydrology and Earth System Sciences*, 24(9), 4413–4440. <https://doi.org/10.5194/hess-24-4413-2020>
- BGR. (2018). *Soil map of Germany 1:200 000 (BÜK200)*. Bundesanstalt für Geowissenschaften und Rohstoffe (BGR).
- Birkel, C., & Soulsby, C. (2015). Advancing tracer-aided rainfall–runoff modelling: A review of progress, problems and unrealised potential. *Hydrological Processes*, 29(25), 5227–5240. <https://doi.org/10.1002/hyp.10594>
- Birkel, C., Soulsby, C., & Tetzlaff, D. (2014). Developing a consistent process-based conceptualization of catchment functioning using measurements of internal state variables. *Water Resources Research*, 50(4), 3481–3501. <https://doi.org/10.1002/2013wr014925>
- Brauman, K. A. (2015). Hydrologic ecosystem services: Linking ecohydrologic processes to human well-being in water research and watershed management. *WIREs Water*, 2(4), 345–358. <https://doi.org/10.1002/wat2.1081>
- Campolongo, F., Saltelli, A., & Cariboni, J. (2011). From screening to quantitative sensitivity analysis. A unified approach. *Computer Physics Communications*, 182(4), 978–988. <https://doi.org/10.1016/j.cpc.2010.12.039>
- Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., et al. (2017). The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism. *Hydrology and Earth System Sciences*, 21(7), 3427–3440. <https://doi.org/10.5194/hess-21-3427-2017>
- Craig, H., & Gordon, L. I. (1965). Deuterium and oxygen 18 variations in the ocean and the marine atmosphere. In *Stable isotopes in Oceanographic studies and paleotemperatures* (pp. 277–374). Laboratoro di geologia nucleare.
- Dupas, R., Musolff, A., Jawitz, J. W., Rao, P. S. C., Jäger, C. G., Fleckenstein, J. H., et al. (2017). Carbon and nutrient export regimes from headwater catchments to downstream reaches. *Biogeosciences*, 14(18), 4391–4407. <https://doi.org/10.5194/bg-14-4391-2017>
- Evaristo, J., Kim, M., van Haren, J., Pangle, L. A., Harman, C. J., Troch, P. A., & McDonnell, J. J. (2019). Characterizing the fluxes and age distribution of soil water, plant water, and deep percolation in a model tropical ecosystem. *Water Resources Research*, 55(4), 3307–3327. <https://doi.org/10.1029/2018wr023265>
- Faticchi, S., Pappas, C., & Ivanov, V. Y. (2016). Modeling plant–water interactions: An ecohydrological overview from the cell to the global scale. *WIREs Water*, 3(3), 327–368. <https://doi.org/10.1002/wat2.1125>
- Fenicia, F., McDonnell, J. J., & Savenije, H. H. G. (2008). Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resources Research*, 44(6), W06419. <https://doi.org/10.1029/2007wr006386>
- Gibson, J. J., Eby, P., Stadnyk, T. A., Holmes, T., Birks, S. J., & Pietroniro, A. (2021). Dataset of ^{18}O and ^2H in streamflow across Canada: A national resource for tracing water sources, water balance and predictive modelling. *Data in Brief*, 34, 106723. <https://doi.org/10.1016/j.dib.2021.106723>
- Gillefalk, M., Tetzlaff, D., Hinkelmann, R., Kuhlemann, L. M., Smith, A., Meier, F., et al. (2021). Quantifying the effects of urban green space on water partitioning and ages using an isotope-based ecohydrological model. *Hydrology and Earth System Sciences*, 25(6), 3635–3652. <https://doi.org/10.5194/hess-25-3635-2021>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Hale, V. C., & McDonnell, J. J. (2016). Effect of bedrock permeability on stream base flow mean transit time scaling relations: I. A multiscale catchment intercomparison. *Water Resources Research*, 52(2), 1358–1374. <https://doi.org/10.1002/2014wr016124>
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018). *ERA5 hourly data on pressure levels from 1979 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.bd0915c6>
- Holmes, T., Stadnyk, T. A., Kim, S. J., & Asadzadeh, M. (2020). Regional calibration with isotope tracers using a spatially distributed model: A comparison of methods. *Water Resources Research*, 56(9), e2020WR027447. <https://doi.org/10.1029/2020wr027447>
- Hrachowitz, M., Soulsby, C., Tetzlaff, D., & Speed, M. (2010). Catchment transit times and landscape controls—Does scale matter? *Hydrological Processes*, 24(1), 117–125. <https://doi.org/10.1002/hyp.7510>
- Hunt, A., Faybishenko, B., & Ghanbarian, B. (2021). Non-linear hydrologic organization. *Nonlinear Processes in Geophysics*, 28(4), 599–614. <https://doi.org/10.5194/npg-28-599-2021>
- IAEA. (2012). *Monitoring isotopes in rivers: Creation of the global network of isotopes in rivers (GNIR)*. International Atomic Energy Agency.

