EcoTWIN 1.0: A Fully Distributed Tracer-Aided Ecohydrological Model Tracking Water, Isotopes, and Nutrients

Songjun Wu[1], Doerthe Tetzlaff[1,2], Yi Zheng[3], Chris Soulsby[4]

1 Department of Ecohydrology and Biogeochemistry, Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

2 Geography Institute and IRI THESys, Humboldt University of Berlin, Berlin, Germany

3 School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China

4 Northern Rivers Institute, School of Geosciences, University of Aberdeen, UK

Corresponding author: Songjun Wu (songjun.wu@igb-berlin.de)

Abstract

The value of stable water isotopes in constraining process representation in hydrological models is well acknowledged with numerous tracer-aided hydrological models developed in recent years, yet few have leveraged these benefits for more robust water quality modelling. Therefore, we introduce EcoTWIN, a fully distributed tracer-aided ecohydrological model that simultaneously ***tracks water***, ***isotope***, and ***nutrient*** fluxes. A thorough model test was conducted by calibrating EcoTWIN against discharge, in-stream isotopes, and NO_3-N concentrations (1980-2024) in 17 large-scale (10^3 - 10^5 km^2) European catchments spanning a wide range of geographic and climatic gradients. Furthermore, three reanalysis products (ERA5 snow depth, MODIS evapotranspiration, and GRACE surface water anomaly) were employed to further validate the capacity of EcoTWIN to reproduce associated but uncalibrated internal water fluxes. Results showed good model performance of both calibrated in-stream targets and uncalibrated internal fluxes in most catchments. Therefore, we conclude that EcoTWIN is a flexible, transferable modelling tool for prediction and process inference in terrestrial ecosystems ranging from boreal to subtropic climates. Constrained by tracer simulations, the model not only captures the celerity, but also the velocity of hydrological fluxes, thus providing spatio-temporally-explicit estimations of water ages and travel times. Such information provides opportunities to bridge catchment hydrology and water quality by linking travel times with biogeochemical processing kinetics. We demonstrate this with a proof of concept using Damköhler Number in nitrogen modelling.

1 Introduction

The development of ecohydrological models has been accelerating in the recent decades towards frameworks that are more spatially distributed (instead of lumped or semi-distributed) and complex (integrating more ecohydrological processes) (Pechlivanidis et al., 2011; Wellen et al., 2015). A few examples include SWAT (Arnold et al., 2012), HYPE (Lindström et al., 2010), and mHM-Nitrate (Yang et al., 2018), which have been widely applied worldwide. As process-based models, they are used not only as prediction tools for specific variables, but also as learning tools for model inference, i.e., to

track the internal states/fluxes from available observations (Wang et al., 2024). This, however, poses challenges due to the considerable uncertainties in model inference.

Inference of internal processes is naturally uncertain due to the lack of direct observations, though such uncertainty can be constrained to some extent by rigorous split-sample calibration and validation. The reason we use "~~somehow~~to some extent" here is based on the fact that most models are calibrated to a minimal number of variables, and 81% of calibrations used data from a single gauge (mostly at a catchment outlet) as reviewed in Wellen et al., (2015). Additionally, from a technical perspective, "equifinality" further adds to the inference uncertainty due to the potential misinformation in data (uncertainty in model forcing and observations) and model structure (the use of simplified, abstract mathematics to simulate real world processes) (Beven, 2006). This can result in inaccurate process representations yielding deceptively good results through error compensation, thus leading to overconfidence in a model's ability to reproduce within-basin dynamics (Wen et al., 2024; Wu et al., 2025a). As acknowledged by the hydrological community, models calibrated solely against discharge at the catchment outlet reflect only the celerity of hydrological systems (pressure wave propagation), yet constituent transport in water quality modelling relies on the velocity (mass flux of the water) (McDonnell & Beven, 2014). Failure to reconcile these differences can lead to questionable process inferences from many ecohydrological and water quality models.

One way to strengthen model inference is to include auxiliary data for calibration (Efstratiadis & Koutsoyiannis, 2010). However, there is a paradox in multi-criteria calibration, as on the one hand, more auxiliary data will feed unique information to the calibration process, thus effectively constraining the model behaviour from an ecohydrological perspective; yet on the other hand, it increases the dimensionality of calibration thus resulting in degraded performance or failure of calibration from a technical perspective. The "curse" of dimensionality in ecohydrological modelling is universal for all the commonly used algorithms under both Bayesian and Pareto theories as demonstrated in Wu et al., (2025c). Therefore, modellers should expect the selected auxiliary data to contain as much information as possible (Nearing et al., 2020). For distributed modelling, the auxiliary data should reflect the cumulative contribution of all upstream reaches/regions, rather than variables that are highly dependent on local condition/processes (e.g. point-scale soil moisture and evapotranspiration measurements etc.).

Stable water isotopes, in this context, have powerful potential in cumulative flux tracking. As conservative tracers, ^{2}H and ^{18}O are independent of biogeochemical reactions and naturally integrate landscape heterogeneity, thus providing effective constraints on spatially distributed (dis)connections of hydrological flow paths as well as velocity of the hydrological systems which reflect flux-storage interactions (Jung et al., 2025; Tetzlaff et al., 2015). The value of tracers has long been recognised by hydrologists (Hooper et al., 1988), with many tracer-aided hydrological models developed and evolved in recent years from lumped (Birkel et al., 2011; Godsey et al., 2010), to semi-distributed (van Huijgevoort et al., 2016; Nan et al., 2021), and distributed structure (Kuppel et al., 2018; Remondi et al., 2018). However, few attempts have been made to integrate a tracer-aided hydrological structure into water quality modelling (Birkel & Soulsby, 2015; Jung et al., 2025), despite the need being evident

for nearly four decades (Neal et al., 1988). Moreover, existing pioneering models are mostly conceptualised/lumped (Benettin et al., 2015; Dick et al., 2015) and/or loosely coupled via external tracer/water quality modules (Pesántez et al., 2023; Yang et al., 2024; Zhang et al., 2020). The external coupling of model chains transfer necessary internal states and fluxes between sub-models (e.g. hydrological fluxes for constituent mixing in water quality or isotopic modules) via online in-memory coupling (instead of offline on-disk coupling), thus significantly increasing the resources consumption in input/output operations. Such model chains, though providing useful scientific insights, can become problematic for large-scale applications owing to the exponential growth in computational and storage requirements. Therefore, there remains a need to develop a fully distributed, computationally efficient ecohydrological model that combines hydrological, isotopic, and water quality simulations.

This research gap motivated the development of EcoTWIN, the model that we present in this paper. To our knowledge, the model is one of the first distributed tracer-aided ***ecohydrological*** models that ***tracks water***, ***isotopic***, and ***nutrient*** fluxes simultaneously in a C++-based framework. For a thorough testing of EcoTWIN, 17 large European catchments were selected for calibration against discharge, in-stream isotopes, and NO_3-N concentrations. These catchments span over a wide range of geographic (Alpine to lowland plain) and climatic (from snow-dominated to Mediterranean) gradients. In addition, the robustness of modelled inference on uncalibrated internal fluxes were also compared with three remote sensing products (snow depth, evapotranspiration, and water storage). Given the overall good integrated performance in most catchments, EcoTWIN is presented as an ecohydrological modelling framework applicable for terrestrial ecosystems ranging from boreal to temperate and subtropical climates across a wide range of geographical environments. The subsequent sections are organised as follows: Section 2 and 3 introduce the model structure of EcoTWIN and details in calibration and validation; the model performance is shown in Section 4; in Section 5 we show the advantages of a tracer-aided ecohydrological framework with an example of how water ages bridge catchment hydrology and water quality models; finally, the current limitations and planned future development of EcoTWIN are also discussed.

2 Model description

EcoTWIN is fully distributed ecohydrological model implemented in C++. The model consists of hydrological, isotopic, and nitrogen modules, which simulate major ecohydrological states and fluxes from canopy to groundwater (Figure 1). Aided by tracer simulations, the model is additionally able to track the water movement vertically and laterally via the calculation of water ages and travel times. For detailed information of model parameters please refer to Table S1.

2.1 Hydrological module

EcoTWIN follows a typical multi-layer, top-down, bucket-type approach that resolves the water balance sequentially for the vegetation canopy, three soil layers, unsaturated zone, and groundwater.

As the foundation of solute transport, the hydrological module employs a selective disassembly structure with multiple alternative conceptualisations possible for specific important hydrological processes. The configuration can be specified *a priori* based on the goal of modelling and prior knowledge of the studied catchment(s).

2.1.1 Soil properties

Before iterative simulations, soil characteristics are estimated using appropriate pedotransfer functions. Three different alternatives are provided, each of which requires different levels of inputs but all were found to provide robust estimation of soil porosity (θ_s), field capacity (θ_{fc}), wilting point (θ_{wp}), and hydraulic conductivity (Ks). All the soil properties are required for each soil layer/depth. This can be achieved via three alternative options: (i) assigning identical properties across the whole soil column, (ii) calculating separately for each depth based on depth-dependent inputs, or (iii) extrapolating deeper profile characteristics from the top soil properties based on a depth-dependent equation in Maneta & Silverman, (2013).

The distribution of soil types and land use is assigned from raster file in EcoTWIN. This can be specified as a static boundary condition; alternatively, the distributions can also be updated dynamically via a user-specified interval to reflect any temporal changes due to land management.

2.1.2 Vertical fluxes

The vertical fluxes are resolved for storages in the canopy, soil layers, unsaturated zone, and groundwater. The mass balance of canopy storage (ΔC) follows:

$$\Delta C = P - I - Th \tag{1}$$

where P, I, Th are precipitation, interception, and throughfall, respectively. The throughfall is calculated as the exceedance of current canopy storage from the maximum storage calculated by Leaf Area index LAI and a correlation parameter α.

$$C_{max} = \alpha * LAI \tag{2-1}$$

Alternatively, the maximum canopy storage can be estimated with explicit consideration of precipitation intensity (Landgraf et al., 2023):

$$C_{max} = \alpha * LAI * (1 - \frac{1}{1 + SCF * P/(\alpha * LAI)}) \tag{2-2}$$

where SCF is the vegetation cover fraction calculated by LAI and an extinction coefficient (rE) adopted from HYDRUS-1D (Šimůnek et al., 2013):

$$SFC = 1 - \exp(rE * LAI) \tag{3}$$

Then throughfall is calculated as the exceedance of canopy storage from the maximum:

$$Th = (P + C) - C_{max}\ if\ (P + C) > C_{max}\ else\ 0 \tag{4}$$

After reaching land surface, throughfall becomes ponding water (S_{Pond}) or snow (S_{snow}) depending on a temperature threshold for separation ($Thre_{SN}$). Snow will melt and recharge the ponding water in warm conditions (air temperature Ta exceed $Thre_{sn}$) following a degree-day model.

$$melt = S_{snow} * min(dd_{min} + dd_{inc} * Th * (Ta - Thre_{SN}), dd_{max}) \quad (5)$$

Where dd_{min} and dd_{max} are the minimum and maximum of degree day factor, while dd_{inc} denotes the rate of increase in the degree-day factor per degree Celsius rise in temperature.

The available ponding water infiltrates into the top soil layer using Green-Ampt model (Kale & Sahoo, 2011; Maneta & Silverman, 2013), with infiltration capacity first calculated as a function of average soil moisture over the hydrologically active depth:

$$I_f = Ks * (1 + \frac{\psi * \theta_s * (1 - (\theta_1 - \theta_{wt})/(\theta_s - \theta_{wt}))}{\theta_1 * d_1}) \quad (6)$$

Where θ_1, θ_s, θ_{wt}, and d_1 are the moisture content, porosity, wilting point, and depth in top soil layer; ψ is a parameter representing soil air entry pressure in m. Then potential infiltration (F_p) is determined from the lesser between the available ponding water (S_{pd}) and potential infiltration rate integrated over time before ponding occurs ($I_f * t_p$).

The actual infiltration (F) is solved iteratively using the Newton–Raphson scheme:

$$F = \Delta\theta * d_1 = F_p + Ks * w_{Ks} * (\Delta t - t_p) - \psi\Delta\theta * \ln(\frac{\psi\Delta\theta + \Delta\theta d_1}{\psi\Delta\theta + F_p}) \quad (7)$$

where w_{Ks} is anisotropy ratio of vertical to horizontal Ks.

The soil storage in each layer is conceptualised as two water pools – a gravitational, free-flowing pool and a capillary, soil-bound pool. The two pools are separated based on field capacity (Maneta & Silverman, 2013), and percolation happens when soil storage exceeds the threshold. Three alternative schemes are included in EcoTWIN.

In the first scheme, all water in excess of field capacity percolates to deeper layer:

$$Pc_i = (\theta_i - \theta_{fc}) * d_i \quad (8\text{-}1)$$

where Pc_i, θ_i and d_i depict the percolation, moisture content and depth from/in ith soil layer in m.

The second scheme additionally considers the hydraulic conductivity (Ks) following the conceptualisation in SWAT (Arnold et al., 2012):

$$Pc_i = (\theta_i - \theta_{fc,i}) * d_i * (1 - \exp(\frac{-\Delta t * Ks}{\theta_{s,i} - \theta_{fc,i}})) \quad (8\text{-}2)$$

The third scheme relates percolation to the extent of soil saturation with an exponential parameter β (Kumar et al., 2013; Samaniego et al., 2010):

$$Pc_i = (\theta_i - \theta_{fc,i}) * d_i * (1 - \exp(\beta * \log(\theta_i/\theta_{s,i}))) \quad (8\text{-}3)$$

For evapotranspiration, soil evaporation and transpiration are estimated separately. The separation of PET is realised by surface cover fraction introduced above:

$$PT = PET * SCF; \quad PE = PET - PT \tag{9}$$

Soil evaporation is simulated in the top soil layer based on the soil saturation:

$$Evap_s = PE * \min\left(\frac{\theta_1}{\theta_{fc,1}}, 1\right) \tag{10}$$

Transpiration is simulated in all soil layers based on the fractions ($f_{root,i}$) of root density ($D_{root,i}$) in each layer partitioned by soil depth and a parameter (γ_{root}):

$$Tr_i = PT * f_{root,i} * \frac{\theta_1 - \theta_{wp,1}}{\theta_{fc,1} - \theta_{wp,1}} \tag{11}$$

$$f_{root,i} = D_{root,i} / \sum_{j=1}^{m} D_{root,j} \tag{12}$$

$$D_{root,i} = \left(1 - \gamma_{root}^{\left(\sum_{j=1}^{m} d_j\right)}\right) - \left(1 - \gamma_{root}^{\left(\sum_{j=1}^{i} d_j\right)}\right) \tag{13}$$

Channel evaporation is also estimated using Penman equation, which relies on net radiation, wind speed, air pressure, and air temperature as inputs.

The last soil layer percolates to an unsaturated storage in unsaturated zone (S_{unsat}). The compartment stores the excess water from soil and percolates either downward to groundwater storage (S_{GW}) or laterally downstream. The percolation to groundwater Pc_{GW} is determined by a weighting parameter p_{GW} as a proportion of unsaturated storage:

$$Pc_{GW} = S_{unsat} * p_{GW} \tag{14}$$

Additionally, irrigation is conceptualised in EcoTWIN, which is realised via the water extraction from river or groundwater. The source is determined by the geographic location: for a grid cell with channel network, water is extracted directly from river, and local groundwater is used as irrigation source for non-channel grids. The amount of extraction is estimated from a predefine coefficient for crop water demands (w_{irr}) from which the deficit is calculated for each of the m soil layers.

$$deficit = \sum_{i=1}^{m} \left(\theta_{fc,i} - \theta_{wp,i}\right) * w_{irr} * d_i \tag{15}$$

Note that the irrigation can switch to groundwater extraction if river storage cannot fill the deficit.

2.1.3 Lateral fluxes

In EcoTWIN, grid cells are connected laterally at three levels - surface, unsaturated zone, and groundwater. Note that some models omit the unsaturated storage and directly calculate excess water to drain based on the saturation extent of the bottom soil layer (e.g., EcH_2O-iso, Kuppel et al.,

2018). EcoTWIN did not follow this conceptualisation because in reality, the lateral drainage is focused in the saturated zone, and thus the bottom of the soil layer instead of the whole soil profile. The drainage of an entire soil layer thus brings considerable uncertainty to the velocity of lateral transport when the lower boundary of the soil is a parameter to tune in calibration. For instance, a large soil depth will dramatically reduce the velocity of interflow drainage and slow down the mixing of constituents, though this might still perfectly reproduce the celerity (hydrograph) for purely hydrological modelling. Our conceptualisation (an independent unsaturated compartment) aligns with most hydrological models (Arnold et al., 2012; Yang et al., 2018) and fits the recent analysis supporting the dominant role of lateral drainage over vertical transports globally (Mcmillan et al., 2025).

By the end of each timestep, ponding water receives upstream inputs and contributes to channel storage if the grid is connected to the channel network, while non-channel grid has $Ovf_C = 0$:

$$Ovf_C = (Ovf_{T,in} + S_{pond}) * p_{Ovf} * dx_C/dx_T \quad (16)$$

dx_C and dx_T are the channel length and size of terrestrial grid cell; p_{Ovf} is a weighting parameter for channel recharge. Then the remaining ponding water routes to downslope terrestrial grid following the topographic gradient. In none-channel grid cells, all available ponding storage routes lateral downstream ($Ovf_C = 0$):

$$Ovf_{T,out} = \left(Ovf_{T,in} + S_{pond}\right) - Ovf_C \quad (17)$$

Similarly, unsaturated storage contributes first to channel storage in grid cells within channel network, while non-channel grid cells have $Inf_C = 0$:

$$Inf_C = \left(Inf_{T,in} + S_{vadose}\right) * K_{vadose} * \left(1 - e^{-1*exp_{Inf}*\left(Inf_{T,in}+S_{vadose}\right)}\right) * p_{Inf} \quad (18)$$

where K_{vadose} is the effective conductivity of lateral transport in the unsaturated zone; while exp_{Inf} is an exponential parameter determining the strength of positive correlation between recharge and current unsaturated storage. Then the remaining unsaturated storage is partially routed to downslope grid cell following a linear approximation of Kinematic wave equation, which assumes gravitational flux per unit width $Inf_{T,out}$ is proportional to the subsurface hydraulic conductivity (K_{vadose}) and bedrock slope ($slope$ approximated from the surface slope):

$$Inf_{T,out} = \left(Inf_{T,in} + S_{unsat} - Inf_C\right) * \left(1 + \alpha * \frac{dt}{dx}\right) * \alpha * \frac{dt}{dx} \quad (19)$$

$$\text{where } \alpha = K_{unsat} * \sin(slope)$$

Groundwater routing is similar to that of interflow, with channel recharge followed by terrestrial transport. Note that the terrestrial groundwater flow does not consider the bedrock slope as groundwater storage is generally much large than unsaturated storage, and thus independent from topographic gradients:

$$GWf_C = \left(GWf_{T,in} + S_{GW}\right) * K_{GW} * \left(1 - e^{-1*exp_{GWf}*\left(GWf_{T,in}+S_{GW}\right)}\right) * p_{GWf} \quad (20)$$

$$GWf_{T,out} = \left(GWf_{T,in} + S_{GW} - GWf_C\right) * \left(1 + \alpha * \frac{dt}{dx}\right) * \alpha * \frac{dt}{dx} \quad (21)$$

where $\alpha = K_{vadose}$

The channel routing is realised using Kinematic wave equation based on a scaled channel roughness parameter (Maneta & Silverman, 2013).

2.2 Isotopic module

The isotopic module in EcoTWIN tracks the composition of stable water isotopes in all water storage compartments following hydrological mixing and fractionation. The module also provides estimation of water age and travel time conceptualised as the time since water molecules enter the catchment as precipitation, and the time water molecules need to travel through the specific storage.

2.2.1 Mixing

The mixing and transport of isotopes (^{2}H and ^{18}O, both noted as C) are governed by the velocity of hydrological fluxes with a complete mixing strategy for most water storages:

$$\frac{d(V * C)}{dt} = \sum_{k=1}^{N_{in}} q_{in,k} * C_{in,k} - \sum_{k=1}^{N_{out}} q_{out,k} * C \quad (22)$$

Where V and C are the volume and composition/concentration of the storage, while N_{in} and N_{out} denote the number of influx and outflux. Such strategy is built on two assumptions: (i) constitutes (i.e., isotopes) are fully mixing within each timestep; (ii) the composition/concentration in outflow equals to that in storage. Additional mixing between ponding and upper soil water storage is allowed (with proportion determined by a parameter $SrfMixing$), as nutrients in top soils can be flushed out in large hydrological events (Seybold et al., 2022).

The full-mixing assumptions have been widely used and shown to be reasonable for storages with relatively small volumes in many mixing/water quality models (Arnold et al., 2012; Yang et al., 2018). However, some studies show that a complete mixing strategy can be problematic for large storages such as groundwater as they are generally poorly constrained (e.g. Soulsby et al., 2015). Therefore, the mass conservation equation used in the INCA-N model and mHM-Nitrate is employed to calculate the mixing of groundwater storages with influxes (i.e., percolation from unsaturated storage and lateral groundwater inflow).

$$\frac{dC}{dt} = \frac{1}{V + V_r} * \left(\sum_{k=1}^{N_{in}} q_{in,k} * C_{in,k} - \sum_{k=1}^{N_{out}} q_{out,k} * C\right) \quad (23)$$

where V_r is the retention storage. The equation is solved by the fourth order Runge-Kutta technique.

2.2.2 Fractionation

As conservative tracers, the composition of isotopes in water storages/fluxes is only changed by kinetic fractionation apart from hydrological mixing. The process is accompanied by evaporation, resulting in the preferential loss of lighter isotopes (^{1}H and ^{16}O) to the vapor phase and a corresponding enrichment of heavier isotopes (^{18}O and ^{2}H) in the residual water. In EcoTWIN, the fractionation is simulated along with evaporation of top soil water and river storage based on the Craig-Gordon model (Craig et al., 1964; Kuppel et al., 2018), while transpiration is assumed to be a non-fractionating process (Dawson & Ehleringer, 1991; Kuppel et al., 2018).

$$C = C^* - (C^* - C) * \left(\frac{S - Evap}{S}\right)^m \tag{24}$$

where C^* and m are the limiting isotopic composition (in ‰) and the dimensionless enrichment slop that are estimated via the following equations in (Good et al., 2014):

$$C^* = \frac{h_a C_a + h_s \varepsilon^+ + \varepsilon^k}{h_s - h_a + \varepsilon^k/1000} \tag{25}$$

$$m = \frac{h_a - (h_s \varepsilon^+ + \varepsilon^k)/1000}{h_s - h_a + \varepsilon^k/1000} \tag{26}$$

where h_a is the relative humidity above the soil surface normalised from atmospheric relative humidity (h), air temperature (T_a), and soil temperature (T_s estimated from Amato & Giménez, 2024). C_a is the isotopic composition of ambient air moisture estimated from precipitation composition:

$$C_a = (C_{rain} - \varepsilon^+)/\alpha^+ \tag{27}$$

where ε^+ is the equilibrium fractionation factor (Skrzypek et al., 2015); α^+ is a temperature factor estimated from T_a.

$$\varepsilon^+ = (1 - 1/\alpha^+) * 1000 \tag{28}$$

The factor of diffusion-controlled kinetic isotopic separation ε^k is calculated based on the relative humidity of soil surface (h_a) and soil pore (h_s).

$$\varepsilon^k = (h_s - h_a) * \left(1 - \frac{D_i}{D}\right) * n \tag{29}$$

Where D_i and D denote the diffusivities of water vapor molecules containing heavier isotope and the lighter isotope, respectively. The ratio can be acquired in Horita et al., (2008) for ^{2}H (0.9877) and ^{18}O (0.9859). n is an advection term ranging between 0.5 (in saturated soils) and 1 (in dry soils). The factor is included in calibration for the fractionation of top soil evaporation yet fixed as 0.5 for that of channel evaporation.

2.2.3 Water age and travel time

EcoTWIN can track the age of water i.e., the time since water enters the catchment as precipitation, in each storage. In age tracking, precipitation is defined as new water with age of zero. At the end of each time step, water ages of all storages are advanced based on the temporal resolution (for instance one day if the model is set up for daily timesteps). Note that in some circumstances, the modellers might need to disable the age evolution of specific storage(s) (e.g., groundwater storage) as the storage can be too large to achieve steady states in model spin-up. Similar to isotopes, water ages are only controlled by hydrological transport with the same mixing strategy (i.e., complete mixing except for groundwater).

The water ages in EcoTWIN are the mean values averaged from all water molecules in the storage, which might be dominated by the inflow of very old water that obscure the age distribution of the young water (e.g., the groundwater input to top soils due to the groundwater extraction for irrigation). Therefore, EcoTWIN additionally provides the estimation of travel time - the time of water molecule travelling through each storage. The simulation is similar to that of water ages. The only difference is that the transition of water between storges (e.g., percolation into deeper soil layers) resets the travel time to zero. Accordingly, all the water enters a new storage becomes new water instead of just precipitation in water age tracking.

2.3 Nitrogen module

The nitrogen module describes the mass balance of nitrogen, particularly nitrate as the main form of dissolved nitrogen, which is dominated by the interaction of hydrological transport and biogeochemical transformations.

For each timestep, the nitrate concentration is simulated in each storage following three processes – hydrological transport/mixing, nitrogen inputs, and biogeochemical transformations. Fully integrated with hydrological module, nitrate transport also aligns with hydrological fluxes following the same mixing strategy as in the isotopic simulation. For nitrogen sources, EcoTWIN considers the inputs from fertiliser, manure, and plant residues, whose annual inputs can be specified via configuration. Notably, fertilization can be parameterised via spatial raster inputs if corresponding dataset is available. The timing and extent of nitrogen addition of all sources are determined following the implementation in HYPE (Lindström et al., 2010), which distributes the annual sum across a specified period (e.g., the period between planting and harvest for crops). Additionally, wet deposition is conceptualised as the atmospheric nitrogen source, whose concentration can be specified via spatial raster and simply as a constant value.

The biogeochemical transformations are mainly modified from the mHM-Nitrate model (Yang et al., 2018), and the HYPE model (Lindström et al., 2010), which are conceptualised for the soil profile and channel network. In the soil profile, three nitrogen pools are conceptualised for each soil layer, including an inactive nitrogen pool (SN_i), an active nitrogen pool (SN_a), and a dissolved nitrate pool

(DN). The soil transformations include degradation (Dgd_s, from SN_i to SN_a), mineralisation ($Minr_s$, from SN_a to DN), denitrification ($Deni_s$, from DN to gaseous N_2), and plant uptake ($Uptk_s$, DN removal).

$$Dgd_s = SN_i * ref_{Dgd,s} * f_{Ta} * f_\theta / dt \quad (30)$$

$$Minr_s = SN_a * ref_{Minr,s} * f_{Ta} * f_\theta / dt \quad (31)$$

$$Deni_s = DN * ref_{Deni,s} * f_{Ta} * f_{\theta,deni} * f_{conc,s} / dt \quad (32)$$

where $ref_{Dgd,s}$, $ref_{Minr,s}$, $ref_{Deni,s}$ are the parameters representing the reference rates of soil degradation, mineralisation, and denitrification. f_{Ta} and f_θ are the regulating factors of temperature and moisture.

$$f_{Ta} = 2^{(T_a-20)/10} * \omega \quad where\ \omega = \begin{cases} 1 & T_a > 5 \\ T_a/5 & 0 \le T_a \le 5 \\ 0 & T_a > 0 \end{cases} \quad (33)$$

$$f_\theta = min\left[\frac{\left(1 - p_{\theta,deni}\right) * (\theta_{fc,i} - \theta_i)}{p_{\theta,fc} * d_i}, \frac{(\theta_i - \theta_{wp,i})}{p_{\theta,wp} * d_i}\right] \quad (34)$$

where$p_{\theta,fc}$ and $p_{\theta,wp}$ are the empirical factors that are fixed as 1.2, 0.8 based on literature values (Lindström et al., 2010; Yang et al., 2018). $p_{\theta,deni}$ is the saturation threshold for soil denitrification ranging between 0.4 – 0.85 (Yang et al., 2018). A different moisture factor considering a saturation threshold (θ_{thres}) is employed for denitrification, as denitrification is more sensitive to the soil wetness condition:

$$f_{\theta,deni} = \left[(\theta_i/\theta_{fc,i} - \theta_{thres})/(1 - \theta_{thres})\right]^{2.5} \quad (35)$$

The process is additionally controlled by the concentration level in the storage $f_{conc,s} = C/(C + 10)$. Plant uptake is simulated using a three-parameter logistic growth equation in (Eckersten et al., 1994; Lindström et al., 2010).

Currently, in-stream denitrification is the only process considered in EcoTWIN.

$$Deni_w = ref_{Deni,w} * f_{T_w} * f_{conc,w} * A/dt \quad (36)$$

where $ref_{Deni,w}$ is the reference in-stream denitrification rates. The actual rates are regulated by a concentration factor $f_{conc,w} = C/(C + 1.5)$ and a temperature factor f_{T_w} (the same equation for f_{Ta} with inputs substituted by river temperature f_{T_w}, simplified as the rolling-average of 20-day air temperature).

It should be noted that the calibrated soil depth in this study is about 2.5 m, with intermittent saturation occurring in the deeper layer. This means that terrestrial denitrification is a combination of soil and groundwater processes in this study, though this might change in other applications if a shallow soil depth is assigned.

3 Model calibration and validation

A robust model application should not only reproduce observed variables through calibration but also yield realistic estimates of internal states and fluxes that are not included in the calibration process. This is essential to avoid situations where inaccurate process representations produce deceptively good results through error compensation. Therefore, we evaluate EcoTWIN from both perspectives. First, we assess the model's ability to reproduce observations via calibration (methods and results in Sections 3.2 and 3.4). Then, we examine the model's capacity to simulate uncalibrated internal states and fluxes by comparing the simulated snow depth, evapotranspiration, and total water storage with corresponding remote-sensing products (methods and results in Sections 3.3 and 3.5).

To ensure model generality, 17 catchments were selected for calibration and validation depending on the data availability (particularly stream stable water isotopes and nitrate), which span a wide range of characteristics in geography, climate, and anthropogenic managements (~~Figure 2 and Table 1).~~catchment locations, characteristics, and major hydrological and nitrogen inputs are shown in Figure 2, Table 1, and Figure S1, respectively). Anthropogenic management practices have a less dramatic effect than climate and geography in most catchments due to the relatively low proportion of urbanized areas. However, a few notable exceptions—such as the Rhine, Elbe, and Danube catchments—are included in the analysis, as these densely populated regions hold critical ecological, agricultural, and economic importance for Europe, and are subject to intensive human interventions in water management. This also provides a chance to examine the applicability of EcoTWIN in human-affected catchments.

3.1 Model setup

EcoTWIN was setup for each of the 17 catchments for calibration with a spatial resolution of 5 km^2 and a temporal resolution of daily timesteps from 1980 to 2024 (with first two years for spin-up). As a fully distributed model, gridded GIS inputs are used in the model setup, including a digital elevation model, flow direction, slope, channel width, channel length, proportion of each land use type (Winkler et al., 2021), proportions of each soil type (world soil map, WRB2014), and soil properties (e.g., depth-dependent proportions of clay, sand, silt, and organic matter from SOILGRIDS). All spatial inputs were acquired with finer resolution (50 m or above) and resampled to the resolution of this application (5 km).

The climatic variables used to drive EcoTWIN include precipitation, air temperature, potential evapotranspiration, relatively humidity, and a few variables that are optional required for the calculation of channel evaporation (air pressure, net radiation, and wind speed). These climatic variables are available from the reanalysis products ERA5 and E-OBS, while PET is calculated using FAO Penman-Monteith equation. For nitrogen simulations, additional inputs are needed including the fertilization map (Grizzetti et al., 2021) and nitrate concentration of rainfall (Zhu et al., 2025) as the boundary of nitrogen addition from agricultural activities and wet deposition.

3.2 Model calibration

Method. The calibration was conducted separately for each catchment to test the applicability of EcoTWIN under different geological and climatic contexts. Three commonly used variables for hydrological and water quality modelling (discharge, stream water isotope composition, and in-stream NO_3-N concentrations) are employed for calibration. Their long-term time series were acquired at daily steps from different sources (discharge from GRDC, isotopes from Wateriso and GNIR, and NO_3-N concentration from global water quality database, GEMStat), and then compared with simulation results at multiple sites for each catchment. Here ^{18}O was selected for isotopic validation due to its higher precision and data abundance. Given the discrepancy in duration of observations (especially for isotopes and NO_3-N), a separate calibration and validation based on a split-sample approach is difficult. Therefore, the full timescale (1982 - 2024) was used for calibration (and the validation introduced in Section 3.3).