- Iorgulescu, I., Beven, K. J., & Musy, A. (2007). Flow, mixing, and displacement in using a data-based hydrochemical model to predict conservative tracer data. *Water Resources Research*, 43(3), W03401. <https://doi.org/10.1029/2005wr004019>
- Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., & Marshall, L. A. (2009). Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale. *Water Resources Research*, 45(4), W04428. <https://doi.org/10.1029/2008wr007225>
- Jiang, S., Jomaa, S., & Rode, M. (2014). Modelling inorganic nitrogen leaching in nested mesoscale catchments in central Germany. *Ecohydrology*, 7(5), 1345–1362. <https://doi.org/10.1002/eco.1462>
- Krause, P., Boyle, D. P., & Båse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, 89–97. <https://doi.org/10.5194/adgeo-5-89-2005>
- Kuppel, S., Tetzlaff, D., Maneta, M. P., & Soulsby, C. (2018). ECH₂O-iso 1.0: Water isotopes and age tracking in a process-based, distributed ecohydrological model. *Geoscientific Model Development*, 11(7), 3045–3069. <https://doi.org/10.5194/gmd-11-3045-2018>
- Kuppel, S., Tetzlaff, D., Maneta, M. P., & Soulsby, C. (2020). Critical zone storage controls on the water ages of ecohydrological outputs. *Geophysical Research Letters*, 47(16), e2020GL088897. <https://doi.org/10.1029/2020gl088897>
- Landgraf, J., Tetzlaff, D., Dubbert, M., Dubbert, D., Smith, A., & Soulsby, C. (2022). Xylem water in riparian Willow trees (*Salix alba*) reveals shallow sources of root water uptake by in situ monitoring of stable water isotopes. *Hydrology and Earth System Sciences*, 2021, 1–27. <https://doi.org/10.5194/hess-2021-456>
- Lischeid, G. (2008). Combining hydrometric and hydrochemical data sets for investigating runoff generation processes: Tautologies, inconsistencies and possible explanations. *Geography Compass*, 2(1), 255–280. <https://doi.org/10.1111/j.1749-8198.2007.00082.x>
- Lutz, S. R., Krieg, R., Müller, C., Zink, M., Knöller, K., Samaniego, L., & Merz, R. (2018). Spatial patterns of water age: Using young water fractions to improve the characterization of transit times in contrasting catchments. *Water Resources Research*, 54(7), 4767–4784. <https://doi.org/10.1029/2017wr022216>
- Maneta, M. P., & Silverman, N. L. (2013). A spatially distributed model to simulate water, energy, and vegetation dynamics using information from regional climate models. *Earth Interactions*, 17(11), 1–44. <https://doi.org/10.1175/2012ei000472.1>
- McDonnell, J. J., & Beven, K. (2014). Debates—The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. *Water Resources Research*, 50(6), 5342–5350. <https://doi.org/10.1002/2013wr015141>
- McDonnell, J. J., McGuire, K., Aggarwal, P., Beven, K. J., Biondi, D., Destouni, G., et al. (2010). How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis. *Hydrological Processes*, 24(12), 1745–1754. <https://doi.org/10.1002/hyp.7796>
- McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., et al. (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, 43(7), W07301. <https://doi.org/10.1029/2006wr005467>
- McDonnell, J. J., Spence, C., Karran, D. J., van Meerveld, H. J., & Harman, C. J. (2021). Fill-and-spill: A process description of runoff generation at the scale of the beholder. *Water Resources Research*, 57(5), e2020WR027514. <https://doi.org/10.1029/2020wr027514>
- Mein, R. G., & Larson, C. L. (1973). Modeling infiltration during a steady rain. *Water Resources Research*, 9(2), 384–394. <https://doi.org/10.1029/wr009i002p00384>
- Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2), 161–174. <https://doi.org/10.2307/1269043>
- Nan, Y., Tian, L., He, Z., Tian, F., & Shao, L. (2021). The value of water isotope data on improving process understanding in a glacierized catchment on the Tibetan Plateau. *Hydrology and Earth System Sciences*, 25(6), 3653–3673. <https://doi.org/10.5194/hess-25-3653-2021>
- Neill, A. J., Birkel, C., Maneta, M. P., Tetzlaff, D., & Soulsby, C. (2021). Structural changes to forests during regeneration affect water flux partitioning, water ages and hydrological connectivity: Insights from tracer-aided ecohydrological modelling. *Hydrology and Earth System Sciences*, 25(9), 4861–4886. <https://doi.org/10.5194/hess-25-4861-2021>
- Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79, 214–232. <https://doi.org/10.1016/j.envsoft.2016.02.008>
- Rode, M., Halbedel née Angelstein, S., Anis, M. R., Borchart, D., & Weitere, M. (2016). Continuous in-stream assimilatory nitrate uptake from high-frequency sensor measurements. *Environmental Science & Technology*, 50(11), 5685–5694. <https://doi.org/10.1021/acs.est.6b00943>
- Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46(5), W05523. <https://doi.org/10.1029/2008wr007327>
- Sivapalan, M. (2005). Pattern, process and function: Elements of a unified theory of hydrology at the catchment scale. In *Encyclopedia of hydrological sciences*. <https://doi.org/10.1002/0470848944.hsa012>
- Skøien, J. O., Blöschl, G., & Western, A. W. (2003). Characteristic space scales and timescales in hydrology. *Water Resources Research*, 39(10), 1304. <https://doi.org/10.1029/2002wr001736>
- Smith, A., Tetzlaff, D., Kleine, L., Maneta, M., & Soulsby, C. (2020). Isotope-aided modelling of ecohydrologic fluxes and water ages under mixed land use in Central Europe: The 2018 drought and its recovery. *Hydrological Processes*, 34(16), 3406–3425. <https://doi.org/10.1002/hyp.13838>
- Smith, A., Tetzlaff, D., Kleine, L., Maneta, M., & Soulsby, C. (2021). Quantifying the effects of land use and model scale on water partitioning and water ages using tracer-aided ecohydrological models. *Hydrology and Earth System Sciences*, 25(4), 2239–2259. <https://doi.org/10.5194/hess-25-2239-2021>
- Smith, A., Tetzlaff, D., Landgraf, J., Dubbert, M., & Soulsby, C. (2022). Modelling temporal variability of in-situ soil water and vegetation isotopes reveals ecohydrological couplings in a willow plot. *Biogeosciences*, 19(9), 2465–2485. <https://doi.org/10.5194/bg-2021-278>
- Smith, A., Tetzlaff, D., Maneta, M., & Soulsby, C. (2022). Critical zone response times and water age relationships under variable catchment wetness states: Insights using a tracer-aided ecohydrological model. *Water Resources Research*, 58(4), e2021WR030584. <https://doi.org/10.1029/2021wr030584>
- Soulsby, C., Birkel, C., Geris, J., Dick, J., Tunaley, C., & Tetzlaff, D. (2015). Stream water age distributions controlled by storage dynamics and nonlinear hydrologic connectivity: Modeling with high-resolution isotope data. *Water Resources Research*, 51(9), 7759–7776. <https://doi.org/10.1002/2015wr017888>
- Soulsby, C., Birkel, C., Geris, J., & Tetzlaff, D. (2015). The isotope hydrology of a large river system regulated for hydropower. *River Research and Applications*, 31(3), 335–349. <https://doi.org/10.1002/rra.2740>
- Soulsby, C., Birkel, C., & Tetzlaff, D. (2016). Characterizing the age distribution of catchment evaporative losses. *Hydrological Processes*, 30(8), 1308–1312. <https://doi.org/10.1002/hyp.10751>
- Soulsby, C., Tetzlaff, D., & Hrachowitz, M. (2010). Spatial distribution of transit times in montane catchments: Conceptualization tools for management. *Hydrological Processes*, 24(22), 3283–3288. <https://doi.org/10.1002/hyp.7864>

- Stadnyk, T. A., Delavau, C., Kouwen, N., & Edwards, T. W. D. (2013). Towards hydrological model calibration and validation: Simulation of stable water isotopes using the isoWATFLOOD model. *Hydrological Processes*, 27(25), 3791–3810. <https://doi.org/10.1002/hyp.9695>
- Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., & Tallaksen, L. M. (2011). Comparison of hydrological model structures based on recession and low flow simulations. *Hydrology and Earth System Sciences*, 15(11), 3447–3459. <https://doi.org/10.5194/hess-15-3447-2011>
- Stevenson, J. L., Birkel, C., Neill, A. J., Tetzlaff, D., & Soulsby, C. (2021). Effects of streamflow isotope sampling strategies on the calibration of a tracer-aided rainfall-runoff model. *Hydrological Processes*, 35(6), e14223. <https://doi.org/10.1002/hyp.14223>
- Tetzlaff, D., Buttle, J., Carey, S. K., Kohn, M. J., Laudon, H., McNamara, J. P., et al. (2021). Stable isotopes of water reveal differences in plant – Soil water relationships across northern environments. *Hydrological Processes*, 35(1), e14023. <https://doi.org/10.1002/hyp.14023>