The DiffeRential Evolution Adaptive Metropolis algorithm (DREAM) was selected for parameter optimisation due to its relatively efficient and effective performance for high-dimensional problems (as benchmarked in Wu et al., 2025c). The algorithm was implemented separately for each catchment with the same prior distribution of parameters (Table S1). The maximum iteration was set as 100,000 for each catchment (20 chains with maximum chain length of 5000), from which 40 best simulations were selected from the posterior distribution. The Kling-Gupta efficiency (KGE) statistic was used to construct an informal likelihood function for DREAM optimisation.

$$l = \left[\sum_{i=1}^{N_{obs}} \sum_{j=1}^{N_{site}} (1 - KGE) * w_{i,j} \right]^{-m} \quad (37)$$

Where l is the likelihood; N_{obs} and N_{site} are the number of observation types (discharge, isotopes, and nitrate) and sites. The weight $w_{i,j}$, defined for observation type i at site j, is assigned equally across sites such that the total weight for each observation type sums to 1/3. m is an exponentially coefficient to stretch the likelihood surface that is often set based on the number of observation points. After prior test run, m was set as 500. Finally, the likelihood function is transformed to logarithmic form for numeric stability. The calibration was validated using Kling-Gupta efficiency (KGE), Root Mean Square Error (RMSE), Pearson Correlation Coefficient (Coefficient), and Percent bias (Pbias) (Table 2).

Calibration performance. EcoTWIN successfully reproduced the observed discharge in all 17 catchments with KGE exceeding 0.5 at most site (Figure ~~3a~~3A&D; Table 2). Such performance is comparable to or better than previous continental calibration of hydrological models (e.g., ParFlow, Naz et al., 2023; E-HYPE, Donnelly et al., 2016).

Similarly, isotopic and nitrate simulations also produced acceptable performances at most sites (Figure ~~3b~~3B; Table 2). However, there are a few exceptions. The failure of isotopic simulations was found at two sites within the ~~Alphine~~Alps region ~~(bottom left corner of figure 3b).~~ This can be attributed to the uncertainty in precipitation isotopes and snowmelt isotopes (due to the lack of snowmelt elution

fractionationsnow elusion fractionation; Ala-aho et al., 2017), the incorrect isotopic composition in groundwater, or the reduced applicability of degree-day model for mountainous areas in Europe. Such simulation deviation due to the uncertainty in data and boundary initialisation is often reported in previous calibration (Smith et al., 2021).

In general,As for nitrogen simulation (Figure 3C), the model produces comparable performances to existing nitrogen modelling at catchment (Wu et al., 2022, 2025b; Yang et al., 2018) and continental scales (Jones et al., 2023; Mikayilov et al., 2015). However, nitrateThe degraded nitrogen simulations failed to capture the observations at three sites, though as is shownwere found coarsely across Europe, but they are mostly in Figure S1, these all have relativelyrivers/stream with low levels of NO_3-N concentrations given the low RMSE in Table 2. Such low average values can easily trigger the degradation in KGE as one of the sub-components of KGE is highly sensitive to the mean deviation, though the absolute deviation remained low (Figure S1 and Table 2).. Overall, we concluded that EcoTWIN has good capacity to reproduce in-stream components for a wide range of catchments and for relatively long periods.

3.3 Model validation

Method. Remote sensing or reanalysis products were further employed to validate uncalibrated internal model states or fluxes from three important perspectives in ecohydrological modelling – snow depth from ERA5, evapotranspiration from MODIS, and surface water mass anomaly from GRACE (as a storage proxy). The simulated variables corresponding to these products are, respectively, the depth of snow pack, the sum of soil evaporation, channel evaporation, and transpiration from all soil layers, and the anomaly of total water storage above groundwater (i.e., the sum of canopy storage, snow, soil water storages, and unsaturated storage). The validation was realised via resampling the remote sensing products to 5 km and comparing grid-to-grid with the modelled outputs. r^2 was used as the performance metrics, as KGE is not applicable for time series with zero average, yet the average of surface mass anomaly is close to 0.

Note that all three products may contain considerable uncertainties. ERA5 is a reanalysis product that combines historical observations into global estimates using modelling and data assimilation approaches, therefore inevitably embeds uncertainties associated with model structure and observational coverage (Hersbach et al., 2020). MODIS evapotranspiration is derived from remotely sensed spectral information, energy partitioning approaches and the Penman–Monteith framework, whose uncertainty may exceed 30% depending on spatial scale and environmental conditions (Mu et al., 2011). GRACE infers changes in terrestrially stored water masses from spatial and temporal variations in the Earth's gravity field; however, its coarse spatial resolution can introduce substantial uncertainty when used for hydrological validation, particularly at basin or sub-basin scales (Tapley et al., 2004). Nevertheless, good agreement between simulations and remote sensing or reanalysis products can enhance confidence in the robustness of simulated spatial and temporal patterns, although it does not necessarily imply accurate representation of absolute magnitudes.

Validation performance. First, we compared the sum of soil evaporation, channel evaporation, and transpiration to MODIS evapotranspiration in each grid cell. The results in Figure ~~4~~S2 and Table 3 shows a general good fit between simulation and MODIS records with ~~KGE and r²~~coefficient of determination (R^2) above 0.5 in most ~~regions. From the subplots~~selected catchments. This is further validated after zooming into two representative catchments (W5 Nemunas and W10 Rhine) in Figure ~~4, we can see that~~4A-B, where KGE reaches above 0.5 in >80% of the catchment domain. The other evaluation metrics (RMSE, Pearson coefficient, RMSE, and percent bias shown in Table 3) also demonstrated the ~~seasonality and magnitude~~good performance of evapotranspiration ~~were well captured though the peaks in summer were slightly underestimated~~simulations.

Then, the surface water ~~storage~~mass anomaly from GRACE was compared to the anomaly of simulated surface storage, i.e., the sum of canopy storage, snow, soil water storages, and unsaturated storage. The grid-to-grid comparison in Figure ~~4~~4C-D shows ~~a general good~~an acceptable fit ~~in most regions~~ with ~~r² close to or~~R^2 above 0.~~5.~~3 in around half regions in catchment Nemunas (W5) and Rhine (W10), as well as other European catchments (Figure S3 and Table 3). However, more degradation in performance was found compared to ~~the performance in~~ evapotranspiration~~, especially in coastal regions.~~ simulation. For instance, the validation in Rhine shows degraded water storage simulation in Alpine regions. This is likely due to the locally complex topography and the simplified representation of snow accumulation and melt processes in the model, which struggle to capture the high spatial variability of storage in mountainous terrain. Besides, we also found degraded performance in coastal regions. For example, GRACE exhibited considerably increasing trends in water storage between 2005 to 2015 in two Nordic catchments (W2 and W3; Figure S3), yet our simulations only showed a moderate increasing trend. Similar degraded performance was found in the coastal catchments (e.g., three British catchments W6-8 in UK; Figure S3), though the magnitudes of simulation and GRACE data fit well. This is possibly attributed to the coarse resolution of GRACE which ~~additional~~additionally considered the storage mass from ocean in coastal region yet not included in this terrestrial-explicit modelling.

Finally, the simulated snow depth was compared to the daily snow depth in ERA5 reanalysis products (ERA5 post-processed daily statistics on single levels; 10.24381/cds.4991cf48). Results in Figure ~~4~~4E-F show a generally good agreement between simulations and ERA5 records with $R^2 > 0.5$ in most regions ~~with r2 > 0.5, though~~in Nemunas and Rhine. Though degradation was found in a few other catchments~~. Note that, the poor performances were generally found in~~ (e.g., W14-17 in Figure S4), these catchments ~~with~~generally experienced limited snow accumulation~~, e.g., W14-17 in subplots in Figure 4.~~. In the other words, the absolute deviation was relatively limited for snow depth simulation~~.~~ demonstrated by low RMSE in Table 3

4 Water age simulation and its link to water quality

Like many existing distributed hydrological and water quality models (e.g., SWAT, mHM, EcH$_2$O-iso, HYPE etc.), EcoTWIN can provide estimation of the main ecohydrological fluxes at high spatial and

temporal resolutions, including canopy interception, snow melt-accumulation, infiltration, percolation through soil layers, groundwater recharge, and lateral flux routing at different horizontal phrases. Among these variables, a unique trait of EcoTWIN lies in its capacity to track water fluxes via isotopes, thus being able to provide a consistent estimate of water age and travel times. Therefore, in Figure 5, ~~both~~these variables are shown as the long-term average from 1982 to 2024 for soil profile and stream water in Nemunas (W5) and Rhine (W10).

Generally, the magnitudes of water ages follow the geographic and climatic gradients, with younger water found in catchments with higher annual precipitation inputs. Those regions locate in ~~the~~Alpine regions – for instance around Alps in Southern Rhine, and in catchment W12 and W17 (Figure S6-7). The north-west coast of Europe (Figure ~~5), particularly for Nordic catchments~~S5) also exhibited young water age, where lower temperature and net radiation ~~further~~ limit the level of potential evapotranspiration, leading to larger percolation to deeper soil layers and groundwater. Such high turnover rates of water in these catchments (~~W1,~~ W2, W3, ~~W4, W5~~W6, and ~~W8~~W7) are ~~also demonstrated as the simulated~~reflected by short travel time in soil profile with average values remaining below 500 days~~.~~ (Table 4).

A similar pattern was also found in mountainous regions with higher precipitation and lower potential evapotranspiration compared to lowland areas. Two clear examples are W12 and W17 located in the Alps and the Taurus Mountains where water ages and travel time remained below 500 days (Table 4 and Figure ~~5~~S5-7). In specific wet periods, the soil water ages and travel time can be reduced to just days, suggesting the rapid response of saturated hydrological systems ~~(e.g., the wet year 1999 in Europe in Figure S5-7). .~~ In contrast, the lowlands in central-west Europe showed much slower turnover rates~~, with the~~. A typical example is Rhine (Figure 5), where mean water ages ~~reaching almost~~reached up to 10 years in ~~some specific regions. A few examples could be~~lowlands. Such phenomenon was also found in ~~the three major representative catchments in Central Europe –~~other European catchment (e.g., W9 Elbe, ~~Rhine,~~ and W11 Danube ~~(W9-11). Such old water ages and long travel time are further exacerbated during dry years (e.g., 2004, a drought year for much of Europe shown in Figure S5-7~~).

Note that though water ages and travel time share similar magnitudes and spatial patterns. It is partly attributed to the fact that the travel time in the conceptualised storages increases exponentially in a sequential order. Taking the Rhine as an example, the average travel time in top soil layer, median soil layer, deep soil layer are 65, 225, 1291 days, respectively. Such a depth-dependence profile makes the overall ages/travel time follow the magnitude of bottom layer and leads to similarity between water ages and travel time. However, large discrepancies are possible between the two indices if a shallow lower boundary is adopted.

The estimation of travel time and water ages further provides opportunities to link hydrology and water quality processes in the modelling framework. The simplest and most intuitive way is to compare travel times and simulated biogeochemical process kinetics. Taking denitrification as an example, we applied linear regression and Spearman's correlation test to investigate the potential

correlation between travel time of soil water and denitrification rates. The results in ~~Figure 6~~Table 4 showed the strong positive correlations in most agricultural-dominated catchments (W5, ~~W5,~~ W6~~, W8~~, W9, W10, W11, and W13) yet only weak or no positive correlation in remaining pristine watersheds. This suggests that travel time might be a key control on soil nitrogen removal in European croplands.

More insights can be gained via examination of the Damköhler Number, which quantifies the ratio between timescales of solute transport and biogeochemical transformation. Here in our modelling framework, it can be calculated as the ratio between the travel time of soil water and the time for all soil NO_3-N storage to be removed under the simulated denitrification rates. Damköhler numbers <1 mean that soil water nitrogen cannot be ~~fully~~effectively removed during time of residence, indicating the dominance of transport over removal processes and the potential of nitrogen leaching. As shown in Figure ~~7c~~6E-F, the long-term averages of Damköhler number remain below 1 in most croplands~~,~~ in Nemunas, Rhine, and other European catchments (Figure 6 and Figure S8), supporting the conclusion from the linear regression ~~(~~in Table 4 that travel time is a major limiting factor on soil nitrogen removal~~).~~. Via the spatial- and temporal-explicit estimation of Damköhler number, EcoTWIN provides the opportunity to bridge the catchment hydrological and water quality with travel time.

5 Discussion

5.1 Structural and Functional Merits of EcoTWIN

As a new tracer-aided ecohydrological model, EcoTWIN has novel advantages compared to previous models. In this section, we briefly introduced the merits in model structure, applicability, and insights from tracer-aided simulation.

5.1.1 Integrated C++ framework

Applications of large-scale modelling have been increasingly popular due to the accelerating development of observation networks and availability of remote sensed data. However, it severely increases the computational burden of ecohydrological modelling. Especially for fully distributed models, increasing size of the model domain can lead to exponential increase in computation demands. In this context, an integrated framework in C++ can significantly accelerate the modelling tasks, as all computation can be conducted within memory thereby avoiding the additional input/output overhead associated with disk-based operations in loosely coupled model chains (e.g., EcH$_2$O-iso-nitrate; Yang et al., 2024). A standard test was not performed, but based on our modelling experience in the same catchment with different models, the speed of EcoTWIN (~5 seconds for a simulation with 285 grid cells and 30 years at daily timestep) is close to the water quality model mHM-Nitrate (~5 seconds yet without isotopic simulations; Wu et al., 2022) and easily outperforms EcH$_2$O-iso-nitrate (7 minutes; Wu et al., 2025b).

5.1.2 Selective disassembly structure

EcoTWIN incorporates a wide range of ecohydrological processes from canopy to groundwater, which not only include natural processes but also anthropogenic activities like irrigation. Land managements

549 can also be represented by dynamic parametrisation, thus enabling EcoTWIN to function as a learning
550 tool to investigate the impacts of changes in anthropogenic management over natural ecosystems;
551 for instance, the land use distribution was updated every 10 years in our test examples to reflect the
552 moderate increases in afforestation in the past 45 years in Europe. More importantly, unlike hard
553 coded process representations/equations in most ecohydrological models, EcoTWIN has a selective
554 disassembly structure, which provides alterative conceptualisations for several important hydrological
555 processes (canopy interception, percolation, groundwater recharge, as well as three pedotransfer
556 functions for initialising soil properties). Modellers can benefit from such flexible model structures by
557 either selecting process representations best suited to field knowledge or data prior to calibration, or
558 integrating module selection into the calibration thus enabling simultaneous optimisation of model
559 structure and corresponding parameters. The latter aspect, i.e., the optimisation of model structure,
560 can be realised together with the recently developed optimisation algorithm $DREAM_{(LOAX)}$ that aims to
561 identify the deficits in model structure during calibration (Wu et al., 2025a).

562 ***5.1.3 Transferability to contrasting geographic and climatic contexts***

563 To thoroughly test the applicability of EcoTWIN, 17 catchments with different climatic and
564 geographical contexts were selected for calibration and validation, spanning over most biomes in
565 Europe, from snow-dominated watersheds in Nordic or alpine regions, to agricultural-influenced
566 lowlands catchments, and Mediterranean ecosystems (Figure 2 and Table 1). Through multi-criteria
567 calibration against three objectives at multiple sites, the model successfully reproduced the
568 seasonality and peaks of discharge, in-stream isotopes, and NO_3-N concentrations in most catchments.
569 Such performance is comparable or better than the previous model benchmarks at similar scales
570 (Bajracharya et al., 2023; Mikayilov et al., 2015; Rakovec et al., 2016, 2019). Note that the
571 concentration of NO_3-N was used for calibration, whose accurate simulation is more difficult than NO_3-
572 N loads given the naturally good performance in discharge. In the other words, hydrological simulation
573 is often the least problematic part in integrated water quality modelling, as it is mostly dominated by
574 natural catchment properties while nitrogen cycling is more interfered by anthropogenic
575 managements (e.g., fertilization and irrigation) (Wu et al., 2025b). Additionally, the simulated internal
576 fluxes were also compared to three reanalysis products in hydrological simulations, corresponding to
577 the key fluxes or storage states in hydrological cycling (snow melt-accumulation, evapotranspiration,
578 and water storage). The results show that constrained by isotopes, EcoTWIN was able to reproduce
579 comparable hydrological modelling results to the remote sensing data without direct calibration
580 regarding magnitudes, spatial patterns, and temporal dynamics. The only degraded performance was
581 found in GRACE surface mass anomaly in coastal regions. There are two potential reasons: (i) the
582 coarse resolution of GRACE might account for mass shifts in both ocean and land, yet EcoTWIN only
583 produces mass anomaly in terrestrial systems; (ii) bidirectional fluxes across the land-ocean interface
584 might drive key changes in coastal systems, which is not considered in current version of EcoTWIN.
585 Nonetheless, given the relatively good agreement with most available data, we conclude that
586 EcoTWIN is applicable across a range of terrestrial ecosystems from boreal to temperate and
587 subtropical climate.

5.1.4 Bridging hydrology and water quality with water ages

Further to the inference of hydrological and nitrogen processes that is also available in other distributed water quality models (Wellen et al., 2015), a unique trait of EcoTWIN lies in its capacity to track water fluxes and ages with stable water isotopes. As a tracer-aided model, EcoTWIN not only simulates the celerity of catchment response, but tracks the velocity of water via different flow paths. The importance of delineating flow paths within catchments has long been recognized by hydrologists, and has motivated the development of many indices to describe the movement of water molecule at catchment-scale and estimate associated timescales (Sprenger et al., 2019). A few examples are water ages, transient time distribution, and young water fractions (Benettin et al., 2015; Hrachowitz et al., 2013; Jasechko et al., 2016). However, those indices are mainly calculated in a lumped manner where different flow paths in the catchment are characterised as a black box, thus characterising the overall input-output dynamics yet potentially omitting important spatio-temporal variability of hydrological boundary conditions. Instead, EcoTWIN, benefiting from the gridded-based structure, can utilise the increasingly available spatial information (e.g., gridded remote sensing datasets) thus characterising the water ages and travel time in a spatially-explicit manner. Note that simulations of water age/travel time, like other ecohydrological processes, are sensitive to spatial resolution. The coarse resolution used for large catchments (e.g., 5 km in this study) may obscure the sub-grid heterogeneity. For instance, local hydrological hotspots characterized by short travel times and young water ages can be damped or averaged out at coarser resolutions, as reported in modelling studies using EcH2O-iso (Smith et al., 2021; Yang et al., 2023b). However, this limitation can be mitigated by increasing spatial resolution, and it does not undermine the utility of EcoTWIN for water-tracking.

Compared to water age which quantifies the age of water within the overall system, travel time, accounting for the water age within a specific storage, is more important in understanding the links between hydrological and nutrient cycles. Such an index, also known as transit time or exposure time, forms one of the fundamental components of water quality modelling. Therefore, the travel time estimated by EcoTWIN has potential to improve the simulation of biogeochemical transformations in water quality models interfaced with simplified hydrological modules (e.g., MONERIS; Bonchkovskyi & Osadcha, 2024). Moreover, travel time can be used as a proxy to bridge hydrological processes and biogeochemical transformations. Here we presented a simple framework to calculate the Damköhler Number for denitrification. By using the simulated travel time and reaction timescale (i.e., the time for full removal of nitrogen storage under current denitrification rates), estimation of Damköhler Number was achieved in a spatially- and temporally-explicit manner (Figure ~~7~~6E-F), which can highlight where and when soil nitrogen removal is constrained by the limited exposure time in the catchment. Such high-resolution information is unique, as the use of this index has been largely restricted to steady-state groundwater systems or riparian/hyporheic zones due to the difficulty in quantifying processing time and residence time at larger scales (Ocampo et al., 2006; Wu et al., 2022).

5.2 Limitations and roadmap for future development

Despite these advances, EcoTWIN has limitations. In this section, the uncertainties in model structure and conceptualisation are introduced, as well as the potential roadmaps for future developments.

5.2.1 Potential towards physics-based conceptualisation of groundwater

Groundwater in EcoTWIN is characterised as two conceptual storages linking with adjacent upstream and downstream storages following the topographic gradients. Such conceptualisation, although has been widely employed in hydrological models (e.g., SWAT, mHM, EcH_2O, STARR, etc.), does not align with the physical mechanisms of groundwater routing, as groundwater flow direction follows the hydraulic gradients which may not entirely coincide with topographic gradients (Condon et al., 2021). Such simplified routing has less effect in large catchments with clear topographic gradients (e.g., Rhine starting from Alps to North plain), yet might cause biased estimation in water mass balance for flatter headwater catchments (Yang et al., 2025). Therefore, we plan to further incorporate an additional groundwater module to realise physics-based routing following Darcy's Law in future.

5.2.2 Revisiting mixing strategies

Mixing strategy is a key component in water quality or tracer models describing the flux-storage behaviours along specific flow paths. There has long been a debate on different mixing assumptions and theories. A typical example is the two-water-world hypothesis, where water storage in the soil profile is differentiated into a tightly-bound pool and a mobile-water pool (McDonnell, 2014). Such conceptualisation is close to the definition of soil matrix flow and preferential flow: the existence of free-flowing preferential flow will bypass the soil matrix vertically and accelerate the lateral drainage via direct connection with channel network (Hrachowitz et al., 2013; Sprenger et al., 2019). However, a complete mixing strategy is often regarded as a reasonable first approximation in many situations and is used in most water quality and tracer models (Jung et al., 2025). This is not only attributed to its computational simplicity, but also the difficulty in conceptualising preferential flow in an evidenced-based manner. In the other words, even with the recognition of preferential flow, its calculation is often hindered by the subsurface heterogeneity in soils and bedrock; a good visualisation is given in Figure 7 in Sprenger et al., (2019). Alternatively, partial mixing has been developed for ecohydrological models (e.g., Hrachowitz et al., 2013), which could be added as a complementary mixing strategy in EcoTWIN. However, as benchmarked in Hrachowitz et al., (2013), the partial mixing brings only moderate improvements in simulations yet can introduce challenges to model spin-up (the increasing instability of storage ages due to the exchange between bypass and storage compartment). Moreover, the realisation of partial mixing, like preferential flow, relies on additional parameters to describe the timing and extent of mixing thus introducing additional parametric uncertainty. Therefore, we recommend a rigorous evaluation of the necessity of partial mixing before any application.

5.2.3 Complementing the in-stream biogeochemical processes

Transformation is as crucial as transport in inland-water nitrogen cycling (Wang et al., 2024). In the current version of EcoTWIN, denitrification is the only in-stream process of nitrogen loss. However, recent studies have shown that other processes are involved which may be important for aquatic nitrogen cycling. An example originates from Wang et al., (2024), where global inland-water modelling shows that in-stream denitrification only accounts for a minor fraction of NO_3-N removal compared to biological uptake. Though their modelling considers lakes and reservoirs where primary production of benthic plants and algae is usually greater than that in rivers, in-stream assimilation might still play a significant role, particularly, in slow-flowing river systems. This is supported by a recent modelling study that estimated nitrogen retention at 15-min interval based on high-frequency NO_3-N data (Yang et al., 2023a). Therefore, we plan to further compliment EcoTWIN with in-stream assimilation conceptualisation, as well as other potentially important riverine processes (e.g., nitrogen burial in sediments; Akbarzadeh et al., 2019).

5.2.4 Integrated calibration framework to embrace equifinality

Strictly speaking, equifinality is not specifically linked to EcoTWIN, but remains a universal problem for calibration or parameter tunning for almost all ecohydrological models. It is reflected in multiple parameters sets yielding similarly good model performance, thus increasing the uncertainty in process inference. The extent of equifinality is primarily controlled by the magnitude of parameters and observation/objectives (Wu et al., 2025c). Unfortunately, conceptualisations across diverse process domains (e.g. for hydrology, isotopes and N-cycling) in EcoTWIN also lead to a relatively large number of parameters. Such risk in equifinality can be potentially constrained via sensitivity analysis, but can still remain an issue given the ubiquitous epistemic uncertainty in data and model structure (Beven, 2006, 2015). Alternatively, the recently developed calibration algorithm $DREAM_{(LoAX)}$ provides an opportunity to embrace equifinality by tunning parameters based on the limits-of-acceptability theory under the equifinality thesis (Wu et al., 2025a). The integrated modelling framework of EcoTWIN and $DREAM_{(LoAX)}$ can potentially increase the robustness of model calibration and inference.

6 Conclusions

Uncertainty is a central concern in ecohydrological modelling, as models are not only used for prediction of specific variables, but also for process inference (backtracking internal processes from available observations) that are inherently embedded within considerable uncertainty. Stable water isotopes can help effectively constrain hydrological fluxes due to their conservative nature, motivating the increased development of tracer-aided models. However, few attempts have been made to incorporate a tracer-aided hydrological framework into water quality models.

Therefore, we introduced EcoTWIN, a fully distributed tracer-aided ***ecohydrological*** model that ***tracks water***, ***isotopic***, and ***nutrient*** fluxes simultaneously in an integrated C++-based framework. To thoroughly validate the model, 17 large European catchments were selected with a wide range of

geographic and climatic gradients (from snow-dominated watersheds in Nordic or alpine regions, to agricultural-influenced lowlands catchments, and Mediterranean ecosystems). The model was calibrated against long-term observations of discharge, in-stream isotopes, and NO_3-N concentrations during 1980-2024 in each of the 17 catchments. Additionally, uncalibrated internal states and fluxes were also compared with three remote sensing products (ERA5 snow depth, MODIS evapotranspiration, and GRACE surface water anomaly) to validate the credibility of process inference.

The generally good agreements in both calibrated in-stream components and uncalibrated internal flux-states demonstrated that EcoTWIN is a transferable, flexible prediction and learning tool for process inference across biomes ranging from boreal to subtropical climate. Constrained by tracer simulations, the model not only reproduces the celerity of hydrological systems, but also tracks the velocity. Water ages and travel time are embedded in EcoTWIN to provides spatio-temporal-explicit insights into *when*, *where*, and *how* water moves in the system. Such indices further provide the opportunities to efficiently bridge hydrology and water quality at large catchment-scales. An example was presented using the Damköhler Number to identify regions where denitrification was limited by fast turnover rates of water.

Following this "proof of concept" we also see numerous areas where future developments can improve the limitations in the 1.0 version of the model.

Code and data availability

The initial version (v1.0) of EcoTWIN is archived in https://doi.org/10.5281/zenodo.16747633 (Wu et al., 2025d). For further development please refer to GitHub repository: https://github.com/songjun-wu/EcoTWIN. The geographic data were acquired from Catchment Characterisation and Modelling database (CCM2, version 2.1). The climatic forcing was acquired from E-OBS database (https://www.ecad.eu/download/ensembles/ensembles.php). The LAI were acquired from MODIS database (http://doi.org/10.5067/MODIS/MOD15A2H.006). Long-term observation of discharge was acquired from GRDC (https://grdc.bafg.de/); in-stream isotopic observations were available from Wateriso database (https://wateriso.utah.edu/waterisotopes/index.html) and GNIR database (https://www.iaea.org/services/networks/gnir); In-stream NO_3-N concentration were acquired from global water quality database, GEMStat (https://gemstat.org/).

Acknowledgements

Tetzlaff's and Wu's contributions were partly funded through the WETSCAPES2.0 project (DFG TRR410/1 2025) and the Einstein Research Unit "Climate and Water under Change" from the Einstein Foundation Berlin and Berlin University Alliance (grant no. ERU-2020-609). Tetzlaff was also partly funded through Leibniz Excellence project ISOSCALE and received funding from the "Wasserressourcenpreis 2024" of the Rüdiger Kurt Bode-Foundation. This research was also supported by the BMBF (funding code 033W034A), which supported the IGB stable isotope Laboratory.

Contributions from Soulsby were supported by Leibnitz Association Germany in the project Wetland Restoration in Peatlands.

Author contribution

Conceptualization: SW, DT, YZ, CS

Data curation: SW

Methodology: SW

Software: SW

Investigation: SW, DT, YZ, CS

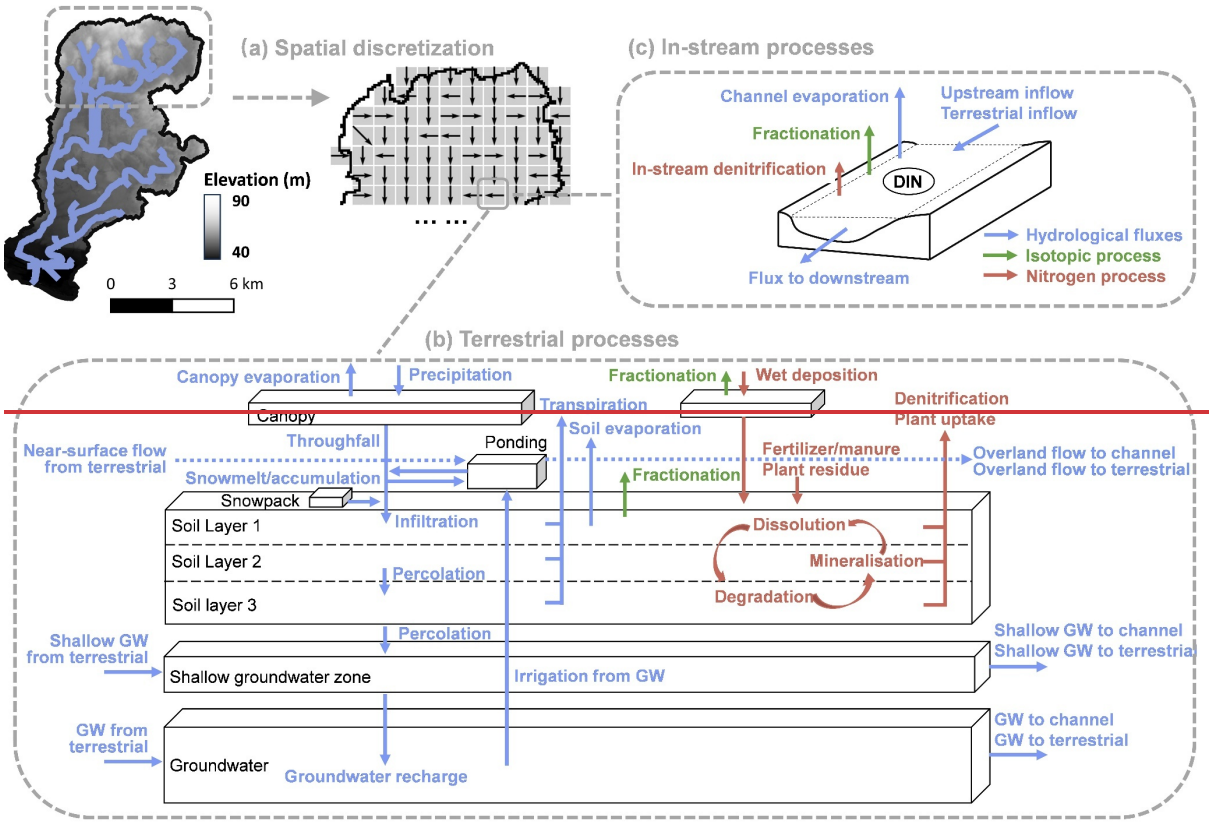
Visualization: SW

Supervision: DT, CS

Writing (original draft preparation): SW

Writing (review and editing): SW, DT, YZ, CS

Figures



~~Figure 1. Model structure of EcoTWIN. As a distributed model, EcoTWIN disentangles the spatial domain into grid cells (Panel a). In each grid cell, hydrological, isotopic, and nitrogen processes were simulated in canopy, snow, soils, shallow groundwater, and groundwater (panel b) and river channel if channels are present (Panel c).~~

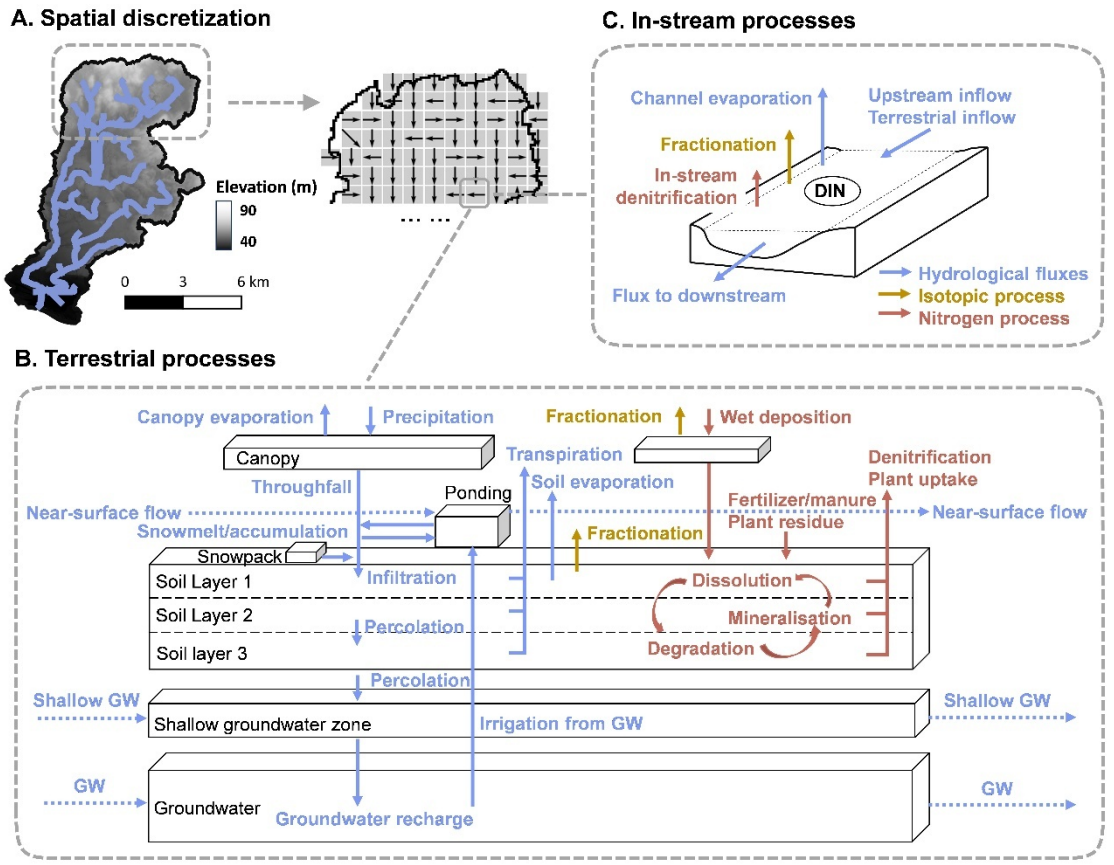


Figure 1. Model structure of EcoTWIN. As a distributed model, EcoTWIN disentangles the spatial domain into grid cells (Panel A). In each grid cell, hydrological, isotopic, and nitrogen processes were simulated in canopy, snow, soils, shallow groundwater, and groundwater (Panel B) and river channel if channels are present (Panel C).