- Tetzlaff, D., McDonnell, J. J., Uhlenbrook, S., McGuire, K. J., Bogaart, P. W., Naef, F., et al. (2008). Conceptualizing catchment processes: Simply too complex? *Hydrological Processes*, 22(11), 1727–1730. <https://doi.org/10.1002/hyp.7069>
- Tetzlaff, D., Soulsby, C., Hrachowitz, M., & Speed, M. (2011). Relative influence of upland and lowland headwaters on the isotope hydrology and transit times of larger catchments. *Journal of Hydrology*, 400(3), 438–447. <https://doi.org/10.1016/j.jhydrol.2011.01.053>
- Thirel, G., Andréassian, V., & Perrin, C. (2015). On the need to test hydrological models under changing conditions. *Hydrological Sciences Journal*, 60(7–8), 1165–1173. <https://doi.org/10.1080/02626667.2015.1050027>
- Wang, H., Tetzlaff, D., & Soulsby, C. (2019). Hysteretic response of sap flow in Scots pine (*Pinus sylvestris*) to meteorological forcing in a humid low-energy headwater catchment. *Ecohydrology*, 12(6), e2125. <https://doi.org/10.1002/eco.2125>
- Wang, L., von Freyberg, J., van Meerveld, I., Seibert, J., & Kirchner, J. W. (2019). What is the best time to take stream isotope samples for event-based model calibration? *Journal of Hydrology*, 577, 123950. <https://doi.org/10.1016/j.jhydrol.2019.123950>
- Weber, U., Attinger, S., Baschek, B., Boike, J., Borchardt, D., Brix, H., et al. (2022). MOSES: A novel observation system to monitor dynamic events across Earth compartments. *Bulletin of the American Meteorological Society*, 103(2), E339–E348. <https://doi.org/10.1175/bams-d-20-0158.1>
- Wollschläger, U., Attinger, S., Borchardt, D., Brauns, M., Cuntz, M., Dietrich, P., et al. (2016). The Bode hydrological observatory: A platform for integrated, interdisciplinary hydro-ecological research within the TERENO Harz/Central German Lowland Observatory. *Environmental Earth Sciences*, 76(1), 29. <https://doi.org/10.1007/s12665-016-6327-5>
- Yang, X., Jomaa, S., & Rode, M. (2019). Sensitivity analysis of fully distributed parameterization reveals insights into heterogeneous catchment responses for water quality modeling. *Water Resources Research*, 55(12), 10935–10953. <https://doi.org/10.1029/2019wr025575>
- Yang, X., Jomaa, S., Zink, M., Fleckenstein, J. H., Borchardt, D., & Rode, M. (2018). A new fully distributed model of nitrate transport and removal at catchment scale. *Water Resources Research*, 54(8), 5856–5877. <https://doi.org/10.1029/2017wr022380>
- Yang, X., Rode, M., Jomaa, S., Merbach, I., Tetzlaff, D., Soulsby, C., & Borchardt, D. (2022). Functional multi-scale integration of agricultural nitrogen-budgets into catchment water quality modeling. *Geophysical Research Letters*, 49(4), e2021GL096833. <https://doi.org/10.1029/2021gl096833>
- Yang, X., Tetzlaff, D., Soulsby, C., & Borchardt, D. (2022a). EcH₂O-iso v3.1: The tracer-aided ecohydrological model for larger scale catchments [Software]. Zenodo. <https://doi.org/10.5281/zenodo.7117112>
- Yang, X., Tetzlaff, D., Soulsby, C., & Borchardt, D. (2022b). Upscaling tracer-aided ecohydrological EcH₂O-iso model in larger catchments: Model setup and model simulations in the Selke catchment, central Germany [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.6628889>
- Yang, X., Tetzlaff, D., Soulsby, C., Smith, A., & Borchardt, D. (2021). Catchment functioning under prolonged drought stress: Tracer-aided ecohydrological modeling in an intensively managed agricultural catchment. *Water Resources Research*, 57(3), e2020WR029094. <https://doi.org/10.1029/2020wr029094>
- Zambrano-Bigiarini, M. (2020). hydroTSM: Time series management, analysis and interpolation for hydrological modelling. R package version 0.6-0. <https://doi.org/10.5281/zenodo.839864>
- Zhang, X., Yang, X., Jomaa, S., & Rode, M. (2020). Analyzing impacts of seasonality and landscape gradient on event-scale nitrate-discharge dynamics based on nested high-frequency monitoring. *Journal of Hydrology*, 591, 125585. <https://doi.org/10.1016/j.jhydrol.2020.125585>
- Zhang, Z.-Y., Schmidt, C., Nixdorf, E., Kuang, X., & Fleckenstein, J. H. (2021). Effects of heterogeneous stream-groundwater exchange on the source composition of stream discharge and solute load. *Water Resources Research*, 57(8), e2020WR029079. <https://doi.org/10.1029/2020wr029079>