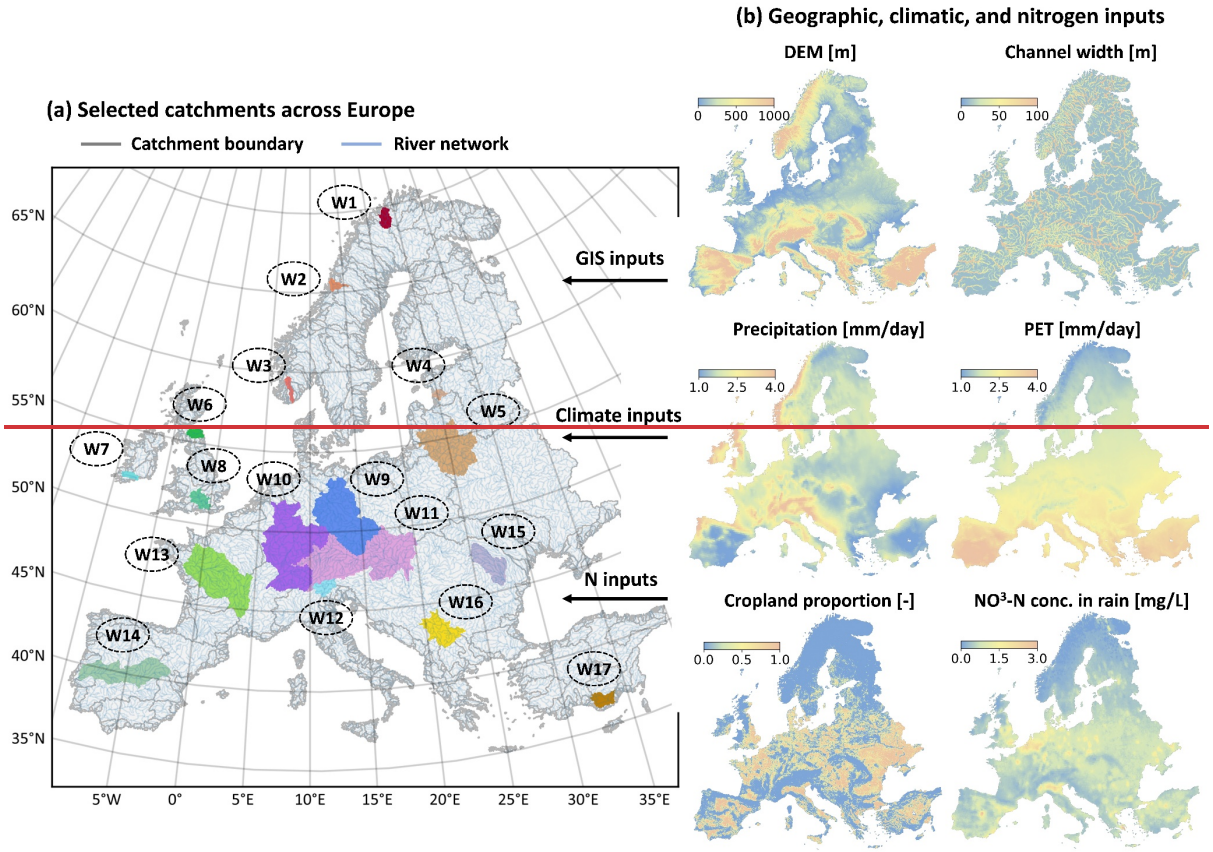


Figure 2. The selected catchments for model validation (Panel a) and an overview of key geographic, climatic, and nitrogen inputs (Panel b).

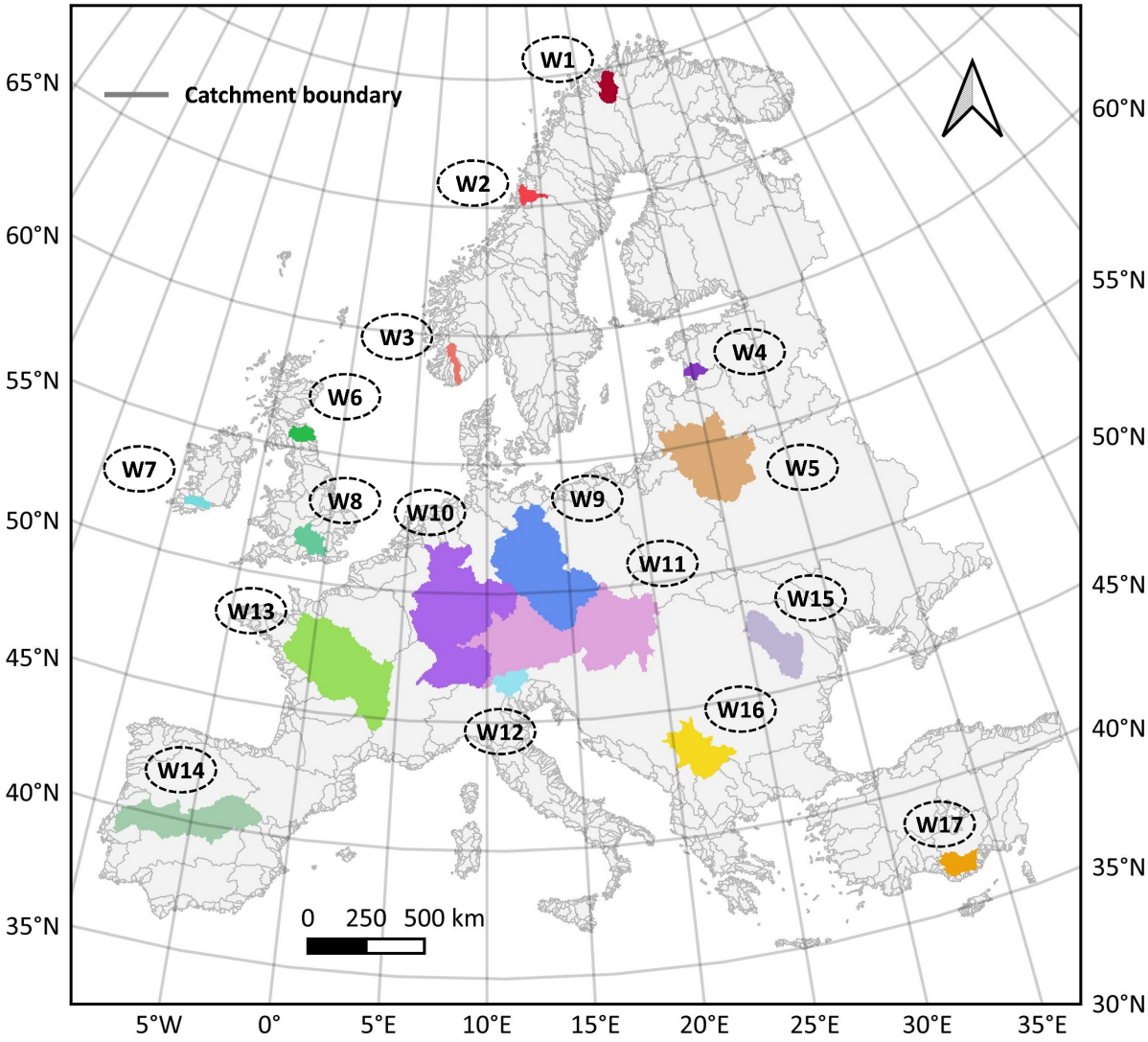
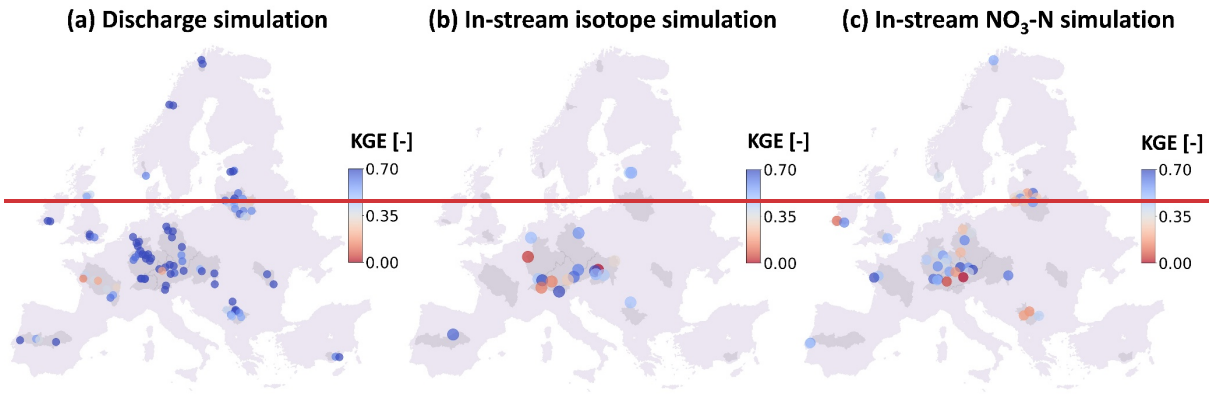


Figure 2. The selected catchments for model calibration and validation. The overview of key geographic, climatic, and nitrogen inputs are show in Figure S1.



~~Figure 3. The simulation performance of discharge, in-stream isotope, and in-stream NO_3-N. The simulated and observed time series are shown in Figure S1.~~

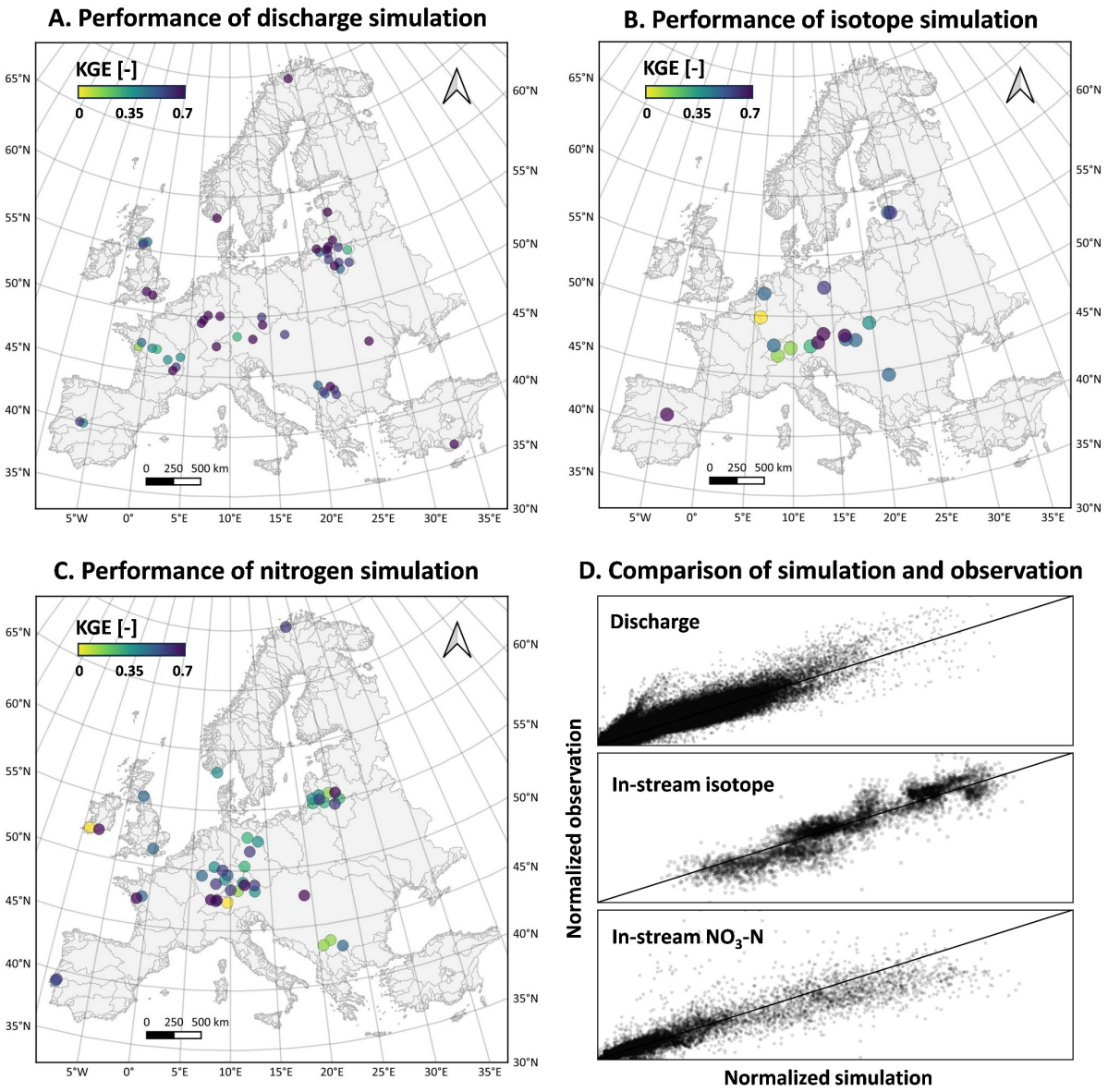
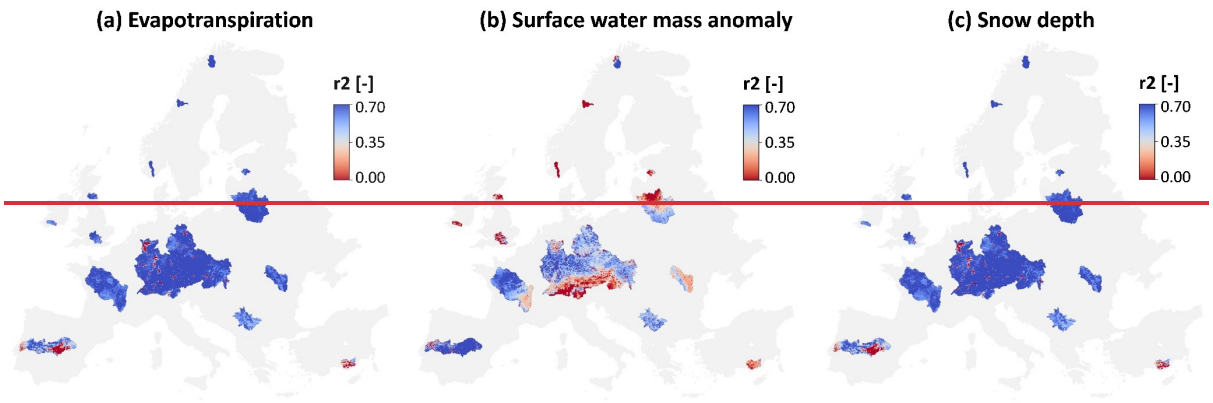


Figure 3. The simulation performance of discharge (A), in-stream isotope (B), and in-stream NO_3-N (C) evaluated using KGE. The comparison between simulation and observation values are shown in Panel D. Catchment-specific performances are summarised in KGE, RMSE, Pearson correlation coefficient, and Pbias in Table 2.



~~Figure 4. The grid-to-grid comparison between time series of simulated internal states/fluxes and the ones extracted from remote sensing/reanalysis products, including evapotranspiration from MODIS (Panel a), surface water mass anomaly from GRACE (Panel b), and snow depth from ERA5 (Panel c). The time series of simulation and remote sensing/reanalysis products in each catchment are shown Figure S2-4.~~

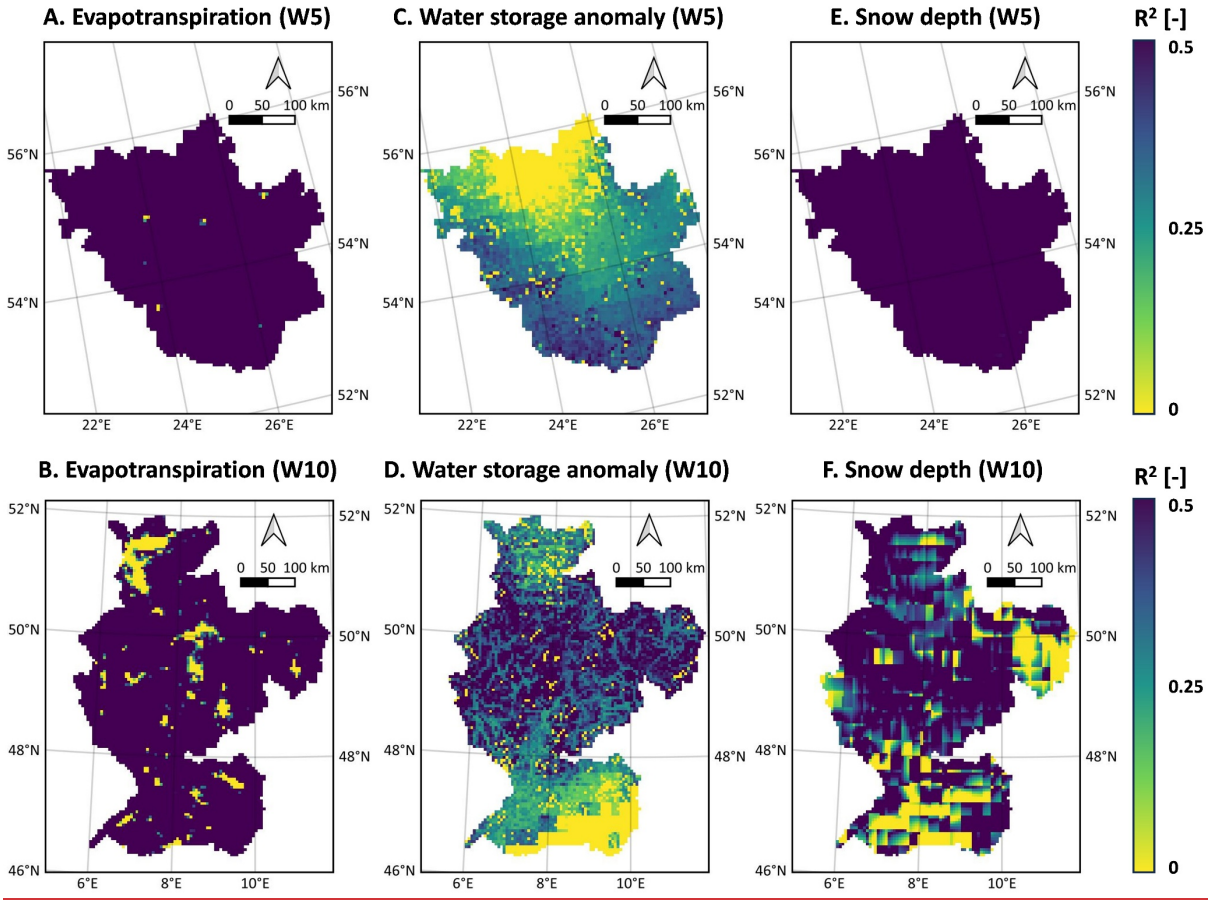


Figure 4. The grid-to-grid comparison (in coefficient of determination R^2) between time series of simulated internal states/fluxes and the corresponding records extracted from remote sensing/reanalysis products, including evapotranspiration from MODIS (Panel A-B), surface water mass anomaly from GRACE (Panel C-D), and snow depth from ERA5 (Panel E-F). The validation of two representative catchments is shown here (W5 Nemunas and W10 Rhine), while the spatial maps of all catchments were shown in Figure S2-4. The evaluation metrics in each catchment were summarised in Table 3.

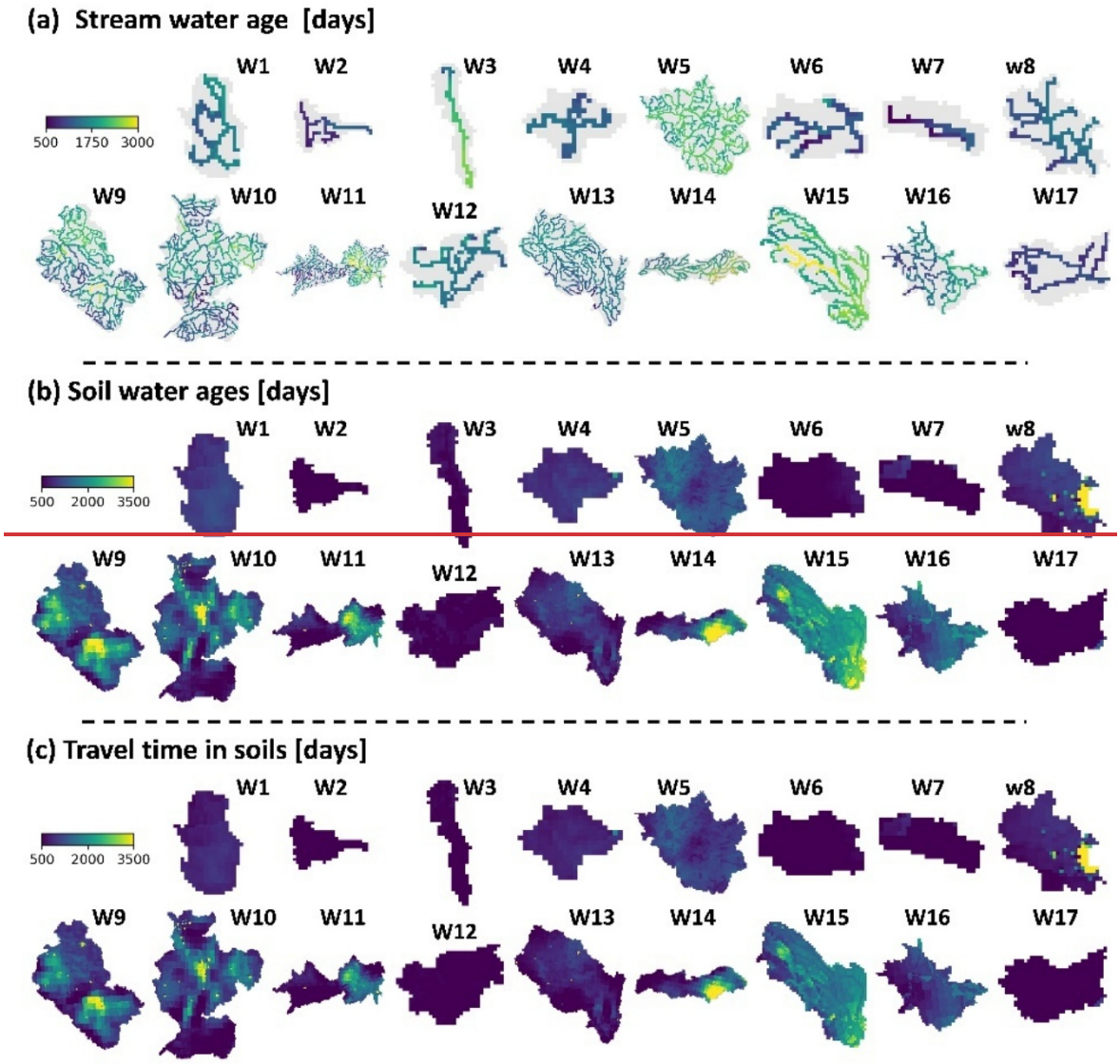


Figure 5. The simulated long-term average (1982-2024) of water age and travel time in channel and soil profile. Water ages represent the time since water enters the catchments as precipitation, while travel times depict the residence time of water within the specific storage.

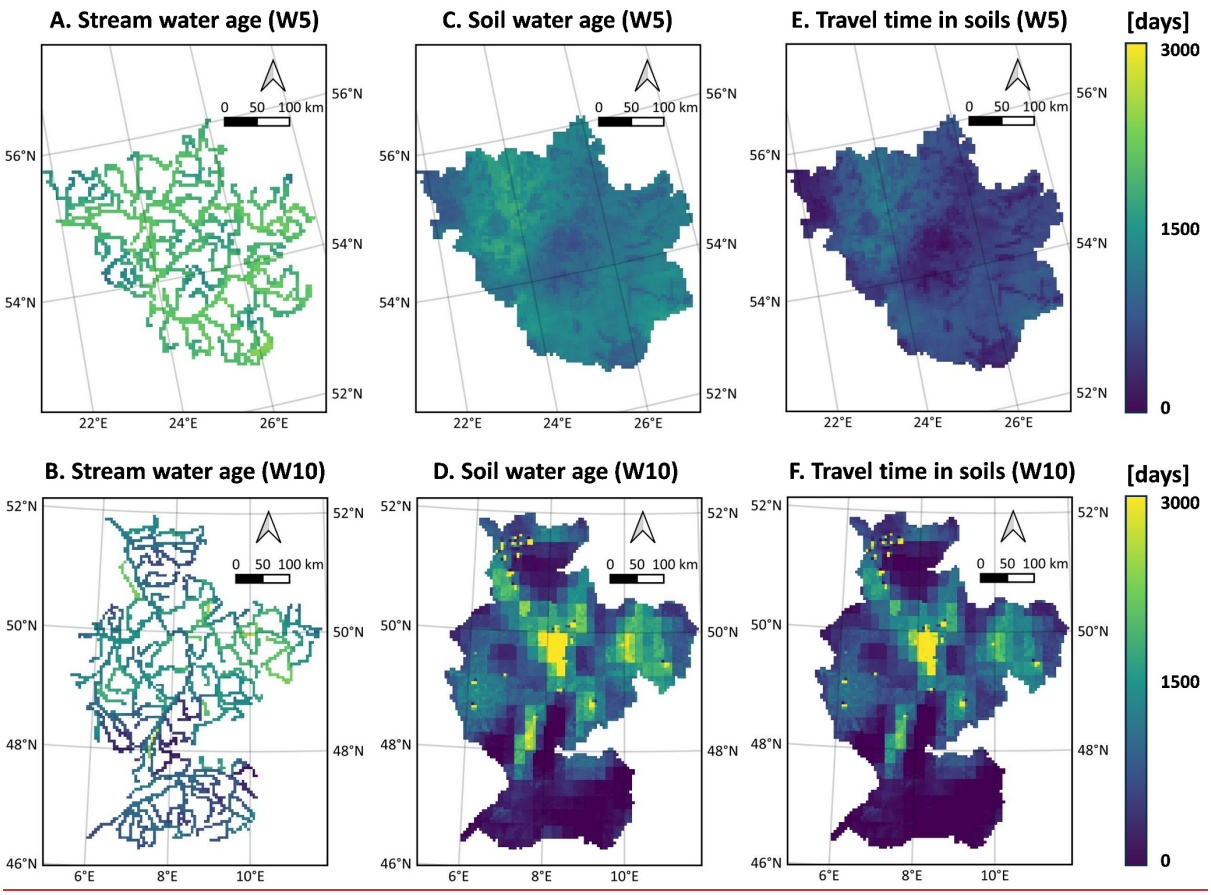


Figure 5. The simulated long-term average (1982-2024) of stream water age (Panel A-B), soil water age (Panel C-D), and travel time of soil water (Panel E-F). Water age represents the time since water enters the catchments as precipitation, while travel time depicts the residence time of water within the specific storage (i.e., the soil profile in this figure).

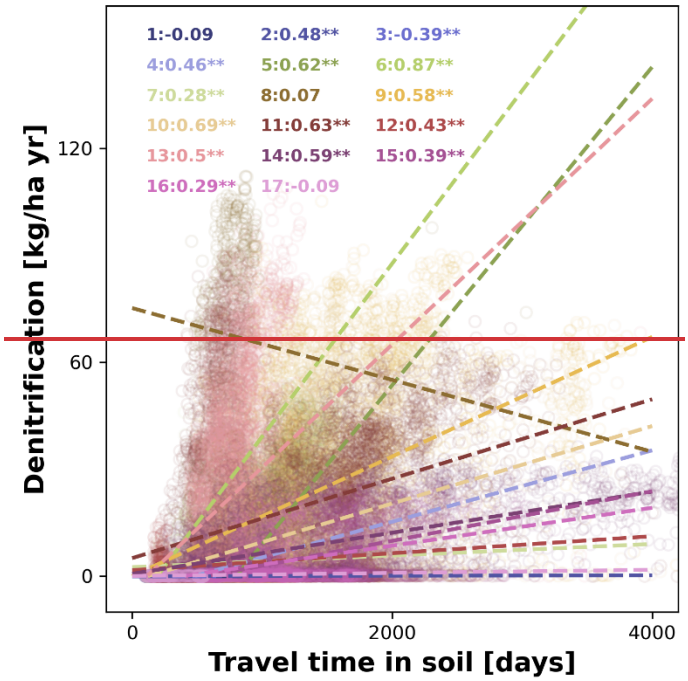


Figure 6. The correlations between travel time and annual denitrification. The text depicts the spearman correlation coefficients and p values (* = less than 0.05, ** = less than 0.01) in each catchment.

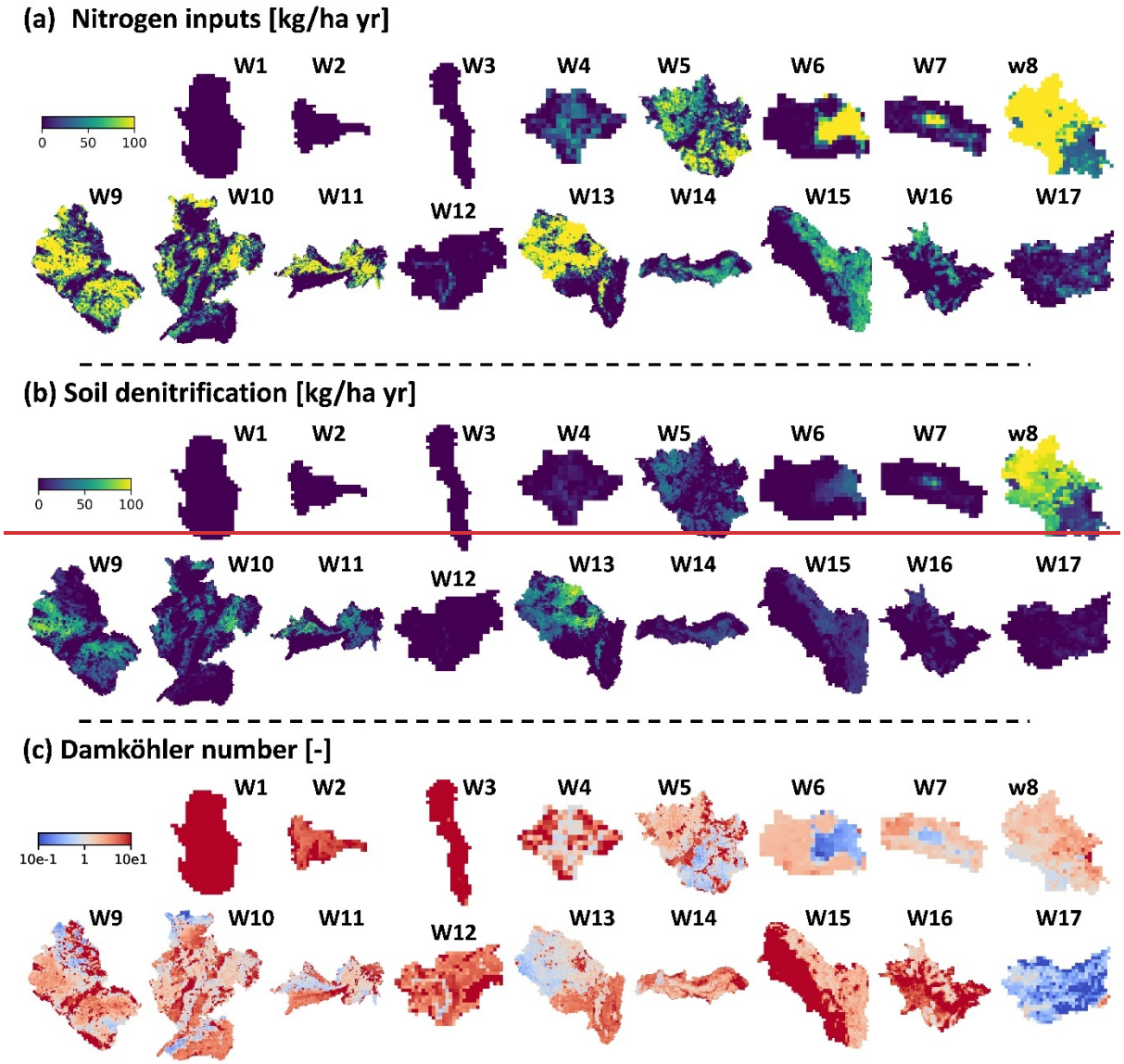
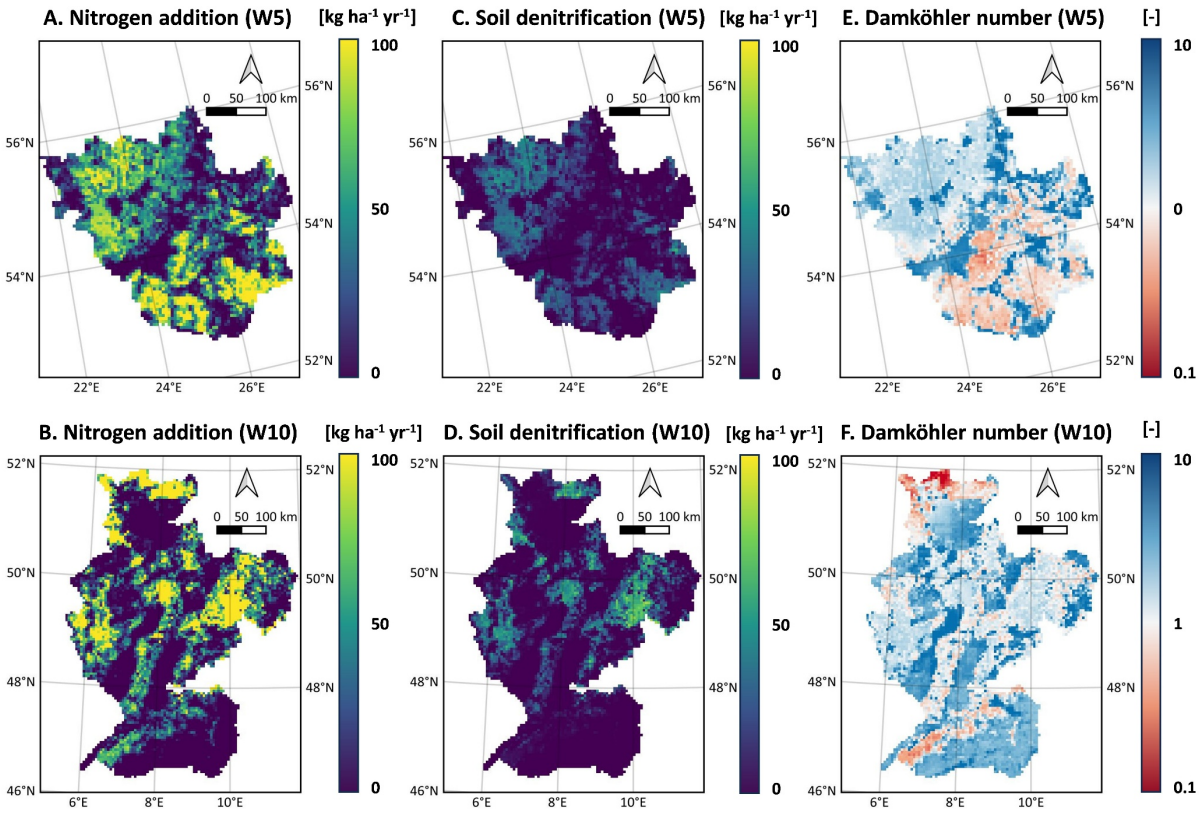


Figure 7. The simulated long-term average (1982-2024) of nitrogen inputs, soil denitrification, and Damköhler number.



814

815 Figure 6. The simulated long-term average (1982-2024) of nitrogen addition from fertilization (Panel
816 A-B), soil denitrification (Panel C-D), and Damköhler number (Panel E-F). Note that logarithmic scale
817 is used for Damköhler number.

818

Tables

Table 1. Characteristics of the selected catchments. **Lat** depicts the latitude of upper left corner of the catchment. **DEM** and **Area** are the mean elevation in m.a.s.l. and catchment size in km^2. **Precip**, **Temp**, and **PET** are the annual averages of precipitation, air temperature, and potential evapotranspiration in mm/yr. f_{crop}, f_{forest}, and f_{urban} are the fractions of cropland, forest, and urbanized areas in 2019 in %. Null means no name is assigned for the catchment in the Catchment Characterisation and Modelling (CCM) database.

ID	Name	Lat	Area	DEM	Precip	Temp	PET	f_{crop}	f_{forest}	f_{urban}
1	Null	70.0	8725	468.5	448.5	-1.8	442.8	<1	1.9	<1
2	Vefsna	65.9	5475	636.5	1260.8	0.7	433.6	<1	24.6	<1
3	Null	59.8	5225	742.3	1400.9	3.0	545.7	<1	39.8	<1
4	Null	58.3	4350	67.4	654.5	6.4	667.9	17.4	60.5	1.4
5	Nemunas	56.6	97550	147.9	599.3	7.1	730.6	33.0	39.1	4.5
6	Tweed	55.9	6250	264.3	1023.4	7.9	600.9	21.8	18.9	1.4
7	Null	52.3	4300	175.4	1218.1	10.1	645.2	12.1	17.4	1.9
8	Thames	52.2	11900	112.0	700.7	10.4	782.9	44.9	14.1	22.0
9	Elbe	53.5	130225	318.3	626.9	8.8	836.3	41.2	34.6	10.8
10	Rhine	52.0	170175	508.3	943.3	8.9	821.0	21.5	41.2	17.1
11	Danube$_{(a)}$	50.5	197600	618.0	843.6	8.3	857.4	28.5	37.1	11.5
12	Adige	47.2	11600	1771.9	1002.3	4.5	809.7	<1	48.8	3.2
13	Loire	48.7	122125	298.9	778.7	11.0	887.4	37.6	25.7	6.5
14	Tajo	40.4	75575	686.2	549.5	14.3	1359.6	26.7	34.5	3.3
15	Danube$_{(b)}$	48.4	37975	533.3	534.7	8.1	869.5	32.0	41.3	6.4
16	Danube$_{(c)}$	44.8	37725	653.3	684.7	9.7	994.2	12.8	44.3	5.8
17	Null	37.6	12650	1384.5	454.1	12.2	1256.9	5.9	4.3	<1

~~Table 2. The calibration performance of discharge (Q), in-stream isotopes (^{18}O, Iso), and nitrate (NO^3-N). Evaluation metrics include Kling-Gupta efficiency ( ), Root Mean Square Error (m^3/s, ‰, and mg/L for discharge, isotopes, and nitrate, respectively), Pearson Correlation Coefficient ( ), and Percent bias (%).~~

~~Metric~~	~~Unit~~	~~Min~~	~~Max~~	~~Mean~~	~~Median~~
~~Kling-Gupta efficiency (Q)~~	~~-~~	~~0.14~~	~~0.89~~	~~0.65~~	~~0.69~~
~~Kling-Gupta efficiency (Iso)~~	~~-~~	~~-0.03~~	~~0.86~~	~~0.45~~	~~0.48~~
~~Kling-Gupta efficiency (NO^3-N)~~	~~-~~	~~-0.36~~	~~0.72~~	~~0.42~~	~~0.44~~
~~Pearson correlation coefficient (Q)~~	~~-~~	~~0.49~~	~~0.92~~	~~0.79~~	~~0.81~~
~~Pearson correlation coefficient (Iso)~~	~~-~~	~~0.14~~	~~0.87~~	~~0.51~~	~~0.54~~
~~Pearson correlation coefficient (NO^3-N)~~	~~-~~	~~-0.26~~	~~0.86~~	~~0.55~~	~~0.6~~
~~Root Mean Square Error (Q)~~	~~m^3/s~~	~~3.99~~	~~677.08~~	~~123.02~~	~~68.51~~
~~Root Mean Square Error (Iso)~~	~~‰~~	~~0.31~~	~~1.51~~	~~0.72~~	~~0.73~~
~~Root Mean Square Error (NO^3-N)~~	~~mg/L~~	~~0.02~~	~~2.82~~	~~0.83~~	~~0.57~~
~~Percent bias (Q)~~	~~%~~	~~0.52~~	~~79.88~~	~~17.44~~	~~9.53~~
~~Percent bias (Iso)~~	~~%~~	~~-11.28~~	~~-0.07~~	~~-4.3~~	~~-4.42~~
~~Percent bias (NO^3-N)~~	~~%~~	~~0.18~~	~~49.25~~	~~15.52~~	~~10.89~~

Table 2. Calibration statistics for discharge (*q*), in-stream isotope (*iso*), and in-stream NO_3–N (*no3*) across the 17 selected catchments. Evaluation metrics represent the mean values computed within each catchment.

ID	Number of sites			Kling-Gupta efficiency			Pearson Correlation			Root Mean Square Error			Percent bias		
	q -	*iso* -	*no3* -	*q* -	*iso* -	*no3* -	*q* -	*iso* -	*no3* -	*q* m^3/s	*iso* ‰	*no3* mg/L	*q* %	*iso* %	*no3* %
W1	2	0	1	0.7		0.5	0.9		0.6	56		0.0	16.4		0.2
W2	2	0	0	0.9			0.9			47			3.3		
W3	1	0	1	0.6		0.4	0.7		0.4	62		0.0	22.5		2.3
W4	3	2	0	0.7	0.5		0.7	0.5		12	0.8		9.5	2.5	
W5	13	0	9	0.6		0.4	0.6		0.6	64		1.4	15.8		22.2
W6	3	0	1	0.4		0.4	0.9		0.4	42		0.4	48.3		4.3
W7	2	0	2	0.8		0.4	0.9		0.5	22		0.5	3.6		18.8
W8	3	0	1	0.7		0.5	0.9		0.5	18		1.3	17.0		6.6
W9	7	1	4	0.7	0.6	0.4	0.8	0.7	0.5	114	0.7	0.9	8.1	6.3	21.5
W10	18	6	10	0.8	0.3	0.5	0.9	0.4	0.6	232	0.8	0.8	8.0	5.4	13.7
W11	12	8	9	0.7	0.4	0.4	0.8	0.5	0.5	231	0.6	0.6	15.6	3.1	19.3
W12	2	1	0	0.8	0.9		0.8	0.9		43	0.3		2.5	0.9	
W13	8	0	2	0.4		0.6	0.8		0.6	120		1.1	52.9		14.1
W14	4	1	2	0.6	0.7	0.5	0.7	0.7	0.6	107	0.8	0.3	23.7	5.5	5.9
W15	2	0	0	0.7			0.8			113			5.8		
W16	9	1	3	0.6	0.5	0.3	0.7	0.5	0.4	59	1.5	0.5	20.4	-	3.7
W17	2	0	0	0.7			0.7			56			8.4		

Table 3. Validation performance of evapotranspiration (against MODIS records), water storage anomaly (against GRACE records), and snow depth (against ERA5 records). Performance is evaluated using the mean Kling–Gupta efficiency (KGE), Pearson correlation coefficient (r), root mean square error (RMSE), and percent bias (Pbias) across all grid cells within each of the 17 catchments.

ID	Evapotranspiration				Water storage anomaly				Snow depth			
	KGE	*r*	*RMSE*	*Pbias*	*KGE*	*r*	*RMSE*	*Pbias*	*KGE*	*r*	*RMSE*	*Pbias*
	-	-	*m*	-	-	-	*m*	-	-	-	*m*	-
W1	0.88	0.95	0.00	0.3	-	0.79	0.03	8.7	0.78	0.90	0.04	57.5
W2	0.79	0.96	0.00	0.2	-	0.77	0.14	37.5	0.19	0.81	0.12	162.5
W3	0.80	0.94	0.00	0.3	-	0.45	0.08	18.5	0.65	0.82	0.06	1159.7
W4	0.67	0.88	0.00	0.3	-	0.51	0.05	13.1	0.52	0.85	0.01	14.6
W5	0.70	0.88	0.00	0.3	-	0.49	0.05	258.6	0.81	0.90	0.01	99.9
W6	0.65	0.90	0.00	0.3	-	0.46	0.03	18.4	0.40	0.82	0.00	1.9
W7	0.60	0.91	0.00	0.2	-	0.46	0.03	10.0	0.07	0.67	0.00	46.3
W8	0.58	0.89	0.00	0.4	-	0.65	0.03	8.3	0.28	0.81	0.00	10.7
W9	0.08	0.86	0.00	0.9	-	0.59	0.04	4.0	0.65	0.82	0.01	28.0
W10	0.44	0.90	0.00	0.6	-	0.62	0.05	4.7	0.25	0.78	0.03	8021.9
W11	0.39	0.90	0.00	0.6	-	0.53	0.05	4.3	0.36	0.80	0.02	99.4
W12	0.65	0.92	0.00	0.3	-	0.57	0.07	14.0	0.54	0.73	0.07	125.8
W13	0.67	0.90	0.00	0.3	-	0.64	0.04	3.2	0.62	0.78	0.00	434.3
W14	0.35	0.60	0.00	0.7	-	0.74	0.03	5.2	0.06	0.66	0.00	158.1
W15	0.67	0.90	0.00	0.3	-	0.50	0.06	2.5	0.16	0.79	0.01	536.2
W16	0.59	0.82	0.00	0.3	-	0.65	0.05	2.2	0.21	0.81	0.01	18.6
W17	0.18	0.78	0.00	0.4	-	0.43	0.06	4.3	-0.05	0.78	0.02	62.4

Table 4. Mean soil travel time and soil denitrification in selected catchments during 1982–2024. The regression slope of the fitted relationship between soil travel time and denitrification, along with the P-values of the Spearman test, are also reported.

ID	Avg. Soil travel time [days]	Soil denitrification [kg ha^{-1}]	Slope [-]	P-value [-]
W1	808.56	0.01	-0.01	<0.05
W2	206.20	0.01	0.07	<0.05
W3	434.02	0.01	0.00	<0.05
W4	804.22	3.33	10.01	<0.05
W5	1067.89	11.98	44.63	<0.05
W6	296.05	5.02	48.53	<0.05
W7	388.63	3.19	1.62	<0.05
W8	1063.87	64.50	-10.08	0.12
W9	1279.90	21.51	16.77	<0.05
W10	1099.18	10.64	10.86	<0.05
W11	948.22	15.64	11.14	<0.05
W12	320.98	0.21	0.43	0.06
W13	719.82	20.88	34.47	<0.05
W14	1336.50	8.40	5.76	<0.05
W15	1681.02	7.72	6.96	<0.05
W16	1039.28	3.46	5.32	<0.05
W17	419.38	2.22	2.38	<0.05

Reference

Akbarzadeh, Z., Maavara, T., Slowinski, S., & Van Cappellen, P. (2019). Effects of Damming on River Nitrogen Fluxes: A Global Analysis. *Global Biogeochemical Cycles*, *33*. https://doi.org/10.1029/2019GB006222

Ala-aho, P., Tetzlaff, D., Mcnamara, J., Laudon, H., & Kormos, P. (2017). Modeling the isotopic evolution of snowpack and snowmelt: Testing a spatially distributed parsimonious approach. *Water Resources Research*, *53*. https://doi.org/10.1002/2017WR020650

Amato, M. T., & Giménez, D. (2024). Predicting monthly near-surface soil temperature from air temperature and the leaf area index. *Agricultural and Forest Meteorology*, *345*, 109838. https://doi.org/10.1016/j.agrformet.2023.109838

Arnold, J., Moriasi, D., Gassman, P., Abbaspour, K., White, M., Srinivasan, R., et al. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)*, *55*. https://doi.org/10.13031/2013.42256

Bajracharya, A., Moghairib, M., Stadnyk, T., & Asadzadeh, M. (2023). Process based calibration of a continental-scale hydrological model using soil moisture and streamflow data. *Journal of Hydrology: Regional Studies*, *47*, 101391. https://doi.org/10.1016/j.ejrh.2023.101391

Benettin, P., Bailey, S., Campbell, J., Green, M., Rinaldo, A., Likens, G., et al. (2015). Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH, USA. *Water Resources Research*, *51*, 9256–9272. https://doi.org/10.1002/2015WR017552

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1–2), 18–36. https://doi.org/10.1016/J.JHYDROL.2005.07.007

Beven, K. (2015). Facets of uncertainty: Epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, *61*, 150407155940006. https://doi.org/10.1080/02626667.2015.1031761

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., Tetzlaff, D., & Dunn, S. (2011). Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models. *Water Resources Research - WATER RESOUR RES*, *47*. https://doi.org/10.1029/2010WR009547

Bonchkovskyi, A., & Osadcha, N. (2024). Modelling of the nutrient load in the Sula River basin using the MONERIS. *Physical Geography and Geomorphology*, *47*, 7–20. https://doi.org/10.17721/phgg.2024.3-4.01

Condon, L., Kollet, S., Bierkens, M., Maxwell, R., Hill, M., Franssen, H.-J., et al. (2021). Global Groundwater Modeling and Monitoring: Opportunities and Challenges. *Water Resources Research*, *57*. https://doi.org/10.1029/2020WR029500

Craig, H., Gordon, L. I., & Horibe, Y. (1964). Isotopic exchange effects in the evaporation of water: 1. Low-temperature experimental results. *Journal of Geophysical Research*, *68*(17), 5079–5087. https://doi.org/10.1029/JZ068I017P05079

Dawson, T., & Ehleringer, J. (1991). Streamside trees that do not use stream water. *Nature*, *350*, 335–337. https://doi.org/10.1038/350335a0

Dick, J. J., Tetzlaff, D., Birkel, C., & Soulsby, C. (2015). Modelling landscape controls on dissolved organic carbon sources and fluxes to streams. *Biogeochemistry*, *122*(2–3), 361–374. https://doi.org/10.1007/S10533-014-0046-3

Donnelly, C., Andersson, J. C. M., & Arheimer, B. (2016). Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe. *Hydrological Sciences Journal*, *61*(2), 255–273. https://doi.org/10.1080/02626667.2015.1027710

Eckersten, H., Jansson, P. E., & Johnsson, H. (1994). SOILN model – user's manual 2nd edition, Division of Agricultural Hydrotechnics Communications, 94:4. Department of Soil Sciences, Swedish University of AgriculturalSciences, 58pp, Uppsala.

Efstratiadis, A., & Koutsoyiannis, D. (2010). One decade of multi-objective calibration approaches in hydrological modelling: A review. *Hydrological Sciences Journal*, *55*(1), 58–78. https://doi.org/10.1080/02626660903526292

Godsey, S. E., Aas, W., Clair, T. A., de Wit, H. A., Fernandez, I. J., Kahl, J. S., et al. (2010). Generality of fractal 1/f scaling in catchment tracer time series, and its implications for catchment travel time distributions. *Hydrological Processes*, *24*(12), 1660–1671. https://doi.org/10.1002/hyp.7677

Good, S. P., Soderberg, K., Guan, K., King, E. G., Scanlon, T. M., & Caylor, K. K. (2014). δ2H isotopic flux partitioning of evapotranspiration over a grass field following a water pulse and subsequent dry down. *Water Resources Research*, *50*(2), 1410–1432. https://doi.org/10.1002/2013WR014333

Grizzetti, B., Vigiak, O., Udias, A., Aloe, A., Zanni, M., Bouraoui, F., et al. (2021). How EU policies could reduce nutrient pollution in European inland and coastal waters. *Global Environmental Change*, *69*, 102281. https://doi.org/10.1016/j.gloenvcha.2021.102281

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*. https://doi.org/10.1002/qj.3803

Hooper, R., Stone, A., Christophersen, N., de Melo, E., & Seip, H. (1988). Assessing the Birkenes Model of Stream Acidification using a multi-signal calibration methodology. *Water Resources Research - WATER RESOUR RES*, *24*, 1308–1316. https://doi.org/10.1029/WR024i008p01308

Horita, J., Rozanski, K., & Cohen, S. (2008). Isotope effects in the evaporation of water: A status report of the Craig-Gordon model. *Isotopes in Environmental and Health Studies*, *44*, 23–49. https://doi.org/10.1080/10256010801887174

Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., & Soulsby, C. (2013). What can flux tracking teach us about water age distribution patterns and their temporal dynamics? *Hydrology and Earth System Sciences*, *17*(2), 533–564. https://doi.org/10.5194/HESS-17-533-2013

Huijgevoort, M., Tetzlaff, D., & Sutanudjaja, E. (2016). Using high resolution tracer data to constrain water storage, flux and age estimates in a spatially distributed rainfall-runoff model. *Hydrological Processes*, *30*. https://doi.org/10.1002/hyp.10902

Jasechko, S., Kirchner, J., Welker, J., & McDonnell, J. (2016). Substantial proportion of global streamflow less than three months old. *Nature Geoscience*, *9*. https://doi.org/10.1038/NGEO2636

Jones, E. R., Bierkens, M. F. P., Wanders, N., Sutanudjaja, E. H., van Beek, L. P. H., & van Vliet, M. T. H. (2023). DynQual v1.0: a high-resolution global surface water quality model. *Geoscientific Model Development*, *16*(15), 4481–4500. https://doi.org/10.5194/gmd-16-4481-2023

Jung, H., Tetzlaff, D., Birkel, C., & Soulsby, C. (2025). Recent Developments and Emerging Challenges in Tracer-Aided Modeling. *Wiley Interdisciplinary Reviews Water*, *12*. https://doi.org/10.1002/wat2.70015

Kale, R., & Sahoo, B. (2011). Green-Ampt Infiltration Models for Varied Field Conditions: A Revisit. *Water Resources Management*, *25*, 3505–3536. https://doi.org/10.1007/s11269-011-9868-0

Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, *49*(1), 360–379. https://doi.org/10.1029/2012WR012195

Kuppel, S., Tetzlaff, D., Maneta, M., & Soulsby, C. (2018). EcH2O-iso 1.0: Water isotopes and age tracking in a process-based, distributed ecohydrological model. *Geoscientific Model Development*, *11*, 3045–3069. https://doi.org/10.5194/gmd-11-3045-2018

Landgraf, J., Tetzlaff, D., Birkel, C., Stevenson, J., & Soulsby, C. (2023). Assessing land use effects on ecohydrological partitioning in the critical zone through isotope-aided modelling. *Earth Surface Processes and Landforms*, *48*. https://doi.org/10.1002/esp.5691

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. (2010). Development and test of the HYPE (Hydrological Predictions for the Environment) model – A water quality model for different spatial scales. *Hydrology Research*, *41*. https://doi.org/10.2166/nh.2010.007

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. (2014). The two water worlds hypothesis: Ecohydrological separation of water between streams and trees? *Wiley Interdisciplinary Reviews: Water*, *1*. https://doi.org/10.1002/wat2.1027

McDonnell, J., & Beven, K. (2014). Debates on Water Resources: 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*. https://doi.org/10.1002/2013WR015141

Mcmillan, H., Araki, R., Bolotin, L., Kim, D.-H., Coxon, G., Clark, M., & Seibert, J. (2025). Global patterns in observed hydrologic processes. *Nature Water*, *3*. https://doi.org/10.1038/s44221-025-00407-w

Mikayilov, F., Rouholahnejad Freund, E., Ashraf Vaghefi, S., Srinivasan, R., Yang, H., & Klöve, B. (2015). A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology*, *524*. https://doi.org/10.1016/j.jhydrol.2015.03.027

Mu, Q., Zhao, M., & Running, S. (2011). Improvements to a MODIS Global Terrestrial Evapotranspiration Algorithm. *Remote Sensing of Environment*, *115*, 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019

Nan, Y., He, Z., Tian, F., Wei, Z., & Tian, L. (2021). Can we use precipitation isotope outputs of isotopic general circulation models to improve hydrological modeling in large mountainous catchments on the Tibetan Plateau? *Hydrology and Earth System Sciences*, *25*(12), 6151–6172. https://doi.org/10.5194/hess-25-6151-2021

Naz, B. S., Sharples, W., Ma, Y., Goergen, K., & Kollet, S. (2023). Continental-scale evaluation of a fully distributed coupled land surface and groundwater model, ParFlow-CLM (v3.6.0), over Europe. *Geoscientific Model Development*, *16*(6), 1617–1639. https://doi.org/10.5194/gmd-16-1617-2023

Neal, C., Christophersen, N., Neale, R., Smith, C., & Reynolds, B. (1988). Chloride in precipitation and streamwater for the upland catchment of River Severn, mid- Wales; some consequences for hydrochemical models. *Hydrological Processes*, *2*, 155–165. https://doi.org/10.1002/hyp.3360020206

Nearing, G., Ruddell, B., Bennett, A., Prieto, C., & Gupta, H. (2020). Does Information Theory Provide a New Paradigm for Earth Science? Hypothesis Testing. *Water Resources Research*, *56*. https://doi.org/10.1029/2019WR024918

Ocampo, C., Sivapalan, M., & Oldham, C. (2006). Hydrological connectivity of upland-riparian zones in agricultural catchments: Implications for runoff generation and nitrate transport. *Journal of Hydrology*, *331*, 643–658. https://doi.org/10.1016/j.jhydrol.2006.06.010

Pechlivanidis, I., Jackson, B., Mcintyre, N., & Wheater, H. (2011). Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *GlobalNEST International Journal*, *13*, 193–214.

Pesántez, J., Birkel, C., Gaona, G., Arciniega-Esparza, S., Murray, D., Mosquera, G., et al. (2023). Spatially distributed tracer-aided modelling to explore DOC dynamics, hot spots and hot moments in a tropical mountain catchment. *Hydrological Processes*, *37*, 1–15. https://doi.org/10.1002/hyp.15020

Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., et al. (2016). Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins. *Journal of Hydrometeorology*, *17*, 287–307. https://doi.org/10.1175/JHM-D-15-0054.1

Rakovec, O., Mizukami, N., Kumar, R., Newman, A., Thober, S., Wood, A., et al. (2019). Diagnostic Evaluation of Large-Domain Hydrologic Models Calibrated Across the Contiguous United States. *Journal of Geophysical Research: Atmospheres*, *124*. https://doi.org/10.1029/2019JD030767

Remondi, F., Kirchner, J., Burlando, P., & Fatichi, S. (2018). Water Flux Tracking with A Distributed Hydrological Model to Quantify Controls on the Spatio-Temporal Variability of Transit Time Distributions. *Water Resources Research*, *54*. https://doi.org/10.1002/2017WR021689

Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.*, *46*. https://doi.org/10.1029/2008WR007327

Seybold, E., Dwivedi, R., Musselman, K., Kincaid, D., Schroth, A., Classen, A., et al. (2022). Winter runoff events pose an unquantified continental-scale risk of high wintertime nutrient export. *Environmental Research Letters*, *17*, 104044. https://doi.org/10.1088/1748-9326/ac8be5

Šimůnek, J., Šejna, M., Saito, H., & Van Genuchten, M. T. (2013). *The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media (Version 4.17)*. Department of Environmental Sciences University of Riverside California, California.

Skrzypek, G., Mydłowski, A., Dogramaci, S., Hedley, P., Gibson, J. J., & Grierson, P. F. (2015). Estimation of evaporative loss based on the stable isotope composition of water using Hydrocalculator. *Journal of Hydrology*, *523*, 781–789. https://doi.org/10.1016/J.JHYDROL.2015.02.010

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*, 2239–2259. https://doi.org/10.5194/hess-25-2239-2021

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

Sprenger, M., Stumpp, C., Weiler, M., Aeschbach, W., Allen, S., Benettin, P., et al. (2019). The Demographics of Water: A Review of Water Ages in the Critical Zone. *Reviews of Geophysics*, *57*. https://doi.org/10.1029/2018RG000633

Tapley, B., Bettadpur, S., Ries, J. C., Thompson, P., & Watkins, M. (2004). GRACE Measurements of Mass Variability in the Earth System. *Science (New York, N.Y.)*, *305*, 503–5. https://doi.org/10.1126/science.1099192

Tetzlaff, D., Buttle, J., Carey, S. K., Mcguire, K., Laudon, H., & Soulsby, C. (2015). Tracer-based assessment of flow paths, storage and runoff generation in northern catchments: A review. *Hydrological Processes*, *29*(16), 3475–3490. https://doi.org/10.1002/HYP.10412

Wang, J., Bouwman, A., Vilmin, L., Beusen, A., Hoek, W., Liu, X., & Middelburg, J. (2024). Global inland-water nitrogen cycling has accelerated in the Anthropocene. *Nature Water*, *2*. https://doi.org/10.1038/s44221-024-00282-x

Wellen, C., Kamran-Disfani, A.-R., & Arhonditsis, G. (2015). Evaluation of the Current State of Distributed Watershed Nutrient Water Quality Modeling. *Environmental Science & Technology*, *49*. https://doi.org/10.1021/es5049557

Wen, Y., Lin, J.-S., Plaza, F., & Liang, X. (2024). Roles of Hydrology and Transport Processes in Denitrification at Watershed Scale. *Water Resources Research*, *60*, 1–24. https://doi.org/10.1029/2023WR034971

Winkler, K., Fuchs, R., Rounsevell, M., & Herold, M. (2021). Global land use changes are four times greater than previously estimated. *Nature Communications*, *12*(1), 2501. https://doi.org/10.1038/s41467-021-22702-2

Wu, S., Tetzlaff, D., Yang, X., & Soulsby, C. (2022). Disentangling the influence of landscape characteristics, hydroclimatic variability and land management on surface water NO_3-N dynamics: spatially distributed modelling over 30 years in a lowland mixed land use catchment. *Water Resources Research*, e2021WR030566. https://doi.org/10.1029/2021WR030566

Wu, S., Tetzlaff, D., Beven, K., & Soulsby, C. (2025a). DREAM(LoAX): Simultaneous Calibration and Diagnosis for Tracer-Aided Ecohydrological Models Under the Equifinality Thesis. *Water Resources Research*, *61*, e2024WR038779. https://doi.org/10.1029/2024WR038779

Wu, S., Tetzlaff, D., Yang, X., Sauter, T., & Soulsby, C. (2025b). Hydrological Connectivity Dominates NO3-N Cycling in Complex Landscapes – Insights From Integration of Isotopes and Water Quality Modeling. *Water Resources Research*, *61*, e2025WR040525. https://doi.org/10.1029/2025WR040525

Wu, S., Tetzlaff, D., & Soulsby, C. (2025c). Revising Common Approaches for Calibration: Insights From a 1-D Tracer-Aided Hydrological Model With High-Dimensional Parameters and Objectives. *Water Resources Research*, *61*, e2024WR037656. https://doi.org/10.1029/2024WR037656

Wu, S., Tetzlaff, D., Zheng, Y., & Soulsby, C. (2025d). EcoTWIN v1.0 release. https://doi.org/10.5281/zenodo.16747633

Yang, C., Condon, L., & Maxwell, R. (2025). Unravelling groundwater–stream connections over the continental United States. *Nature Water*, *3*. https://doi.org/10.1038/s44221-024-00366-8

Yang, X., Jomaa, S., Zink, M., Fleckenstein, J., Borchardt, D., & Rode, M. (2018). A New Fully Distributed Model of Nitrate Transport and Removal at Catchment Scale. *Water Resources Research*, *54*. https://doi.org/10.1029/2017WR022380

Yang, X., Zhang, X., Graeber, D., Hensley, R., Jarvie, H., Lorke, A., et al. (2023a). Large-stream nitrate retention patterns shift during droughts: Seasonal to sub-daily insights from high-frequency data-model fusion. *Water Research*, *243*, 120347. https://doi.org/10.1016/j.watres.2023.120347

Yang, X., Tetzlaff, D., Müller, C., Knöller, K., Borchardt, D., & Soulsby, C. (2023b). 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

Yang, X., Tetzlaff, D., Jin, J., Li, Q., Borchardt, D., & Soulsby, C. (2024). Linking terrestrial biogeochemical processes and water ages to catchment water quality: A new Damköhler analysis based on coupled modeling of isotope tracers and nitrate dynamics. *Water Research*, *262*, 122118. https://doi.org/10.1016/j.watres.2024.122118

Zhang, Z., Chen, X., Cheng, Q.-B., Li, S.-L., Yue, F.-J., Peng, T., et al. (2020). Coupled hydrological and biogeochemical modelling of nitrogen transport in the karst critical zone. *Science of The Total Environment*, *732*, 138902. https://doi.org/10.1016/j.scitotenv.2020.138902

Zhu, J., Jia, Y., Yu, G., Wang, Q., He, N., Chen, Z., et al. (2025). Changing patterns of global nitrogen deposition driven by socio-economic development. *Nature Communications*, *16*(1), 46. https://doi.org/10.1038/s41467-024-55606-y